

Analyzing Technology-Enhanced Learning Processes: What Can Process Mining Techniques Contribute to the Evaluation of Instructional Support?

Inaugural-Dissertation
zur Erlangung der Doktorwürde der
Fakultät für Humanwissenschaften
der
Julius-Maximilians-Universität Würzburg

Vorgelegt von
Christoph Sonnenberg
aus Coburg

2016

Betreuerin: Prof. Dr. Maria Bannert

Erstgutachter: Prof. Dr. Lynn Huestegge

Zweitgutachter: Prof. Dr. Joachim Wirth

Tag des Kolloquiums: 07.07.2017

Acknowledgments

The current dissertation would not have been possible without the professional and personal support of many people. At least, I want to mention the most important persons in the following.

First, many thanks to my supervisor, Prof. Dr. Maria Bannert, who supported me continuously throughout the emergence of this dissertation; from the initial ideas to the publication of my manuscripts. Additionally, she introduced me to the scientific community, which helped significantly to develop and pursue my research agenda.

Second, I would like to thank my former colleagues and research assistants from the Chair of Instructional Media, University of Würzburg, who contributed to a very productive and pleasant working atmosphere. Your door was always open for discussing professional and personal issues, and your support has always motivated me to cope with all challenging situations.

Third, I'm very thankful for having been part of the EARLI Special Interest Group 16 – Metacognition, which comprises many inspiring junior as well as senior researchers. Your feedback on my academic work helped to improve my research constantly and to shape my ideas to substantial manuscripts. Moreover, I want to thank the anonymous reviewers of my publications, whose valuable comments helped me a lot.

Finally, my sincere thanks to my parents and to my partner in life, Nina, who gave me support and motivation throughout all phases of this dissertation.

Gerbrunn, 24.11.2016

Christoph Sonnenberg

Part 1: Synopsis

Contents

1. Zusammenfassung	1
2. Summary	3
3. Research Agenda: In Search of Hidden Treasures	5
4. Fostering Technology-Enhanced Learning Through Instructional Support	7
4.1 The Challenges of Technology-Enhanced Learning	7
4.2 Self-Regulated Learning Models and Theoretical Process Assumptions	8
4.3 Sequential and Temporal Patterns – Self-Regulated Learning as Dynamic Event.....	9
4.4 Instructional Support Through Metacognitive Prompting.....	11
5. The Measurement of Learning Activities on the Micro Level	13
5.1 Overview: Assessment of Self-Regulated Learning and Data Channels.....	13
5.2 Online Measurement Using Think-Aloud Protocols	14
5.3 Analytical Challenges.....	15
6. Process Mining: Foundations and Application on Educational Data	17
6.1 Educational Data Mining and Learning Analytics	17
6.2 Process Mining Foundations.....	19
6.3 The Application of Process Mining.....	22
6.4 The Benefits of Process Mining for Technology-Enhanced Learning Issues.....	24
7. Summary of Findings: Process Mining Analyses	25
7.1 Analysis 1: Research Questions and Results	26
7.2 Analysis 2: Research Questions and Results	27
7.3 Analysis 3: Research Questions and Results	28
8. General Discussion of Findings	29
8.1 The Revealed Potential of Process Mining.....	30
8.2 Limitations of Process Mining.....	31
8.3 Lessons Learned - Recommendations How to Use Process Mining	32
8.4 Future Directions: A Quest for Standards and Guidelines.....	33
8.5 Conclusions: The Final Chapter?!	35
9. References	36
10. Appendix	44

1. Zusammenfassung

Die vorliegende Dissertation beschäftigt sich mit der Analyse technologieunterstützter Lernprozesse unter Verwendung von sogenannten *Process Mining* Methoden. Dabei werden kodierte Protokolle des lauten Denkens der Lerner als Prozessmaß genutzt, um eine Bewertung des Potentials dieses Analyseansatzes für die Evaluation der Effekte instruktionaler Hilfe vornehmen zu können.

Die zunehmende Verbreitung digitaler Medien in der Hochschulbildung und weiteren Ausbildungssektoren schafft neue Potentiale, allerdings auch neue Anforderungen an den Lerner, insbesondere an die Regulation seines Lernprozesses. Um den Lerner dabei zu unterstützen seinen Lernfortschritt optimal zu gestalten, wird ihm während des Lernens instruktionale Hilfe angeboten. Neben der Evaluation mittels Fragebögen und Testverfahren wird die Wirksamkeit der angebotenen Unterstützung zunehmend durch Prozessdaten bewertet. Die Analyse von beobachteten Verhaltensspuren während des Lernens (z.B. Logfiles, Blickbewegungen, Verbalprotokolle) ermöglicht einen detaillierten Einblick in die Lernhandlungen und die Folgen von Unterstützungsmaßnahmen auf den Lernprozess. Allerdings stellen sich auch eine Reihe von neuen analytischen Herausforderungen, wie der Umgang mit zeitlichen Dynamiken und Sequenzen von Lernhandlungen, insbesondere wenn man über Häufigkeitsanalysen der beobachteten Ereignisse hinausgehen möchte. Vor diesem Hintergrund beschäftigt sich die vorliegende Arbeit mit der Anwendung von *Process Mining* Methoden zur detaillierten Betrachtung von Lernprozessen. Insbesondere der Mehrwert dieses Ansatzes gegenüber einer reinen Häufigkeitsanalyse und somit die Potentiale von *Process Mining* für die Evaluation von Fördermaßen sollen herausgestellt werden.

Als Grundlage für die Bearbeitung der Fragestellung diente eine umfangreiche Laborstudie mit 70 Universitätsstudierenden, die durchgeführt wurde um die Effekte einer instruktionalen Fördermaßnahme zu prüfen. Die Probanden der Experimentalgruppe ($n = 35$) erhielten in einer 40-minütigen Hypermedia-Lernsitzung eine Förderung durch metakognitive Prompts, während die Kontrollgruppe ($n = 35$) ohne Hilfe lernte. In einer weiteren Lernsitzung drei Wochen später bearbeiteten alle Teilnehmer eine weitere Lerneinheit, diesmal ohne Unterstützung für alle Probanden. Während des Lernens wurden alle Teilnehmer instruiert, ihre Lernhandlungen kontinuierlich zu verbalisieren. Die kodierte Verbalprotokolle wurden in den folgenden drei Analysen dieser Dissertation detailliert mit Häufigkeits- und *Process Mining* Analysen untersucht.

Die erste Analyse konzentrierte sich auf den Vergleich der Lernhandlungen der Experimental- und Kontrollgruppe während der ersten Lernsitzung. Es wurde den Fragen nachgegangen, ob metakognitive Prompts die Lerner dazu anregen mehr metakognitive Lernhandlungen auszuführen, ob eine höhere Anzahl dieser Lernhandlungen mit dem Lernerfolg zusammenhängt (Mediation) und welche Unterschiede sich in den Abfolgen der Lernhandlungen finden lassen. In der zweiten Analyse wurden die Effekte der einzelnen Prompts sowie die Bedingungen für ihre Wirksamkeit auf einer sehr detaillierten Ebene betrachtet. Zusätzlich zu Process Mining wurde auch eine Data Mining Methode eingesetzt, um deren Befunde zu vergleichen. Im Detail fanden eine Klassifikation der Prompts anhand ihrer Effektivität und eine Untersuchung der kodierten Lernaktivitäten vor und nach der Präsentation instruktionaler Hilfe statt. Schließlich untersuchte die dritte Analyse die langfristigen Effekte metakognitiver Prompts auf den Lernprozess in einer weiteren Lernsitzung ohne Unterstützung. Hier stand die Frage im Mittelpunkt, welche geförderten Lernaktivitäten und Prozessmuster während der zweiten Lernsitzung stabil blieben. Wieder wurde sowohl eine Analyse anhand der Häufigkeiten der Lernaktivitäten als auch eine Process Mining Analyse eingesetzt, um deren Befunde vergleichen zu können.

Insgesamt belegen die Ergebnisse aller drei durchgeführten Analysen den Mehrwert von Process Mining im Vergleich zu reinen häufigkeitsbasierten Analysemethoden. Insbesondere unter Betrachtung des Lernprozesses als dynamische Abfolge von mehreren Lernhandlungen, ermöglicht Process Mining die Identifikation von Regulationsschleifen und zentralen Verzweigungen des Prozesses. Diese Befunde könnten zur Optimierung von Interventionsstrategien verwendet werden. Bevor aus den aufgedeckten Prozessmustern Schlussfolgerungen für die Gestaltung instruktionaler Hilfe gezogen werden können, müssen allerdings weitere Analysen erst noch die Generalisierbarkeit der Befunde untersuchen. Darüber hinaus bleibt die Anwendung von Process Mining herausfordernd, da derzeit keine Richtlinien für analytische Entscheidungen und Parametereinstellungen für technologieunterstützte Lernkontexte vorhanden sind. Darum müssen in Zukunft weitere Studien das Potential von Process Mining und verwandten Analysemethoden betrachten, um Forschern konkrete Anwendungsempfehlungen zur Verfügung stellen zu können. Generell kann Process Mining aber bereits jetzt dazu beitragen, das Verständnis der Auswirkungen instruktionaler Hilfe auf der Prozessebene voran zu treiben.

Schlüsselwörter: Selbstreguliertes Lernen, Metakognitives Prompting, Process Mining, Prozessanalyse, Lautes Denken

2. Summary

The current dissertation addresses the analysis of technology-enhanced learning processes by using so-called *Process Mining* techniques. For this purpose, students' coded think-aloud data served as the measurement of the learning process, in order to assess the potential of this analysis method for evaluating the impact of instructional support.

The increasing use of digital media in higher education and further educational sectors enables new potentials. However, it also poses new challenges to students, especially regarding the self-regulation of their learning process. To help students with optimally making progress towards their learning goals, instructional support is provided during learning. Besides the use of questionnaires and tests for the assessment of learning, researchers make use increasingly of process data to evaluate the effects of provided support. The analysis of observed behavioral traces while learning (e.g., log files, eye movements, verbal reports) allows detailed insights into the student's activities as well as the impact of interventions on the learning process. However, new analytical challenges emerge, especially when going beyond the analysis of pure frequencies of observed events. For example, the question how to deal with temporal dynamics and sequences of learning activities arises. Against this background, the current dissertation concentrates on the application of Process Mining techniques for the detailed analysis of learning processes. In particular, the focus is on the additional value of this approach in comparison to a frequency-based analysis, and therefore on the potential of Process Mining for the evaluation of instructional support.

An extensive laboratory study with 70 university students, which was conducted to investigate the impact of a support measure, served as the basis for pursuing the research agenda of this dissertation. Metacognitive prompts supported students in the experimental group ($n = 35$) during a 40-minute hypermedia learning session; whereas the control group ($n = 35$) received no support. Approximately three weeks later, all students participated in another learning session; however, this time all students learned without any help. The participants were instructed to verbalize their learning activities concurrently while learning. In the following three analyses of this dissertation, the coded think aloud data were examined in detail by using frequency-based methods as well as Process Mining techniques.

The first analysis addressed the comparison of the learning activities between the experimental and control groups during the first learning session. This study concentrated on the research questions whether metacognitive prompting increases the number of

metacognitive learning activities, whether a higher number of these learning activities corresponds with learning outcome (mediation), and which differences regarding the sequential structure of learning activities can be revealed. The second analysis investigated the impact of the individual prompts as well as the conditions of their effectiveness on the micro level. In addition to Process Mining, we used a data mining approach to compare the findings of both analysis methods. More specifically, we classified the prompts by their effectiveness, and we examined the learning activities preceding and following the presentation of instructional support. Finally, the third analysis considered the long-term effects of metacognitive prompting on the learning process during another learning session without support. It was the key objective of this study to examine which fostered learning activities and process patterns remained stable during the second learning session. Again, we conducted a frequency-based analysis as well as Process Mining to compare the results of both approaches.

Overall, all three analyses indicated the additional value of Process Mining in comparison to a frequency-based analysis. Especially when conceptualizing the learning process as a dynamic sequence of multiple activities, Process Mining allows identifying regulatory loops and crucial routing points of the process. These findings might contribute to optimizing intervention strategies. However, before drawing conclusions for the design of instructional support based on the revealed process patterns, additional analyses need to investigate the generalizability of results. Moreover, the application of Process Mining remains challenging because guidelines for analytical decisions and parameter settings in technology-enhanced learning context are currently missing. Therefore, future studies need to examine further the potential of Process Mining as well as related analysis methods to provide researchers with concrete recommendations for use. Nevertheless, the application of Process Mining techniques can already contribute to advance the understanding of the impact of instructional support through the use of fine-grained process data.

Keywords: Self-Regulated Learning, Metacognitive Prompting, Process Mining, Process Analysis, Think-Aloud Data

3. Research Agenda: In Search of Hidden Treasures

Technology-enhanced learning (TEL) increasingly gains in importance for education in the 21st century (Johnson et al., 2016). For example, digital learning environments which are based on intelligent tutor systems, hypermedia learning systems, or computer-supported cooperative learning scenarios characterize learning and instruction in schools and universities. In comparison to traditional learning settings (e.g., teacher-centered classroom instruction), TEL enables new opportunities and potentials from a constructivist approach (Chi & Wylie, 2014; Lawless & Brown, 1997). In particular, these potentials comprise a greater extent of students' autonomy, that is, learning becomes more independent from time and place, and a student has a higher personal responsibility for his or her learning pathways. However, new demands on the individual student come along with the evolving TEL settings. Therefore, researchers and practitioners aim to design, apply, and evaluate various types of instructional support that help students to use the available potentials better. Especially the promotion of self-regulatory competencies and the application of metacognitive knowledge plays a crucial role in the enhancement of TEL processes (e.g., Azevedo, Guthrie, & Seibert, 2004; Bannert, 2007). Metacognition can be defined recursively as cognition about cognition; comprising the function to regulate one's cognition (Flavell, 1979). To better design instructional support and to strive for personalized and adaptive interventions (e.g., Azevedo, Cromley, Moos, Greene, & Winters, 2011; Walker, Rummel, & Koedinger, 2009), the recent research is interested in investigating the student's learning process on a very detailed level. One major issue here is the consideration of how learning unfolds over time and the analysis of the sequential and temporal structure of learning processes (e.g., Molenaar & Järvelä, 2014), as well as the impact of instructional support on these phenomena.

The increasing significance of analyzing TEL processes, particularly the deployment of self-regulatory activities during learning, is evidenced by several recent special issues in the journals of various educational research communities; namely, *Learning Science* (Martin & Sherin, 2013), *Educational Data Mining* (Winne & Baker, 2013), *Metacognition and Learning* (Ben-Eliyahu & Bernacki, 2015; Molenaar & Järvelä, 2014), and *Learning Analytics* (Roll & Winne, 2015). Many contributions of these special issues conceptualize learning as a dynamic interplay of multiple learning activities, and highlight the importance of assessing and analyzing fine-grained traces of students' behavior to understand the impact of instructional support. TEL facilitates the measurement of learning behavior, which can be

recorded in real time on different data channels nowadays (e.g., Azevedo et al., 2013). Nevertheless, new analytical challenges arise when dealing with these fine-grained and large volumes of data. Therefore, current studies in the mentioned special issues address the development, application, and evaluation of innovative analysis approaches, which might help researchers to gain deeper insights into the learning process and its dynamics.

The following analogy, which describes a problem-solving scenario, illustrates the current research efforts to analyze TEL processes. Let us imagine a suspected treasure that lies at the very bottom of the ocean. Adventurers attempt to retrieve this treasure; however, they have a major lack of equipment, which impedes their progress. Moreover, there is no guarantee whether the expensive and time-consuming retrieval is worth the effort, because they only have a blurred picture of the hidden treasure and it might be valueless. Consequently, the adventurers' main concern is to investigate which method is most suited for their endeavor. Of course, there might be several approaches that will result in their goal attainment, the retrieval of the hidden treasure. Dependent on the treasure's characteristics and its location, it might be useful to search for the most effective techniques. Additionally, they first might be interested in using methods that help to visualize the treasure better to decide if it is worth the retrieval and which retrieval approach might be appropriate. Finally, the adventurers have to assess if their individual efforts are sufficient, or if it might be better to start a joint venture with others.

Following this analogy of searching for hidden treasures, the current dissertation contributes to the existing literature by examining the value of an analytical approach called *Process Mining* (PM) for the evaluation of instructional support in TEL settings. PM techniques were applied in the context of an empirical study that was conducted to investigate the effects of so-called metacognitive prompts on self-regulated hypermedia learning (Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015). Although we used a specific setting and a specific type of instructional support, the findings on the contribution of PM might be generalizable to additional TEL contexts. The potential of PM for discovering and testing process patterns as well as the impact of instructional support on these patterns is demonstrated and discussed by three analyses. The findings are compared to traditional frequency-based analyses of learning activities, which still represent the standard in many studies.

The current synopsis is structured as follows. Chapters 4 to 6 summarize the relevant theoretical background for this dissertation. These sections comprise literature on (i) process models of self-regulated learning and instructional support through metacognitive

prompting, (ii) the fine-grained measurement of learning activities, and (iii) the PM approach. Then, Chapter 7 reports the research objectives and findings of each PM analysis. Finally, Chapter 8 concludes this synopsis with a general discussion about the significance of PM for the evaluation of instructional support and draws implications for future directions.

4. Fostering Technology-Enhanced Learning Through Instructional Support

The present dissertation examines the application of PM techniques in the context of TEL and instructional support. Therefore, this chapter presents the theoretical background on these phenomena. First, the challenges of TEL, which students have to face, are presented in Chapter 4.1. Then, the Chapters 4.2 and 4.3 refer to the key assumptions of self-regulated learning (SRL) models and the current perspective of researchers on regulatory processes as a dynamic sequence of events. Finally, one type of instructional support, namely metacognitive prompting, is introduced in Chapter 4.4, whose impact on learning processes was investigated in the analyses of this dissertation.

4.1 The Challenges of Technology-Enhanced Learning

Technology enables new opportunities and potentials for teaching and learning, such as the facilitated distribution of material through online systems, the use of multiple representations and multimedia, and adaptive learning environments (Johnson et al., 2016; Mayer, 2009). However, TEL also confronts students with new challenges, especially in comparison to traditional learning scenarios (e.g., classroom teaching with a high external regulation). Let us imagine a typical learning situation in which students work with topic-specific, computer-presented hypermedia material for a specified period with a particular learning task (e.g., Bannert & Reimann, 2012). Such a scenario is representative of a broad class of interactive information activities (Reimann, Markauskaite, & Bannert, 2014). To successfully master this scenario, students constantly have to make decisions on what to do, where to go next, and to evaluate the retrieved information on their current learning goals (Schnitz, 1998). Moreover, the awareness and control of their manner of learning, or in broader terms the deployment of metacognitive skills and the attempt to regulate their learning actively, plays a key role in TEL and open-ended learning tasks (Azevedo, 2005; Lin, 2001; Lin & Lehman, 1999).

Empirical evidence indicates that the deployment of self-regulatory skills represents an essential prerequisite for successful learning (e.g., Winne & Hadwin, 2008; Zimmerman,

2008). However, studies also demonstrate that learners often show no spontaneous use of metacognitive competencies during learning, which leads to poorer learning outcomes (Azevedo, 2009; Bannert & Mengelkamp, 2013; Greene, Dellinger, Tüysüzoglu, & Costa, 2013). Therefore, it is one major challenge in TEL for educators to train or to activate the students' repertoire of self-regulatory skills. The components and phases of successful regulation are described in SRL models, which are summarized in the following section.

4.2 Self-Regulated Learning Models and Theoretical Process Assumptions

A variety of research that investigates learning in traditional but also in TEL settings builds upon SRL models. As argued above, the regulation of one's learning plays a fundamental role to meet the challenges of the learning task, particularly while being engaged with TEL. SRL models describe the characteristics of successful learning; often in terms of an ideal-typical learning process. In general, these models emphasize an active performance of cognitive, metacognitive, and motivational learning activities, as well as a dynamic interplay of these activities to achieve one's learning goals (Boekaerts, 1997; Schmitz & Wiese, 2006; Winne & Hadwin, 2008; Zimmerman, 2008). The assumptions in SRL models anticipate a time-ordered sequence of activities, whereby no strict order is pre-determined (Azevedo, 2009). Often, three cyclical phases of forethought, performance, and reflection are considered (e.g., Zimmerman, 2008). The COPES model (Winne & Hadwin, 2008) represents the most elaborate description in terms of an information-processing model, which comprises the four phases *task definition, goal setting and planning, studying tactics, and adaptations to metacognition*. Additionally, it considers monitoring and control as key elements of regulated learning. Moreover, SRL models also comprise assumptions concerning the transfer of learning experiences to additional tasks in the future. For example, the successful use of a learning strategy should affect the learning process in similar contexts. Furthermore, this assumption kindles the interest in examining the sustainability and transfer of strategies, which were fostered through instructional support (see analysis 3; Sonnenberg & Bannert, submitted). Empirical findings confirmed that successful learning corresponds with the active deployment of the activities described in SRL models (e.g., Azevedo et al., 2004; Bannert, 2009; Johnson, Azevedo, & D'Mello, 2011; Moos & Azevedo, 2009), but few studies have addressed the stability of SRL between several learning tasks and contexts (e.g., Moos & Miller, 2015).

Bannert (2007) proposed a theoretical framework that describes the learning process during hypermedia learning and that refers to the learning activities in SRL models. In the analyses of this dissertation, this framework represents the basis for the measurement of learning activities and characterizes the coding scheme for analyzing the think-aloud data. Figure 1 shows the determinants and learning activities that affect the learning process and performance during hypermedia learning, according to Bannert's framework. It comprises an ideal sequence of *orientation, planning, goal specification, information search and relevance judgment, information processing, and evaluation of goal attainment*, which is constantly monitored and controlled. However, considering the challenges of a given task, the performance of these activities might be more dynamic. Moreover, the determinants of the assumed learning process are learner characteristics (e.g., prior knowledge) and conditions of the learning environment (e.g., task characteristics).

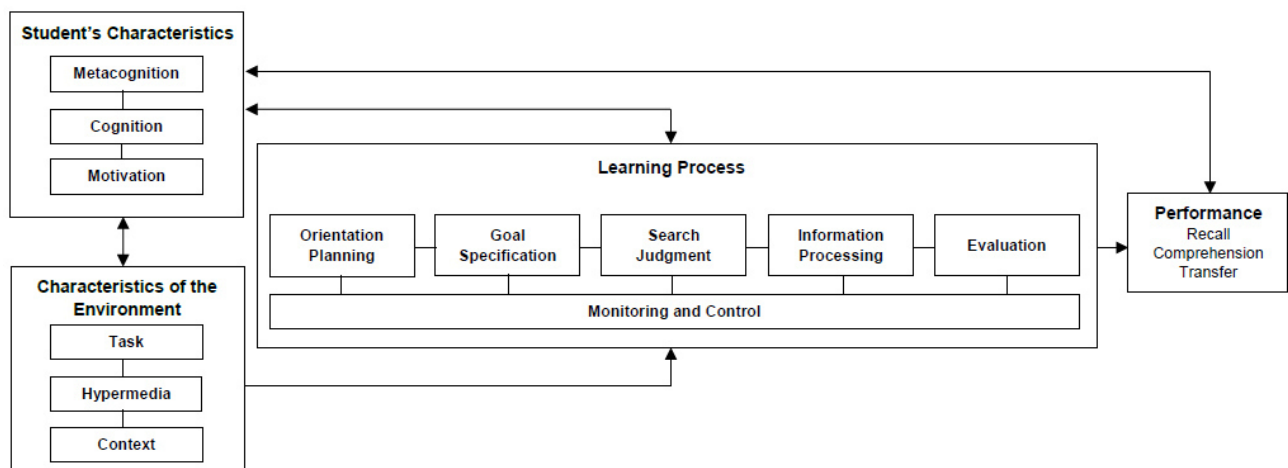


Figure 1. Bannert's proposed framework for self-regulated hypermedia learning (adapted from Bannert, 2007, p. 72).

Recent research also highlights the significance of motivational and emotional processes as well as their measurement during TEL (e.g., Azevedo, 2015; Ben-Eliyahu & Linnenbrink-Garcia, 2015), but the present dissertation focuses on the interplay of cognitive and metacognitive learning activities. For instance, Azevedo and colleagues (2013) introduced the conceptualization of *Cognitive, Affective, Motivational, and Metacognitive Processes* (CAMP) for measuring and assessing learning processes.

4.3 Sequential and Temporal Patterns – Self-Regulated Learning as Dynamic Event

Over the past years, the researchers' perspective has shifted from SRL being an aptitude to a process-orientated view that explains differences among learners with respect to regularities and patterns in the performed learning events (Winne & Perry, 2005). SRL

models such as Zimmerman's phase model (Zimmerman, 2008) and the COPES model (Winne & Hadwin, 2008) already incorporate the process-orientated view by describing SRL as a dynamic interplay of events during learning. Based on this theoretical assumption, researchers increasingly investigate regulatory activities as dynamically unfolding over time during a learning task, particularly focusing on the discovery of sequential and temporal patterns that affect learning performance (Azevedo, 2009; Winne, 2014). Consequently, the measurement of SRL also shifted from questionnaire methods to the recording of directly observable traces of students' behavior (see Chapter 5). For example, according to the COPES model, the learner passes specific states, which correspond to behavior that can be observed as his or her actions or utterances during learning (Reimann et al., 2014). Because the learning process is in general strongly affected by situational cues and demands, the perspective of SRL as an aptitude, which is measured by questionnaires, represents a too static construct.

Moreover, Reimann (2009) compared the variable-centered with the event-centered approach for considering a learning process and drew a conclusion for the investigation of time factors. Although his article concentrates on computer-supported collaborative learning (CSCL) research, the conceptualization is also applicable to other contexts, such as SRL (Bannert, Reimann, & Sonnenberg, 2014). Within the variable-centered perspective, a process represents a set of concepts that mediate between independent and dependent variables. For instance, these concepts could be frequencies of performed learning activities. In this case, the procedure for process analysis would be a coding and counting of the students' activities and the application of a statistical method for the analysis of variance (e.g., see the mediation approach in analysis 1; Sonnenberg & Bannert, 2015). However, the variable-centered approach assumes that the independent variables act continuously on the dependent variables, and it cannot accommodate qualitative changes in the system of variables. Because this perspective does not sufficiently reflect the theoretical concept of a dynamic learning process that unfolds over time, Reimann proposed the event-centered approach as well as a combination of both approaches. Figure 2 illustrates the essence of these approaches. The event-centered approach assumes a discrete event system that describes the course of a learning process from an initial state to a resulting state. This system itself might change over time (e.g., during a semester) or through an intervention (e.g., the presentation of a scaffold). Referring to the framework for self-regulated hypermedia learning described above, the factors displayed in Figure 2 could be internal or external determinants such as prior knowledge, task characteristics, or instructional support, which

affect the system of events during learning. Furthermore, the interplay of cognitive and metacognitive events affects outcome variables such as learning performance.

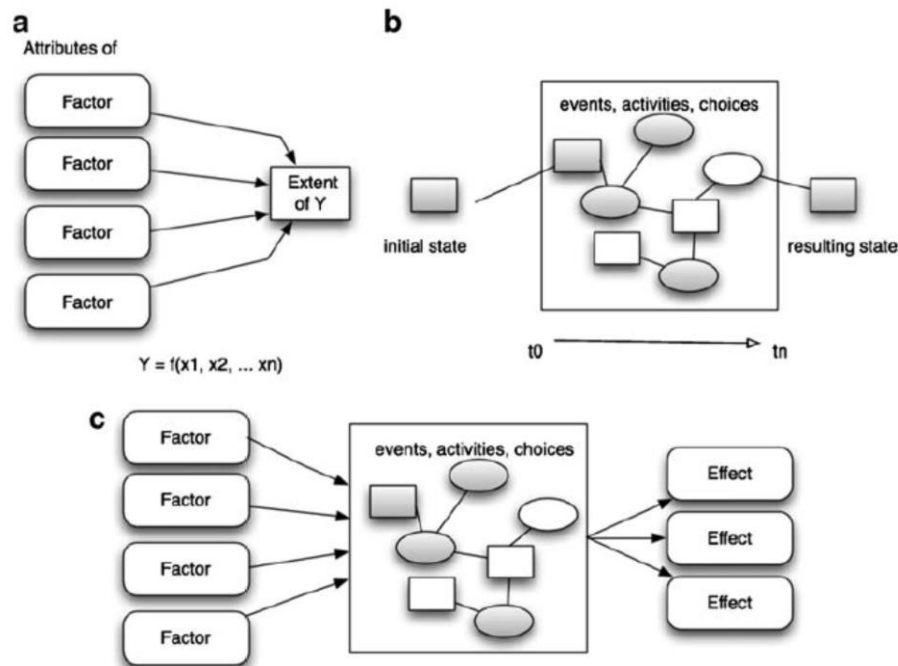


Figure 2. Variable- versus event-based approach (from Reimann, 2009, p. 243). a = variable perspective, b = event perspective, and c = combination of both approaches.

In summary, the process-orientated and event-centered perspective presented above increased the researchers' interest in investigating how the learning process unfolds over time and how scaffolds influence the dynamic nature of regulatory activities. However, the traditional canon of methods from the social science is not sufficient for issues that come from the event-based perspective. Therefore, recent special issues present methodological contributions to the analysis of time and order in learning activities (Martin & Sherin, 2013; Molenaar & Järvelä, 2014). For example, the effects of instructional support on learning activities and their sequential structure could be analyzed on the micro level to enable a precise evaluation and to optimize supporting strategies (e.g., Jeong et al., 2008; Johnson et al., 2011).

4.4 Instructional Support Through Metacognitive Prompting

In general, instructional support attempts to counteract the students' learning difficulties by fostering strategic learning processes. For this purpose, educators make use of different types of scaffolding techniques in traditional as well as TEL settings. The concept of scaffolding (Puntambekar & Hübscher, 2005; Wood, Bruner, & Ross, 1976) comprises the support of desired behavior until a student advances to the scaffolded

activities. Then, a teacher or an educational technology withdraws the support. Ideally, a student also transfers the fostered behavior into new contexts. Detailed information about the students' learning process is needed to diagnose when to reduce support.

One type of instructional support are prompts, which can be defined as scaffolds that induce and stimulate students' cognitive, metacognitive, and motivational activities during learning (Bannert, 2009). They are based on the assumption of a production deficit, that is, students show no spontaneous recall or execution of already acquired processes (e.g., Winne, 1996; Wirth, 2009). Because metacognition represents a key role in SRL models, especially monitoring and controlling one's learning, especially interventions that focus on metacognitive support, such as metacognitive prompts, have the potential to foster students' successful learning (Bannert & Reimann, 2012; Künsting, Kempf, & Wirth, 2013).

Metacognitive prompts attempt to activate the students' repertoire of metacognitive knowledge and learning strategies by requesting them to reflect, monitor, and control their learning process (Bannert, 2007, 2009; Veenman, 1993). They focus students' attention on their thoughts and on understanding the activities in which they are engaged in during learning. Ideally, the prompted requests induce SRL activities such as orientation, goal specification, planning, monitoring and control, and evaluation strategies. It is expected that prompting increases the quantity of these regulatory activities, but they might also affect the sequential order of events during a learning task.

A robust body of research indicates that metacognitive support has beneficial effects on TEL (Devolder, van Braak, & Tondeur, 2012; Zheng, 2016). With respect to metacognitive prompting, empirical findings showed positive effects on learning in different domains and settings, such as hypermedia learning (Azevedo et al., 2011; Bannert & Mengelkamp, 2013; Bannert et al., 2015), writing learning journals (Hübner, Nückles, & Renkl, 2010; Nückles, Hübner, Dümer, & Renkl, 2010), and additional settings (Künsting et al., 2013; Thillmann, Künsting, Wirth, & Leutner, 2009). In general, research has investigated and evidenced the quantitative increase of regulatory processes during learning, but there is also initial evidence that prompting affects temporal dependencies among SRL activities (Johnson et al., 2011).

Despite these promising results of prompting effects, the design and implementation of prompts during learning in open-ended environments remains challenging (Azevedo & Hadwin, 2005). More research is necessary to address the issues of how to determine the presentation times of support optimally, how to calibrate support for the appropriate phase of SRL, and how to gradually withdrawal support. For instance, several studies indicated

that some students still show a poor compliance with provided support, and consequently, they do not benefit as intended by educators (Bannert & Mengelkamp, 2013; Clarebout & Elen, 2006). Therefore, instructional support would benefit from detailed analyses that take into account the learning activities on the micro level, and that provide implications for optimizing a supporting strategy. Moreover, because TEL research strives for adaptive support and real-time interventions, process analyses are needed as the fundament for developing and refining student models and production rules that control the presentation of support in digital learning environments (e.g., Baker & Corbett, 2014; Bouchet, Harley, Trevors, & Azevedo, 2013; Molenaar & Roda, 2008). In conclusion, analysis methods that are capable of evaluating the specific effect of scaffolds by taking into account a process-orientated view, allow to optimize instructional support (e.g., Jeong et al., 2008; Johnson et al., 2011; Molenaar & Chiu, 2014), and they might provide valuable information for the development of SRL theories on the micro level (Molenaar & Järvelä, 2014).

5. The Measurement of Learning Activities on the Micro Level

The investigation of TEL processes with respect to the temporal dynamics of sequences of events requires the measurement of learning activities on a very detailed level. First, Chapter 5.1 provides an overview of SRL measurement and the available data channels. Then, Chapter 5.2 presents the assessment method used in this dissertation, namely concurrent think-aloud protocols. Additionally, Chapter 5.3 discusses the challenges that arise from the analysis of fine-grained process data.

5.1 Overview: Assessment of Self-Regulated Learning and Data Channels

Dependent on the theoretical understanding of a learning process, researchers who investigate SRL make use of measures of aptitudes (e.g., questionnaires) or behavioral process data (Azevedo, 2009, 2015; Bannert, 2009; Veenman, Van Hout-Wolters, & Afflerbach, 2006). This difference resulted in a general classification into *offline* and *online* measures (Veenman, 2005; Wirth & Leutner, 2008). Because the current dissertation focuses on SRL process models, the measurement of observable behavioral traces plays a key role, rather than students' solid abilities. Based on the conceptualization of SRL as being a dynamic sequence of events, researchers need to measure the students' activities online during learning and use a granularity that is appropriate for the research objectives. Moreover, they have to identify observable indicators for specific SRL processes as well as what might determine successful learning. Furthermore, technological advances kindled the

proliferation of fine-grained process data because digital learning environments allow a very detailed and largely unobtrusive recording of learning-related behavior (Azevedo et al., 2013; Winne & Nesbit, 2009). SRL behavior is often measured using online trace methods such as concurrent think-aloud protocols or computer log files (Greene & Azevedo, 2010), but recently, the research experiences a vivid extension of data channels such as eye-tracking (e.g., Trevors, Feyzi-Behnagh, Azevedo, & Bouchet, 2016), physiological parameters (e.g., Azevedo et al., 2013), and neuronal correlates (e.g., De Smedt, 2014).

Another possible classification of assessment methods refers to the level of data granularity. In general, this classification considers the number of individual data points, that is, the measurement unit, as well as the time scale of observation. For instance, researchers might investigate learning by considering macro-level units (e.g., the aggregated number of all learning strategies that were used) or more detailed micro-level units (e.g., a sequence of multiple events that represents various cognitive, metacognitive, and motivational learning activities) (Azevedo, 2015; Dent & Hoyle, 2015). Furthermore, with respect to the time scale, learning might be observed during a short learning episode (e.g., 30 minutes) or a longer period (e.g., during a semester). Additionally, events that reflect human thought might occur over seconds, minutes to hours, or weeks to months (Ben-Eliyahu & Bernacki, 2015).

In general, using fine-grained process data on different data channels allows researchers to analyze learning on the micro level, and to test assumptions regarding the temporal dynamics of TEL. The analyses of this dissertation make use of concurrent think-aloud protocols, which were collected during two 40-minute learning episodes. The following section addresses the online measurement of SRL activities through verbal reports in more detail.

5.2 Online Measurement Using Think-Aloud Protocols

Concurrent think-aloud protocols, also known as verbal reports, represent an online trace method that is frequently used in SRL settings, mainly because it allows a valuable access to the learning events performed during learning (Azevedo, Moos, Johnson, & Chauncey, 2010). The think-aloud technique is based on the work of Ericsson and Simon (1993). According to them, the conscious thoughts stored in the short-term memory can be verbalized, and therefore they are observable for researchers. In this dissertation, students were instructed to verbalize every thought that comes to their minds, without any interpretation or justification. These instructions to think aloud represent level 2 verbalizations; consequently, the technique should not affect the processes of human thought

(i.e., reactivity). For example, Bannert and Mengelkamp (2008) found evidence that thinking-aloud during hypermedia learning did not affect the learning performance. Additionally, a meta-analysis from Fox, Ericsson, and Best (2011) indicated that the think-aloud procedure is nonreactive, but that times to complete a task were increased for participants who had to verbalize their thoughts.

Although think-aloud protocols are not unobtrusive for the learner, such as other online trace methods (e.g., computer log files), they provide a detailed trace of learning activities that is appropriate for investigating the dynamics of TEL and the impact of instructional support on these processes (Azevedo et al., 2010; Schraw, 2010; Veenman, Bavelaar, De Wolf, & Van Haaren, 2014). Moreover, the technique is helpful for the identification of indicators of successful learning as well as theory-building. Dependent on the coding of the think-aloud data, the granularity of events might correspond more directly to the level of theory formulation than other data channels. However, for practical purposes such as real-time interventions in digital learning environments other data channels like interaction logs are needed. Nevertheless, computer log files imply a more difficult interpretation of awareness and intent regarding the students' actions than verbalizations.

In general, the coding of verbal data is a necessary, but time-consuming preparation step before data analysis. For the current dissertation, we followed the procedure that was recommended by Chi (1997). The coding was based on an original scheme from Bannert (2007), which distinguishes between four main categories (i.e., metacognitive, cognitive, motivational, and residual) and several subcategories. The coding scheme builds upon the framework presented above (see Figure 1). Please see Appendix 1 for all categories and descriptions of the original coding scheme. With respect to the research questions and analysis methods of this dissertation, we used aggregated versions of the original coding scheme and we did not consider the valence of events (e.g., the successful or unsuccessful result of monitoring). Reasons for the aggregation of codes were the low frequencies of some categories, the orientation along the granularity in SRL models, and the properties of a PM technique as explained in analysis 1 (Sonnenberg & Bannert, 2015).

5.3 Analytical Challenges

Besides the benefits of the detailed multichannel measurement of learning, new analytical challenges accompany the availability of rich behavioral traces. Especially while examining more than frequency distributions of learning activities, researchers have to address emerging analytical as well as conceptual challenges, such as the questions how to

analyze the temporal dynamics of learning and how to assess quantitative and qualitative effects of instructional support on the structure of learning processes. To address the dynamic sequence of self-regulatory activities during learning and its relationship to learning outcome, advanced analysis approaches are needed, which are appropriate for the conceptualization of SRL as a sequence of multiple events and which can handle the collected large volumes of data (e.g., data mining approaches). Recent special issues comprise contributions that rise to these challenges by presenting exploratory approaches to identify process patterns from SRL data (Ben-Eliyahu & Bernacki, 2015; Molenaar & Järvelä, 2014). Because these analyses assess and model learning in a specific setting or for a specific sample of students, a shift towards the application of confirmatory testing and the validation of resulting process patterns is also necessary in future research (Roll & Winne, 2015; Winne, 2014).

Researchers should consider the following key challenges while measuring and examining fine-grained traces of learning behavior. Winne and Baker (2013) indicated the importance of taking into account how reliable a measurement instrument is for the collection of robust data (i.e., minimal variation when there is no change of state) and the purity of the measured data (i.e., the ratio of signal to noise). Similarly, Ben-Eliyahu and Bernacki (2015) emphasized the challenge of measuring data that are sufficiently precise, but also comprehensive to allow complex statistical analyses. However, instead of just focusing on so-called big data, researchers must assess if the collected data basis is meaningful for their objectives. Otherwise, the statement from computer science “garbage in – garbage out” might become true, when applying a data mining method. Although the application of data mining and machine learning techniques is also dependent on data quality, these methods mainly rely on the availability of large samples. However, educational research often deals with small sample sizes (Dent & Hoyle, 2015).

Moreover, the measurement of authentic learning behavior results in the presence of noisy data because the real world is subjected to stochastic principles (Roll & Winne, 2015). Consequently, analysis methods must deal with indeterminacy and probabilities. Furthermore, data mining approaches that are applied to educational data must account for the multi-level hierarchy and non-independence of measurement units (Baker, 2010). When addressing characteristics of temporality, another important conceptual issue is to consider quantitative terms (e.g., durations, or rates of change) as well as the qualitative structure of behavior (e.g., the relative positioning of actions) (Reimann et al., 2014). Moreover, the consistency of patterns between participants must be addressed when analyzing the learning

activities of a sample. Assuming that a sample of students shows a great extent of variability in their course of learning, the aggregation of all learning processes into a common model might not be appropriate (e.g., Bannert et al., 2014), and a pre-selection of cases or a clustering method might be useful in advance. Finally, in the case of multichannel data, the challenge of the temporal alignment of different channels arise (Azevedo, 2014), along with the issue how SRL constructs appear on the various behavioral levels (e.g., what is an indicator of a monitoring activity in verbal reports versus in computer log files).

In summary, researchers need to identify the markers of successful SRL within their measured data with respect to their research objectives. Furthermore, they must decide which analysis method might be appropriate by taking into account the data quality and analytical challenges of fine-grained data traces. The current dissertation cannot address all the mentioned challenges, but it investigates an analysis method from the field of educational data mining, which might contribute to advance the analysis of fine-grained SRL data. The following chapter presents the approach of PM and discusses how PM techniques address some of the presented key challenges.

6. Process Mining: Foundations and Application on Educational Data

The analysis techniques presented in this dissertation were used in the context of Educational Data Mining (EDM) and Learning Analytics (LA). First, Chapter 6.1 gives a short overview of these fields and their potential for evaluating instructional support. Second, Chapter 6.2 summarizes the foundations of PM, particularly the main concepts and functions for analyzing sequences of events. Third, Chapter 6.3 presents how PM techniques can be applied to the process data that was introduced in the previous chapter. Finally, the potential of using PM to investigate research questions in SRL settings is discussed in Chapter 6.4.

6.1 Educational Data Mining and Learning Analytics

EDM and LA are two emerging disciplines that are fueled by the increasing availability of learners' digital traces through the use of TEL environments. Both disciplines pursue the main objective of improving and personalizing education by using data-intensive approaches (Baker & Inventado, 2014). EDM has a greater focus on the development and application of computational techniques that are suited to work with large-scale educational data sets (Romero, Ventura, Pechenizkiy, & Baker, 2010); whereas LA highlights the significance of visualization and human interpretation of data (Baker & Inventado, 2014),

and considers a wider range of stakeholders as well as the application in instructional systems (Larusson & White, 2014; Martin & Sherin, 2013). However, despite these differences in their point of view, both disciplines use a very similar range of analysis methods. In the past years, the impact of the EDM and LA methods on educational research and practice has steadily grown (Baker & Yacef, 2009; Papamitsiou & Economides, 2014; Romero & Ventura, 2010).

In general, the range of methods comprises statistical, machine-learning, and data-mining algorithms, which are applied to different types of educational data to understand learners and the settings in which they study better (Romero & Ventura, 2010). Referring to a classification of methods proposed by Baker (2010), the key functions of EDM techniques are (i) prediction (e.g., regression and classification), (ii) clustering, (iii) relationship mining (e.g., sequential pattern mining), and (iv) discovery with models (e.g., Bayesian networks). The last function comprises the automated discovery of a student model that can be validated and used in additional analyses, and it has become increasingly popular for the investigation of more complex learning behavior (e.g., Baker & Corbett, 2014; Jeong et al., 2008). The approach of PM, described below, falls into this category of EDM techniques. Discovery with models allows to operationalize and identify specific behaviors, for example, learning activities described in SRL theories, and to analyze learning as it unfolds over time (Winne & Baker, 2013). Through these features, EDM techniques have the potential to advance the discovery of event patterns in SRL. Furthermore, they allow the precise modeling of robust learning and the impact of scaffolds in technology-enhanced settings (Baker & Corbett, 2014). For example, researchers can evaluate the impact of pedagogical support on the improvement of models that represent information about student's knowledge, motivation, metacognition, and attitudes (Baker & Yacef, 2009). Moreover, EDM techniques applied on fine-grained data from SRL settings might advance the understanding of the sequential and temporal characteristics (Martin & Sherin, 2013; Molenaar & Järvelä, 2014) and the dynamic relationship between SRL processes (Ben-Eliyahu & Bernacki, 2015).

Despite the high potential of the analysis methods used in the fields of EDM and LA for gaining deeper insights into TEL, the general challenges of data mining and machine learning during data analysis and interpretation of results need to be considered (Aggarwal, 2015). For example, because of an inductive approach that identifies patterns based on present instances, an overfitting of the resulting model might occur. Consequently, multi-level cross-validation is necessary to determine the extent of model generalizability; at least if it is the researcher's goal to predict the behavior of additional students (Winne & Baker,

2013). In the second analysis of this dissertation (Sonnenberg & Bannert, 2016), we use a cross-validation on the prompt-level as well as on the student-level to measure the generalizability of a learned linear regression model. Although EDM and LA techniques allow researchers to take into account large volumes of educational data, Martin and Sherin (2013) point out that there should be no restriction to data that can easily be collected using online systems (e.g., key press and mouse click data), but the analysis methods are also applicable to more traditional educational data. The selection of data channels and time windows for observed behavior represents another challenge for data mining methods. The research objectives must be considered for those decisions, but not vice versa, that is, data is taken into account just because they are available. Therefore, the analyses presented in the following chapter make use of a data channel whose granularity matches the assumptions in SRL models, namely concurrent think-aloud protocols, rather than taking into account computer log files. In the following, an approach is introduced that can be used in the field of EDM and LA to model sequences of learning events. PM techniques are applied and evaluated in the present dissertation because of the advantages described in the following sections.

6.2 Process Mining Foundations

The origin of PM lies in computer science and business IT, and it is an approach that uses temporally ordered event data to model the underlying process (Bannert et al., 2014; van der Aalst, 2011). Referring to the event-based view of SRL presented above, PM is compatible with the conceptualization of learning as a dynamic sequence of activities. Therefore, a process model that is derived from the student's behavioral traces while using a TEL environment can describe his or her learning process. PM builds on the concept of Petri nets (Reisig, 1985), which represent a discrete event system that comprises places and transitions. Petri nets are formal models with an explicit modeling language and executable semantics, and therefore they are deterministic. To extend these properties to stochastic aspects (e.g., transition probabilities), PM uses variants with weaker semantics and heuristics.

Regarding the scope of functions, PM allows researchers to discover process models inductively from event-based learning activities, to test models by conformance checking with additional data, and to extend existing models (Trčka, Pechenizkiy, & van der Aalst, 2010; van der Aalst, 2011). These functions, as well as the general concept of PM in educational settings, are illustrated in Figure 3. Observed behavioral traces are stored in an

event log. In general, learning activities can be recorded through an electronic system, but in our case, we used the think-aloud procedure to measure and code events. The event log comprises sequences of activities, which represent the instances of the process. Based on these instances, it is possible to discover a process model, which indicates the relationship, more precisely the workflow, between the observed events.

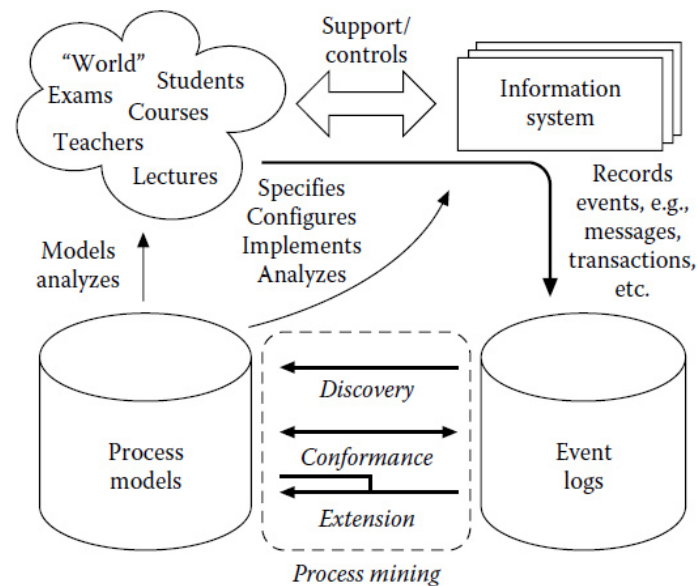


Figure 3. The general concept of using Process Mining in educational settings (from Trčka et al., 2010, p. 124).

Just like other analysis techniques, the application of PM is accompanied by several assumptions. First, PM assumes the presence of a hidden underlying process that produced the observed sequence of events, and that can be discovered from the instances stored in the event log. In the case of SRL, such a process could be a learning strategy or advice from instructional support that directs the student's mental processes (Bannert et al., 2014). Therefore, the regulatory behavior can be seen as driven by a holistic model of a process. Second, the data stored in the event log must be temporally ordered activities of several process instances. Third, the event log must contain a representative sample for the process under investigation. That does not necessarily imply that a large volume of data is better in any case, but that behavioral traces are needed that are meaningful for the research objectives. Finally, the purity of data, that is, the ratio of signal to noise, as well as the heterogeneity of behavior between cases have to be considered. PM algorithms can deal with noise in the data and low-frequency behavior, but a significant amount of noise makes process discovery more difficult. Moreover, because of the attempt to discover a common underlying process, the instances must represent a homogeneous sample. Otherwise, a pre-

selection of cases or a clustering technique might be needed in advance. For example, in the first analysis (Sonnenberg & Bannert, 2015), we used a trace clustering technique to verify this assumption. In general, researchers have to ask themselves critically if it is appropriate to use PM for analyzing their process data by taking into account these assumptions.

We use PM techniques to investigate the impact of instructional support on learning processes because of the following specific characteristics and strengths in comparison to other analysis methods for event sequences. First, we argued that PM is appropriate when conceptualizing SRL from a process-orientated view, which means, the observed behavioral traces are generated by a self-regulation process (Bannert et al., 2014; Reimann, 2009). This theoretical point of view determines the required analysis method. For instance, Reimann (2009) discussed that a holistic view of a process is necessary, not a process-as-sequence perspective. The assumption of an underlying process distinguishes PM from other techniques such as sequence mining, sequential pattern analysis, and stochastic methods. All recorded events are taken into account to generate an end-to-end process model, not only reoccurring sequences of events as in Sequential Pattern Mining (Zhou, Xu, Nesbit, & Winne, 2010). Hidden Markov models (e.g., Jeong et al., 2008) also incorporate a holistic process perspective, but they are time-consuming iterative procedures, and the output model is hard to interpret (van der Aalst, 2011). Moreover, hidden Markov models, as well as transitions graphs, represent a lower level of abstraction compared to the PM modeling language.

Second, it is one distinct feature of PM that the discovery of process models can be designed adaptively because the level of detail of output models can be controlled by parameter settings. That means the granularity of process data does not necessarily be pre-determined, but the model representation can be influenced dynamically by the researcher. For example, by increasing a cut-off value, it is possible to generate a more abstract model with fewer details. In general, the output models are similar to transition graphs (Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007; Siadaty, Gašević, & Hatala, 2016), but PM offers more options to control the level of detail adaptively.

Third, the output models allow the parsing of an activity sequence and the prediction of new behavior. Therefore, models can be compared through formal parameters, and the validity of a model can be tested using additional data (Rozinat & van der Aalst, 2008). The implemented conformance checking allows a relatively simple comparison of a model and a sequence of events, as illustrated in the third analysis of this dissertation (Sonnenberg & Bannert, submitted).

Fourth, PM explicitly deals with noise in the data and helps researchers to concentrate on the main relations among learning activities. When examining complex real-life event logs, analysis methods are needed that are robust to noise in the data. For example, the coded think-aloud data used in our analyses represent no perfect trace of behavior because the participants might not have uttered all learning activities or because of an erroneous code assignment.

Fifth, PM possesses the practical strengths of identifying process models automatically from event logs, even if large volumes of data are stored (Reimann et al., 2014). Finally, a unified framework for data import supports the application of PM, and a variety of algorithms that can be assigned to one of the three main functions described above and that are designed to meet specific requirements, such as dealing with noise, are available to the process analyst. The application of PM is addressed in more detail in the next section.

6.3 The Application of Process Mining

Romero and Ventura (2010) recommended that EDM tools should be designed user-friendly (e.g., comprising wizard tools and intuitive interfaces) because educators are usually non-experts in using data-mining frameworks. PM meets this requirement by providing a comprehensive software framework and a growing community that offers user support (see <http://processmining.org/>). We used the ProM Framework Version 5.2 (2008) for conducting our analyses. It comprises a variety of PM algorithms that can be assigned to the functions of model discovery, conformance checking, and model extension. For data preparation from our coded think-aloud protocols in a spreadsheet format, we used an additional software tool, Fluxicon Disco Version 1.7.2 (2014).

To illustrate the application of PM using coded learning activities from verbal reports, Figure 4 shows a simplified representation of the data input and output, as well as the functions of discovery of models and conformance checking. The starting point is the event log that contains the students' learning activities; more specifically, an ID, time stamps, and activity labels for each case. The learning activities are based on the coding scheme that was described above. The information stored in the event log allows computing the frequency of events, the relative arrangement of multiple events, and the event duration. The latter was not relevant for our purpose because it was only necessary to determine the temporal order of events. Using a PM algorithm for process discovery, the ProM framework allows to generate a process model inductively, for example as displayed on the right side in Figure 4. This output model describes the underlying process based on the event log, and it

is a visual representation of the dependencies of the event classes. The metrics such as transition dependencies that are normally also displayed in the output model are left out in Figure 4. Dependent on the used algorithm, the output is a specific variant of a Petri net. To set the level of abstraction from noise and low-frequency behavior, the researcher can adjust the output model by parameter settings (e.g., thresholds). It is a strength of PM that the output model can be stored and used in additional analyses, all in a unified software framework. For instance, additional data of the same students or another sample of learners can be used to measure the conformance between model and event log.

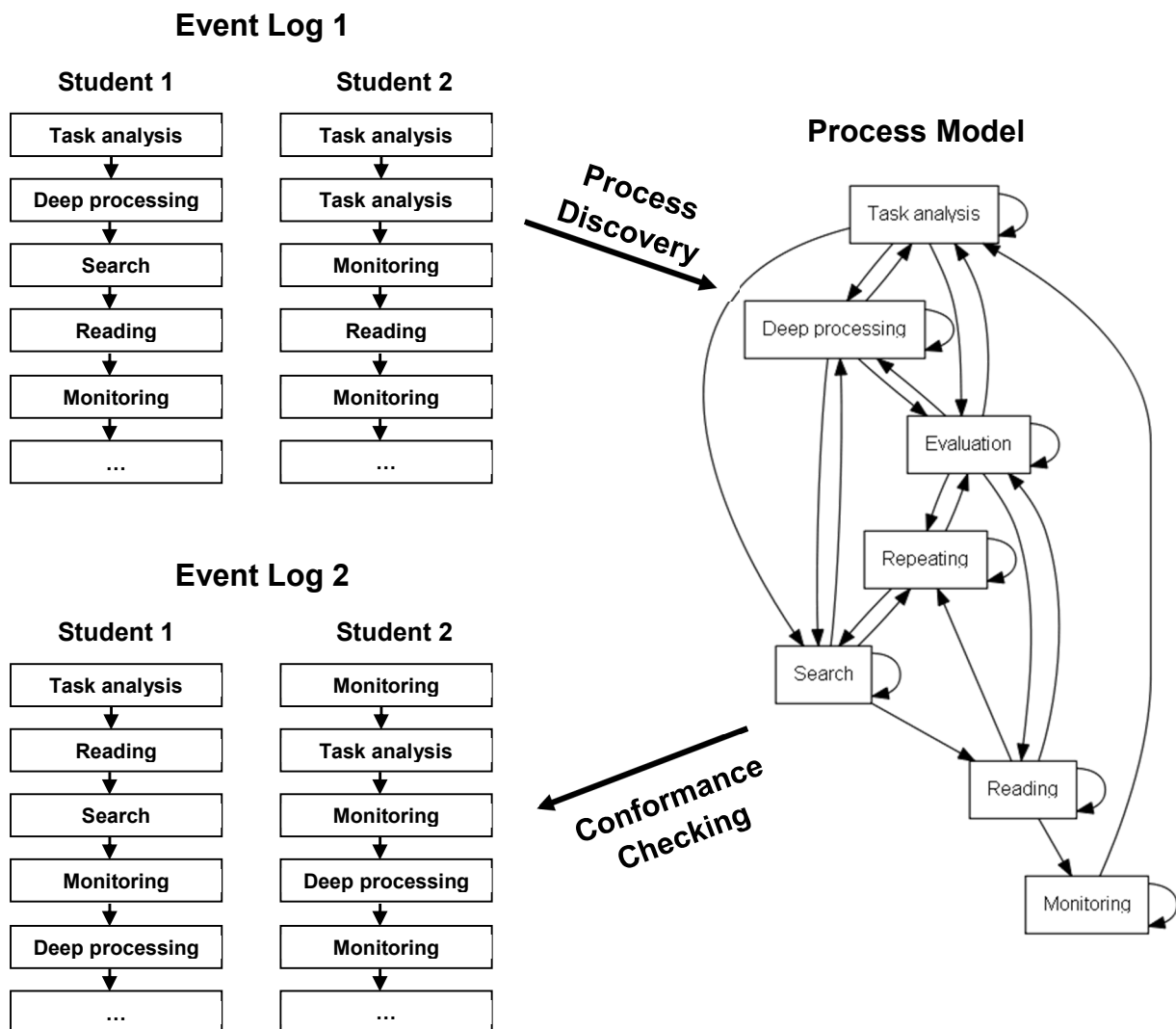


Figure 4. The Process Mining functionality.

In our analyses, we applied two algorithms for discovery and a conformance checking technique. We selected the discovery algorithms based on a comparison of seven state-of-the-art PM techniques on the dimensions accuracy and comprehensibility, which used real-life event logs (De Weerd, De Backer, Vanthienen, & Baesens, 2012). In our first

analysis (Sonnenberg & Bannert, 2015), we used the HeuristicsMiner algorithm (Weijters, van der Aalst, & de Medeiros, 2006) because the output model, a so-called heuristic net, can be automatically converted into a Petri net and used in further analyses. Additionally, the algorithm can be combined with a trace clustering procedure (DWS; de Medeiros et al., 2008) to avoid underfitting, that is, a resulting model is too general. However, this algorithm is reaching its limits when dealing with too many event classes (i.e., categories); therefore, we had to simplify the coding scheme. In contrast, the Fuzzy Miner algorithm (Günther & van der Aalst, 2007), which was used in the second analysis (Sonnenberg & Bannert, 2016), can deal better with less structured data in appearance and allows a flexible simplification of output models. For example, it abstracts from less significant event classes and displays the main features of a process. However, a direct conversion into a Petri net is not possible. In general, both algorithms are robust to noise in the data, and they require the setting of parameters that guide the model discovery. To test the discovered process models from the first analysis with additional data, we applied the conformance checker (Rozinat & van der Aalst, 2008) in our third analysis (Sonnenberg & Bannert, submitted). It allows measuring the conformance of a complete process model, represented as a Petri net, and an event log. Within the ProM framework, other conformance checking techniques are available, but they do not test complete models. For example, in Bannert et al. (2014) we illustrated the use of the LTL Checker (De Beer & van den Brand, 2007) by testing specific sequential patterns that can be derived from SRL models. Finally, PM also offers data visualization techniques, which become more important when dealing with large volumes of data and which can support the selection of measurement units. For instance, in our second analysis, we first visualized the sequences of events using a Dotted chart (Song & van der Aalst, 2007; van der Aalst, 2011) to determine the time window following a prompt in which an impact on metacognition had occurred. The functionality of these PM techniques are explained in more detail in the manuscripts (Sonnenberg & Bannert, 2015, 2016, submitted).

6.4 The Benefits of Process Mining for Technology-Enhanced Learning Issues

As indicated by recent studies that investigate TEL with a focus on self- and co-regulatory learning processes, PM techniques might have the potential to advance the field by allowing researchers to explore and to test process patterns on a micro level. For instance, Bannert et al. (2014) compared the process models of students with high versus low learning performance and revealed differences in the sequential patterns of regulatory processes. Moreover, studies from the CSCL context showed that PM contributes to a deeper

understanding of learning (Malmberg, Järvelä, Järvenoja, & Panadero, 2015; Reimann, Frerejean, & Thompson, 2009; Schoor & Bannert, 2012). In comparison to other analysis techniques, PM allows a comparison of extracted process patterns and assumptions in SRL models, for example, the concept of a time-ordered sequence of activities or dynamic and cyclical patterns. As described above, PM seems to be appropriate for the event-based view of SRL and it has specific characteristics that distinguish PM from other techniques for sequential analysis. Therefore, PM is a promising method in SRL research that has benefits compared to a traditional analysis of frequency distributions by using variance analysis and other EDM and LA approaches (Bannert et al., 2014; Reimann, 2009).

Despite the initial evidence of the benefits of PM for analyzing TEL processes, more research is necessary to gain a deeper understanding of the contribution that these techniques might make. The available rich behavioral traces of learners on various data channels offer new opportunities to advance a supporting strategy and to work towards adaptive interventions. However, that will only be possible by using appropriate analysis techniques for the evaluation of instructional support on the micro level; for instance, to conduct an in-depth investigation of scaffolding effects on learning activities. Therefore, the objective of the present dissertation is to explore the potential of PM techniques for analyzing the impact of instructional support on micro-level processes and gathering more information on the strengths as well as possible weaknesses of this approach in SRL settings. From a methodological point of view, the focus will be especially on the additional value of PM in comparison to a traditional analysis of frequencies. The type of support that is under investigation are metacognitive prompts that were provided during hypermedia learning (e.g., Bannert & Mengelkamp, 2013). The data channel are concurrent think-aloud protocols, which were coded to performed learning activities. In the following, the research objectives and findings of the three PM analyses are reported.

7. Summary of Findings: Process Mining Analyses

This chapter presents the research objectives and results of three analyses that were conducted to investigate the potential of PM for the evaluation of instructional support. Each analysis refers to think-aloud data from an extensive study that aimed at examining the impact of metacognitive prompting during hypermedia learning (Bannert et al., 2015).¹ Figure 5 displays the research design of this study. Overall, the study comprised three

¹ This research was supported by funds from the German Research Foundation (DFG: BA 2044/7-1).

sessions, and process as well as product data. During the first session, we assessed learner characteristics using tests and questionnaires. In the second and third session, students worked through a hypermedia learning environment for 40 minutes while thinking aloud concurrently. During the second session, metacognitive prompts supported participants that were assigned to the experimental group. The prompts were requests to give reasons for node selections. Immediately after working on the learning environment, the performance was measured using three learning tests (see Appendix 2 to 4 for a screenshot of the learning environment with a prompt and the learning tasks).

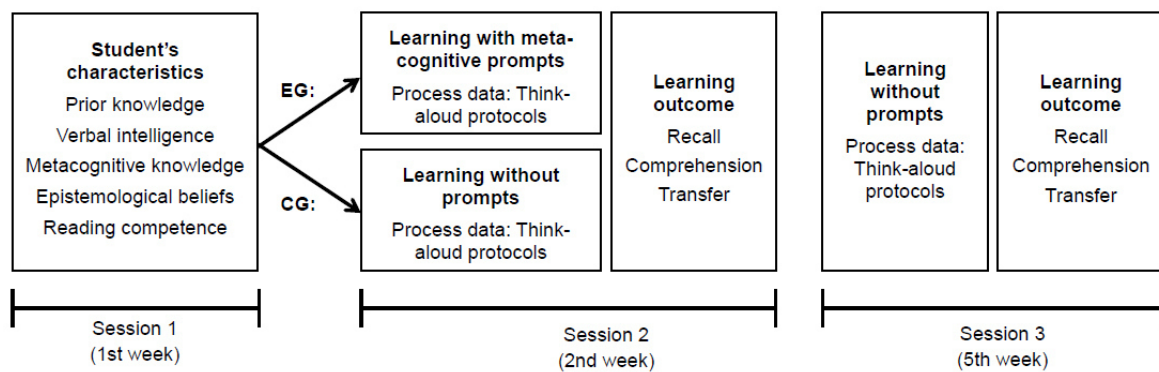


Figure 5. Research design. EG = experimental group, CG = control group.

The analyses make use of the coded think-aloud data as follows. The first analysis uses the data of the second session and compares the learning activities between the experimental and control groups. The second analysis only considers the data of the experimental group during learning when prompted (i.e., session 2), but on a very detailed level, in order to investigate conditions for effective prompts. Finally, the third analysis uses the data of the two learning sessions (i.e., session 2 and session 3) to examine the long-term effects of metacognitive prompting on the students' self-regulatory behavior. Although student characteristics, such as measured in session 1, might affect the learning process, they are not in the scope of this dissertation. The individual research questions and a summary of findings for each analysis are reported in the following sections.

7.1 Analysis 1: Research Questions and Results

The first analysis (Sonnenberg & Bannert, 2015) addressed the impact of metacognitive prompting on the students' self-regulatory behavior during hypermedia learning, particularly on the sequential structure of learning activities. Therefore, we compared the metacognitive, cognitive, and motivational utterances of a prompted experimental group and a control group by using a frequency analysis, a mediation analysis

that also considered learning outcome, and a PM technique. In detail, we investigated (i) whether metacognitive prompting during learning influences SRL processes by engaging students in more metacognitive learning activities, (ii) whether the number of metacognitive events mediates the beneficial effect on learning outcome, and (iii) which sequential patterns are induced through prompting. Our results showed that participants that were supported by metacognitive prompts articulated significantly more metacognitive learning activities, and additionally they achieved better performance in a learning test that measured the application of knowledge. Moreover, the findings of a mediation analysis indicated that prompting increased the number of metacognitive events, especially monitoring activities, which in turn positively affected the learning outcome. We expected these findings based on metacognitive prompting research (e.g., Azevedo et al., 2004; Bannert, 2009) as well as the assumed key mechanism of this type of instructional support (e.g., Bannert & Mengelkamp, 2013). Both analyses were based on the frequencies of coded learning activities. To illustrate the additional value of taking into account the relative arrangement of learning events using PM, we applied an algorithm that inductively discovered specific sequential patterns in the learning process. A comparison of the process models of students in the experimental and the control group revealed two noticeable differences. First, students supported through prompting showed a much better integration of preparing activities (i.e., orientation, planning, and goal specification), whereas these activities were quite isolated in the model of the control group. Second, the process model of the experimental group showed a higher number of loops between cognitive and metacognitive learning activities, which indicates a more active regulation of learning compared to the control group. Again, these results are in line with theoretical assumptions (e.g., Winne & Hadwin, 2008; Zimmerman, 2008), which emphasize the significance of orientation phases and an active regulation for successful learning. Finally, with respect to the potential of PM, the first analysis indicated that the displayed sequential patterns in the students' process models reveal additional information on the effects of metacognitive prompting that could not be shown by a traditional frequency analysis of occurring learning activities.

7.2 Analysis 2: Research Questions and Results

The main focus of the second analysis (Sonnenberg & Bannert, 2016) was on the impact of a single metacognitive prompt on a student's learning process, in order to explore the conditions for its effectiveness. Therefore, time intervals preceding and following each prompt presentation represent the measurement unit of this analysis. In detail, we

investigated (i) whether it is possible to classify the metacognitive prompts in terms of the activation of regulatory activities using the coded think-aloud data, (ii) whether the classification of effectiveness corresponds with learning outcome, and (iii) which conditions affect the induction of metacognitive learning activities through prompting. Based on the classification of prompts, we used a data-mining (DM) approach that considered the frequencies of various learning activities preceding a prompt, as well as a PM algorithm that additionally took into account the sequential patterns of events, to determine learning behavior that affects the effectiveness of a prompt. Our results showed that it is possible to distinguish between effective and non-effective prompts by considering the coded think-aloud data. Moreover, although we observed no significant correlation between the number of induced metacognitive activities and learning performance, we found a positive correspondence between monitoring and transfer performance, which is in line with previous research (e.g., Sonnenberg & Bannert, 2015). Furthermore, the findings of a DM analysis revealed that a high occurrence of orientation and monitoring activities fostered the desired prompting effects. A PM analysis that also took into account the relative arrangement of learning activities preceding a prompt supported and extended these results. We compared two inductively generated process models that represent the learning process preceding the appearance of an effective and a non-effective prompt. In the case of a non-effective prompt, the sequence of learning activities already resembled the performance of successful regulation patterns as described in SRL models (e.g., Winne & Hadwin, 2008; Zimmerman, 2008), which means that an intervention through prompting might not have been necessary. In contrast, the process model of effective prompts resembled a poor regulation of learning, and it comprised metacognitive activities that were not yet well embedded in the course of learning. In general, the second analysis showed how to use process data to evaluate the effectiveness of instructional support on the micro level by applying DM and PM techniques. As in the first analysis, we contrasted the findings of an analysis technique that only considers the frequency of coded activities with a PM techniques that also addresses the sequence of events. Again, it was possible to demonstrate the additional value of PM for the evaluation of metacognitive prompting effects on a very detailed level.

7.3 Analysis 3: Research Questions and Results

The third analysis (Sonnenberg & Bannert, submitted) investigated the long-term effects of metacognitive prompting on students' self-regulatory behavior. Therefore, we made use of the mined process models of the first analysis and applied a PM technique called

Conformance Checking to assess the stability of effects during a further learning task without instructional support. The research questions of this analysis concentrated on (i) the impact of metacognitive prompts on the learning process in a follow-up task three weeks after the intervention and (ii) the discovery of sequential patterns that were transferred to the future task using a PM technique. In contrast to the previous analyses that concentrated on the inductive discovery of process models, this analysis demonstrates an approach to model testing. Again, we first conducted an analysis based on frequencies. This time, we applied a mixed MANOVA that used the treatment (experimental vs. control group) as the independent variable, and the coded learning activities of two learning sessions as repeated-measures variable. The findings showed that the prompted experimental group demonstrated more metacognitive learning activities in both sessions, but also that there was a general downward trend of uttered metacognitive activities. In a second step, we compared the sequence of learning activities between two learning tasks using Conformance Checking. The results of this analysis also indicated that prompting effects remained stable over time by showing high fitness values between the process models of the first learning session and the data of the second session. Moreover, the PM analysis provided a more detailed view on the conformance between model and event log by showing the points of mismatches. However, the metrics that measure the precision of the process models also indicated that they might be too general (i.e., allowing for more behavior than observed). In conclusion, this analysis demonstrated the potential of PM for confirmatory model testing by taking into account the relative arrangement of learning activities in addition to their frequency distribution. Additionally, the findings provide initial evidence for sustainable long-term effects of metacognitive prompting on hypermedia learning.

8. General Discussion of Findings

The research agenda of the present dissertation addressed the potential of PM techniques for the evaluation of instructional support. Three analyses were conducted in the context of supporting hypermedia learning through metacognitive prompting to investigate the additional value of PM. This chapter discusses what could be achieved as well as limitations of PM for educational settings. Furthermore, some recommendations how to use PM for SRL research objectives are presented. Finally, this chapter points out important directions for future work.

8.1 The Revealed Potential of Process Mining

All three analyses indicated the additional value of PM in comparison to a traditional frequency-based analysis. When conceptualizing learning as a dynamic sequence of multiple events, PM contributes to the detection of important routing points and regulatory loops. This can only be achieved by using an analysis method that takes into account the sequential structure of learning activities. Moreover, regarding the assessment of metacognitive prompting effects on hypermedia learning, the findings showed that this type of instructional support affected not only the frequency of regulatory activities but also the deployment of the sequence of events.

In general, PM has the potential to optimize instructional support by discovering process patterns. For instance, the comparison of process models in analysis 1 (Sonnenberg & Bannert, 2015) indicated that the sequential deployment of evaluating activities should be scaffolded in more detail. In particular for the design of adaptive support and adaptive hypermedia systems that tailor support to the student's requirements (e.g., Bouchet et al., 2013; Brusilovsky, 2007; Molenaar & Roda, 2008), a detailed understanding of the learning process is a crucial prerequisite. Therefore, PM might help to increase the diagnosis of the proper timing of scaffolds, the calibration of support to SRL phases, and the gradual reduction of support (Azevedo & Hadwin, 2005). Analysis 2 (Sonnenberg & Bannert, in press) demonstrated how researchers might use PM techniques to examine the conditions for effective prompts. The temporal positioning of a scaffold within the learning process can significantly influence compliance and thereby the effectiveness of prompting (Azevedo et al., 2011; Sitzmann, Bell, Kraiger, & Kanar, 2009; Thillmann et al., 2009). PM findings on the learning process preceding the presentation of instructional support and how these learning activities affect its effectiveness can help to design production rules that trigger a support device adaptively. Additionally, PM techniques do not only allow to discover process patterns but also to test process models using additional data. As illustrated in analysis 3 (Sonnenberg & Bannert, submitted), researchers can compare the conformance between a process model and an event log to investigate the stability of supporting effects on the process level.

Moreover, the representation of the sequential characteristics of a learning process as a visual model that allows a comprehensive understanding of the course of events, is another benefit of PM, especially in comparison to other analysis techniques for sequential data. Possibly, these models might also be used as process feedback for students (Reimann et al., 2009; Sedrakyan, De Weerd, & Snoeck, 2016), because they represent displays of traces

students can interpret (Winne, 2014). Furthermore, the findings of PM might support the refinement of current theoretical models by providing new insight into micro-level processes. Finally, PM algorithms are robust to noise and low-frequent behavior, which are inevitably present in data from authentic learning settings.

8.2 Limitations of Process Mining

Despite the potential of PM for the evaluation of instructional support, there are also several limitations that have to be considered. First, because PM represents an inductive approach, the validity of mined process patterns depends on the data quality and the representativity of the behavioral traces stored in the event log (Reimann et al., 2009). Second, the discovered process models are descriptive models. Although these models can be tested using additional data, conformance checking differs from inferential statistics. PM was originally developed for the practical purpose of optimizing business processes, not for the significance testing of model assumptions. Still, the process models might inform the development and refinement of theories (Bannert et al., 2014). Third, the findings presented in this dissertation are dependent on the learning setting, the sample of students, and the coding scheme for the think-aloud data. Therefore, the PM results are task-specific, and they probably do not represent general patterns. Before drawing conclusions for the design of instructional support, generalizations for other contexts and other samples still have to be verified in future research. For example, another measurement unit or process data from other channels might result in different findings. Fourth, we found indications that the precision of the process models might be improvable (see analyses 3), that means, the sequence of events among the students might comprise a high variance in learning behavior. Consequently, a common process model for all students might not be appropriate. Fifth, the selection of the time window in which instructional support affects the learning process remains challenging. In analyses 1 und 3, we considered the entire learning episode; whereas in analyses 2, we focused on time intervals around the presentation of prompts. We did not take into account the student's current learning goal explicitly when determining the boundaries of a SRL process, as for example demanded by Winne (2014). Possibly, it might be necessary to explore in more detail when the sequential structure of a learning process changes, because then a new process model would be needed. A qualitative analysis of the think-aloud data might provide more insights into the qualitative changes of the learning process during the 40-minute episode. Finally, it is a key feature of PM algorithms to abstract from less frequent behavior and to focus on the main relations of a process. However, if the

object of investigation is such a behavior (e.g., an anomaly), then PM might not be the appropriate analysis method.

8.3 Lessons Learned - Recommendations How to Use Process Mining

Based on the previous experience with the application of PM on educational data, there are some recommendations for researchers. First, it is possible to generate process models in a relatively simple way by using the PM framework; however, before interpreting the visual output, researchers need to familiarize themselves with the PM literature (see van der Aalst, 2011 for an excellent introduction). An understanding of the main assumptions of PM (e.g., the existence of an underlying process) and how parameter settings influence the mining procedure represent essential basics.

Second, the analyses of this dissertation followed the scheme (i) theoretical derivation of research objectives, (ii) data inspection and visualization, (iii) selection of measurement units, and (iv) data analysis as well as interpretation of results. In general, these are common steps for empirical work. However, it is important to emphasize that researchers should not build on purely data-driven results, but embed their procedure in conceptual and theoretical work. For instance, theory and the research objectives determine the required granularity of process data. Moreover, when dealing with large volumes of data, the relevance of pre-selection, pre-analysis, and data visualizations increases. Research considers the selection of measurement units a general challenge (Johnson et al., 2011; Winne, 2014). As shown in analysis 2 (Sonnenberg & Bannert, 2016), data visualization and verification checks can support researchers with this challenge.

Third, learning is a complex and heterogeneous phenomenon (Martin & Sherin, 2013; Reimann et al., 2014), and research indicates a high variance among students' regulatory behavior (Hadwin et al., 2007; Winne, 2014). The complexity and heterogeneity of learning pose a great challenge for process analysis. The PM framework allows to use trace clustering techniques (De Weerd, Vanden Broucke, Vanthienen, & Baesens, 2013; Greco, Guzzo, Pontieri, & Saccà, 2006), which should be considered before computing a process model. For example, in analysis 1 (Sonnenberg & Bannert, 2015), we checked if it is possible to split the event log into more homogeneous subsamples. Alternatively, variable-based clustering that uses the frequency of events (Biswas, Jeong, Kinnebrew, Sulcer, & Roscoe, 2010; Bouchet et al., 2013), and a selection based on learner characteristics or learning outcome (e.g., Bannert et al., 2014) might help to identify similar groups of learners. However, when taking into account an event-based perspective, trace clustering is more

appropriate, and those techniques might become increasingly important in future (e.g., De Weerd et al., 2013).

Fourth, as illustrated in analysis 2 (Sonnenberg & Bannert, in press), it might be useful to combine the results of DM and PM techniques. For instance, researchers can check if the findings of both approaches indicate similar patterns. Recently, the application of PM within a prominent DM framework was facilitated through the development of the RapidProM plugin².

Fifth, PM depends on the representativeness of an event log for the observed behavior. Consequently, the availability of meaningful data is more important than their volume (e.g., Reimann et al., 2014). The think-aloud data that was used in the presented analyses comprised a detailed trace of learning behavior, whose granularity was appropriate for the research objectives. Therefore, a larger number of data points but on another data channel and granularity (e.g., computer log files) would not have been more useful necessarily.

Finally, especially for the selection of measurement units, researchers need to also consider the expected stability of behavior as well as the expected sustainability of instructional support on the learning process. For example, in analysis 3 (Sonnenberg & Bannert, submitted), we selected two learning episodes to investigate the long-term effects of metacognitive prompts on regulatory behavior. PM allows the comparison of a process model and additional event data, but researchers must determine the time window for observation, again with respect to theoretical assumptions.

8.4 Future Directions: A Quest for Standards and Guidelines

Despite the first promising results using PM to discover sequential patterns, the parameter setting as well as further analytical decisions can be challenging due to missing standards and guidelines for this approach in educational settings. Moreover, more information is needed on the validity of identified process patterns and their significance for educational interventions (e.g., for instructional support). Therefore, more research on the development of standards and guidelines for using PM, and for other new analytical approaches taking into account fine-grained process data, is needed. These standards should address analytical and methodical issues like a recommended scheme for process analysis, a unified terminology and framework, and a decision support for selecting measurement units and techniques.

² See <http://www.promtools.org/doku.php?id=rapidprom:home> for more information on RapidProM.

The following two concrete proposals might be a starting point for pursuing the quest for standards and guidelines. First, it is necessary to replicate findings on process patterns using additional data, especially since most of the current analyses are exploratory. Shifting from exploratory to confirmatory analyses that test patterns (Winne, 2014), we might obtain more certainty about the validity of applied analytical approaches. Moreover, the alignment of SRL data on different data channels (Azevedo, 2014), the combination of process and product data, and the integration of findings across several SRL studies (Dent & Hoyle, 2015) might also provide additional information on the validity of analysis methods and their results. Hence, researchers could gain a deeper understanding of new methodological approaches for analyzing process data and the impact of taken analytical decisions.

Second, we should aim at a recommendation catalog for analysis techniques addressing the dynamics of SRL data. This catalog could inform researchers about the appropriateness of a particular technique considering the features of the present data and their research objectives. For example, given a certain level of granularity, or a certain data channel, it is recommended to use technique X to analyze the temporal dynamics of regulation behavior. The working plan for such a catalog could be the application of various analytical approaches to the same data sets, and thereby attempting to observe advantages and disadvantages of specific techniques for these data sets. More precisely, a comparison of different methods for sequential and temporal analyses presented in recent special issues in the journal *Metacognition and Learning* (Ben-Eliyahu & Bernacki, 2015; Molenaar & Järvelä, 2014) could be a starting point.

In addition to these two proposals, future research needs to concentrate on conceptual work and the refinement of theoretical process models of learning. PM allows to test empirically mined or theory-based models (Bannert et al., 2014), but more elaborated process assumptions would be needed that correspond with the granularity of measured event data. The analysis of fine-grained behavioral traces requires a fine-grained theory, which is currently missing in the SRL field (Molenaar & Järvelä, 2014). The conceptualization of learning as a dynamic sequence of events kindled the growing interest in process data and advanced analysis methods, which, in turn, now stimulates the need for theoretical refinements (Ben-Eliyahu & Bernacki, 2015). With respect to the impact of instructional support on TEL processes, there is currently no theoretical model that explains causal effects on the micro level. Furthermore, Reimann et al. (2014) argued that data-intensive approaches such as EDM techniques are not sufficient for theory development because inductive process patterns are only conceptually interesting if they are combined

with theoretical explanations. Therefore, research needs to develop and test causal models and mechanisms that help to understand and predict sequences of learning activities on various data channels as well as granularity levels.

8.5 Conclusions: The Final Chapter?!

In conclusion, PM techniques have the potential to contribute to the advancement of visualizing and understanding the impact of instructional support using fine-grained process data. Detailed process analyses, such as demonstrated in the three contributions of this dissertation, are necessary for the design and evaluation of effective support. PM can contribute not only to the discovery of dynamic learning behavior but also to the implementation of confirmatory analysis that allows examining sustainable long-term effects of scaffolds on the process level. Such analyses provide insights into the robustness of pedagogical interventions and the transfer of competencies to new situations. However, validity issues and comparisons with related techniques need to be addressed for PM, as well as for additional analytical techniques currently used to investigate process patterns.

Returning to the search for hidden treasures, the current dissertation has demonstrated how PM techniques might help to discover one type of treasure, namely the impact of metacognitive prompting on learning activities during hypermedia learning. Although this is the last chapter of my dissertation, research on the dynamics of TEL processes and advanced analysis techniques will have to face many more challenges in future work. Moreover, PM has a vivid community that constantly advances its approach and algorithms. Therefore, the development of analytical approaches is still evolving.

In general, I agree with the statement of Baker and Inventado (2014, p. 71) that “[...] the question is not which methods are best, but which methods are useful for which applications, in order to improve the support for any person who is learning, whenever they are learning”. Additionally, researchers should consider well which approach might be appropriate and to always remain critical towards a chosen analysis method. Finally, more researchers should be encouraged to use PM and related techniques to analyze their process data with respect to the sequential and temporal aspects of learning events. Because only in this way, the field can make progress towards the establishment of advanced analysis techniques.

9. References

- Aggarwal, C. C. (2015). *Data mining. The textbook*. New York: Springer. doi:10.1007/978-3-319-14142-8
- Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, 40(4), 199–209. doi:10.1207/s15326985ep4004_2
- Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning*, 4(1), 87–95. doi:10.1007/s11409-009-9035-7
- Azevedo, R. (2014). Issues in dealing with sequential and temporal characteristics of self- and socially-regulated learning. *Metacognition and Learning*, 9(2), 217–228. doi:10.1007/s11409-014-9123-1
- Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, 50(1), 84–94. doi:10.1080/00461520.2015.1004069
- Azevedo, R., Cromley, J. G., Moos, D. C., Greene, J. A., & Winters, F. I. (2011). Adaptive content and process scaffolding: A key to facilitating students' self-regulated learning with hypermedia. *Psychological Testing and Assessment Modeling*, 53(1), 106–140.
- Azevedo, R., Guthrie, J., & 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(1–2), 87–111. doi:10.2190/DVWX-GM1T-6THQ-5WC7
- Azevedo, R., & Hadwin, A. F. (2005). Scaffolding self-regulated learning and metacognition - Implications for the design of computer-based scaffolds. *Instructional Science*, 33(5), 367–379. doi:10.1007/s11251-005-1272-9
- Azevedo, R., Harley, J. M., Trevors, G., Duffy, M., Feyzi-Behnagh, M., Bouchet, F., & Landis, R. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 427–449). New York: Springer. doi:10.1007/978-1-4419-5546-3_28
- Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist*, 45(4), 210–223. doi:10.1080/00461520.2010.515934
- Baker, R., & Corbett, A. (2014). Assessment of robust learning with educational data mining. *Research & Practice in Assessment*, 9, 38–50.
- Baker, R. S., & Inventado, P. S. (2014). Educational Data Mining and Learning Analytics. In J. A. Larusson & B. White (Eds.), *Learning Analytics: From Research to Practice* (pp. 61–75). New York: Springer.
- Baker, R. S. (2010). Data mining for education. In B. McGaw, P. Peterson, & E. Baker (Eds.), *International Encyclopedia of Education* (pp. 112–118). Oxford, UK: Elsevier.

- Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, 1(1), 3–16.
- Bannert, M. (2007). *Metakognition beim Lernen mit Hypermedia. Erfassung, Beschreibung und Vermittlung wirksamer metakognitiver Lernstrategien und Regulationsaktivitäten [Metacognition and Hypermedia Learning]*. Münster: Waxmann.
- Bannert, M. (2009). Promoting self-regulated learning through prompts: A discussion. *Zeitschrift Für Pädagogische Psychologie*, 23(2), 139–145. doi:10.1024/1010-0652.23.2.139
- Bannert, M., & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. Does the verbalisation method affect learning? *Metacognition and Learning*, 3(1), 39–58. doi:10.1007/s11409-007-9009-6
- Bannert, M., & Mengelkamp, C. (2013). Scaffolding hypermedia learning through metacognitive prompts. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 171–186). New York: Springer. doi:10.1007/978-1-4419-5546-3_12
- Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science*, 40(1), 193–211. doi:10.1007/s11251-011-9167-4
- Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning*, 9(2), 161–185. doi:10.1007/s11409-013-9107-6
- Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior*, 52, 293–306. doi:10.1016/j.chb.2015.05.038
- Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: A focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, 10(1), 1–13. doi:10.1007/s11409-015-9134-6
- Ben-Eliyahu, A., & Linnenbrink-Garcia, L. (2015). Integrating the regulation of affect, behavior, and cognition into self-regulated learning paradigms among secondary and post-secondary students. *Metacognition and Learning*, 10(1), 15–42. doi:10.1007/s11409-014-9129-8
- Biswas, G., Jeong, H., Kinnebrew, J. S., Sulcer, B., & Roscoe, R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology Enhanced Learning*, 5(2), 123–152. doi:10.1142/S1793206810000839
- Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction*, 7(2), 161–186. doi:10.1016/S0959-4752(96)00015-1
- Bouchet, F., Harley, J. M., Trevors, G. J., & Azevedo, R. (2013). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *Journal of Educational Data Mining*, 5(1), 104–146.
- Brusilovsky, P. (2007). Adaptive navigation support. In P. Brusilovsky, A. Kobsa, & W.

- Nejdl (Eds.), *The Adaptive Web. Methods and Strategies of Web Personalization* (pp. 263–290). Berlin Heidelberg: Springer. doi:10.1007/978-3-540-72079-9
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, 6(3), 271–315. doi:10.1207/s15327809jls0603_1
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. doi:10.1080/00461520.2014.965823
- Clarebout, G., & Elen, J. (2006). Tool use in computer-based learning environments: Towards a research framework. *Computers in Human Behavior*, 22, 389–411. doi:10.1016/j.chb.2004.09.007
- De Beer, H. T., & van den Brand, P. C. W. (2007). *The LTL checker plugins. A (reference) manual*. Eindhoven: University of Eindhoven.
- de Medeiros, A. K. A., Guzzo, A., Greco, G., van der Aalst, W. M. P., Weijters, A. J. M. M., van Dongen, B. F., & Saccà, D. (2008). Process mining based on clustering: A quest for precision. *Lecture Notes in Computer Science*, 4928, 17–29. doi:10.1007/978-3-540-78238-4_4
- De Smedt, B. (2014). Advances in the use of neuroscience methods in research on learning and instruction. *Frontline Learning Research*, 2, 7–14. Retrieved from <http://journals.sfu.ca/flr/index.php/journal/article/view/115>
- De Weerd, J., De Backer, M., Vanthienen, J., & Baesens, B. (2012). A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real-life event logs. *Information Systems*, 37(7), 654–676. doi:10.1016/j.is.2012.02.004
- De Weerd, J., Vanden Broucke, S., Vanthienen, J., & Baesens, B. (2013). Active trace clustering for improved process discovery. *IEEE Transactions on Knowledge and Data Engineering*, 25(12), 2708–2720. doi:10.1109/TKDE.2013.64
- Dent, A. L., & Hoyle, R. H. (2015). A framework for evaluating and enhancing alignment in self-regulated learning research. *Metacognition and Learning*, 10(1), 165–179. doi:10.1007/s11409-015-9136-4
- Devolder, A., van Braak, J., & Tondeur, J. (2012). Supporting self-regulated learning in computer-based learning environments: Systematic review of effects of scaffolding in the domain of science education. *Journal of Computer Assisted Learning*, 28(6), 557–573. doi:10.1111/j.1365-2729.2011.00476.x
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, Mass.: MIT Press.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *The American Psychologist*, 34, 906–911.
- Fluxicon Disco (Version 1.7.2) [Software]. (2014). Retrieved from <http://fluxicon.com/disco>
- Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin*, 137(2), 316–344. doi:10.1037/a0021663
- Greco, G., Guzzo, A., Pontieri, L., & Saccà, D. (2006). Discovering expressive process models by clustering log traces. *IEEE Transactions on Knowledge and Data*

- Engineering*, 18(8), 1010–1027. doi:10.1109/TKDE.2006.123
- Greene, J. A., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist*, 45(4), 203–209. doi:10.1080/00461520.2010.515935
- Greene, J. A., Dellinger, K. R., Tüysüzoglu, B. B., & Costa, L.-J. (2013). A two-tiered approach to analyzing self-regulated learning data to inform the design of hypermedia learning environments. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 117–128). New York: Springer. doi:10.1007/978-1-4419-5546-3_8
- Günther, C. W., & van der Aalst, W. M. P. (2007). Fuzzy mining – Adaptive process simplification based on multi-perspective metrics. *Business Process Management - Lecture Notes in Computer Science*, 4714, 328–343. doi:10.1007/978-3-540-75183-0
- Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2(2), 107–124. doi:10.1007/s11409-007-9016-7
- Hübner, S., Nückles, M., & Renkl, A. (2010). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction*, 20(1), 18–29. doi:10.1016/j.learninstruc.2008.12.001
- Jeong, H., Gupta, A., Roscoe, R., Wagster, J., Biswas, G., & Schwartz, D. (2008). Using hidden markov models to characterize student behaviors in learning-by-teaching environments. *Lecture Notes in Computer Science*, 5091, 614–625. doi:10.1007/978-3-540-69132-7_64
- Johnson, A. M., Azevedo, R., & D'Mello, S. K. (2011). The temporal and dynamic nature of self-regulatory processes during independent and externally assisted hypermedia learning. *Cognition and Instruction*, 29(4), 471–504. doi:10.1080/07370008.2011.610244
- Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *The NMC horizon report: 2016 higher education edition*. Austin, Texas: The New Media Consortium. Retrieved from <https://www.nmc.org/publication/nmc-horizon-report-2016-higher-education-edition/>
- Künsting, J., Kempf, J., & Wirth, J. (2013). Enhancing scientific discovery learning through metacognitive support. *Contemporary Educational Psychology*, 38(4), 349–360. doi:10.1016/j.cedpsych.2013.07.001
- Larusson, J. A., & White, B. (Eds.) (2014). *Learning Analytics. From Research to Practice*. New York: Springer. doi:10.1007/978-1-4614-3305-7
- Lawless, K. A., & Brown, S. W. (1997). Multimedia learning environments: Issues of learner control and navigation. *Instructional Science*, 25, 117–131. doi:10.1023/a:1002919531780
- Lin, X. (2001). Designing metacognitive activities. *Educational Technology Research and Development*, 49(2), 23–40. doi:10.1007/BF02504926
- Lin, X., & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching*, 36(7), 837–858.

doi:10.1002/(SICI)1098-2736(199909)36:7<837::AID-TEA6>3.3.CO;2-L

- Malmberg, J., Järvelä, S., Järvenoja, H., & Panadero, E. (2015). Promoting socially shared regulation of learning in CSCL: Progress of socially shared regulation among high- and low-performing groups. *Computers in Human Behavior*, *52*, 562–572. doi:10.1016/j.chb.2015.03.082
- Martin, T., & Sherin, B. (2013). Learning analytics and computational techniques for detecting and evaluating patterns in learning: An introduction to the special issue. *Journal of the Learning Sciences*, *22*(4), 511–520. doi:10.1080/10508406.2013.840466
- Mayer, R. E. (2009). *Multimedia learning (2nd ed.)*. New York: Cambridge University Press.
- Molenaar, I., & Chiu, M. M. (2014). Dissecting sequences of regulation and cognition: Statistical discourse analysis of primary school children's collaborative learning. *Metacognition and Learning*, *9*(2), 137–160. doi:10.1007/s11409-013-9105-8
- Molenaar, I., & Järvelä, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning*, *9*(2), 75–85. doi:10.1007/s11409-014-9114-2
- Molenaar, I., & Roda, C. (2008). Attention management for dynamic and adaptive scaffolding. *Pragmatics & Cognition*, *16*(2), 224–271. doi:10.1075/pc.16.2.04mol
- Moos, D. C., & Azevedo, R. (2009). Self-efficacy and prior domain knowledge: To what extent does monitoring mediate their relationship with hypermedia learning? *Metacognition and Learning*, *4*(3), 197–216. doi:10.1007/s11409-009-9045-5
- Moos, D. C., & Miller, A. (2015). The cyclical nature of self-regulated learning phases: Stable between learning tasks? *Journal of Cognitive Education and Psychology*, *14*(2), 199–218. doi:10.1891/1945-8959.14.2.199
- Nückles, M., Hübner, S., Dümer, S., & Renkl, A. (2010). Expertise reversal effects in writing-to-learn. *Instructional Science*, *38*(3), 237–258. doi:10.1007/s11251-009-9106-9
- Papamitsiou, Z., & Economides, A. A. (2014). Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Educational Technology & Society*, *17*(4), 49–64.
- ProM (Version 5.2) [Software]. (2008). Retrieved from <http://www.processmining.org>
- Puntambekar, S., & Hübscher, 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
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, *4*(3), 239–257. doi:10.1007/s11412-009-9070-z
- Reimann, P., Frerejean, J., & Thompson, K. (2009). Using process mining to identify models of group decision making processes in chat data. In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Computer-supported collaborative learning practices: CSCL2009 conference proceedings* (pp. 98–107). Rhodes, Greece: International Society for the Learning Sciences.

- Reimann, P., Markauskaite, L., & Bannert, M. (2014). E-Research and learning theory: What do sequence and process mining methods contribute? *British Journal of Educational Technology*, *45*(3), 528–540. doi:10.1111/bjet.12146
- Reisig, W. (1985). *Petri Nets. An introduction*. Berlin: Springer.
- Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, *2*, 7–12. Retrieved from <http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/4491>
- Romero, C., & Ventura, S. (2010). Educational Data Mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics*, *40*(6), 601–618. doi:10.1109/TSMCC.2010.2053532
- Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. (2010). *Handbook of educational data mining*. Boca Raton, FL: Chapman&Hall/CRC.
- Rozinat, A., & van der Aalst, W. M. P. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems*, *33*(1), 64–95. doi:10.1016/j.is.2007.07.001
- 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*, 64–96. doi:10.1016/j.cedpsych.2005.02.002
- Schnotz, W. (1998). Strategy-specific information access in knowledge acquisition from hypertext. In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, Tools, and Reasoning. Essays on Situated Cognition* (pp. 336–358). Berlin: Springer.
- Schoor, C., & Bannert, M. (2012). Exploring regulatory processes during a computer-supported collaborative learning task using process mining. *Computers in Human Behavior*, *28*, 1321–1331. doi:10.1016/j.chb.2012.02.016
- Schraw, G. (2010). Measuring self-regulation in computer-based learning environments. *Educational Psychologist*, *45*(4), 258–266. doi:10.1080/00461520.2010.515936
- Sedrakyan, G., De Weerd, J., & Snoeck, M. (2016). Process-mining enabled feedback: “Tell me what I did wrong” vs. “tell me how to do it right.” *Computers in Human Behavior*, *57*, 352–376. doi:10.1016/j.chb.2015.12.040
- Siadaty, M., Gašević, D., & Hatala, M. (2016). Measuring the impact of technological scaffolding interventions on micro-level processes of self-regulated workplace learning. *Computers in Human Behavior*, *59*, 469–482. doi:10.1016/j.chb.2016.02.025
- Sitzmann, T., Bell, B. S., Kraiger, K., & Kanar, A. M. (2009). A multilevel analysis of the effect of prompting self-regulation in technology-delivered instruction. *Personnel Psychology*, *62*, 697–734. doi:10.1111/j.1744-6570.2009.01155.x
- Song, M., & van der Aalst, W. M. P. (2007). Supporting process mining by showing events at a glance. In K. Chari & A. Kumar (Eds.), *Proceedings of 17th Annual Workshop on Information Technologies and Systems (WITS 2007)* (pp. 139–145). Montreal, Canada.
- Sonnenberg, C., & Bannert, M. (2015). Discovering the effects of metacognitive prompts on the sequential structure of SRL-processes using process mining techniques. *Journal of Learning Analytics*, *2*(1), 72–100. Retrieved from <http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/4090>

- Sonnenberg, C., & Bannert, M. (2016). Evaluating the impact of instructional support using data mining and process mining: A micro-level analysis of the effectiveness of metacognitive prompts. *Journal of Educational Data Mining*, 8(2), 51-83.
- Sonnenberg, C., & Bannert, M. (submitted). Using process mining to examine the sustainability of instructional support: How stable are the effects of metacognitive prompting on self-regulatory behavior? *Computers in Human Behavior*.
- Thillmann, H., Küsting, J., Wirth, J., & Leutner, D. (2009). Is it merely a question of “what” to prompt or also “when” to prompt? *Zeitschrift für Pädagogische Psychologie*, 23(2), 105–115. doi:10.1024/1010-0652.23.2.105
- Trčka, N., Pechenizkiy, M., & van der Aalst, W. M. P. (2010). Process mining from educational data. In C. Romero, S. Ventura, M. Pechenizkiy, & R. Baker (Eds.), *Handbook of Educational Data Mining* (pp. 123–142). Boca Raton, FL: Chapman&Hall/CRC.
- Trevors, G., Feyzi-Behnagh, R., Azevedo, R., & Bouchet, F. (2016). Self-regulated learning processes vary as a function of epistemic beliefs and contexts: Mixed method evidence from eye tracking and concurrent and retrospective reports. *Learning and Instruction*, 42, 31–46. doi:10.1016/j.learninstruc.2015.11.003
- van der Aalst, W. M. P. (2011). *Process mining: Discovery, conformance and enhancement of business processes*. Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-19345-3
- Veenman, M. V. J. (1993). *Metacognitive ability and metacognitive skill: Determinants of discovery learning in computerized learning environments*. University of Amsterdam.
- Veenman, M. V. J. (2005). The assessment of metacognitive skills: What can be learned from multi-method designs? In C. Artelt & B. Moschner (Eds.), *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis* (pp. 77–99). Münster: Waxmann.
- Veenman, M. V. J., Bavelaar, L., De Wolf, L., & Van Haaren, M. G. P. (2014). The on-line assessment of metacognitive skills in a computerized learning environment. *Learning and Individual Differences*, 29, 123–130. doi:10.1016/j.lindif.2013.01.003
- 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
- Walker, E., Rummel, N., & Koedinger, K. R. (2009). CTRL: A research framework for providing adaptive collaborative learning support. *User Modeling and User-Adapted Interaction*, 19(5), 387–431. doi:10.1007/s11257-009-9069-1
- Weijters, A. J. M. M., van der Aalst, W. M. P., & de Medeiros, A. K. A. (2006). Process mining with the HeuristicsMiner algorithm. *Cirp Annals-Manufacturing Technology*, 166, 1–34. Retrieved from http://cms.ieis.tue.nl/Beta/Files/WorkingPapers/Beta_wp166.pdf
- Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences*, 8(4), 327–353. doi:10.1016/S1041-6080(96)90022-9
- Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning*, 9(2), 229–237. doi:10.1007/s11409-014-9113-3

- Winne, P. H., & Baker, R. S. (2013). The potentials of educational data mining for researching metacognition, motivation and self-regulated learning. *Journal of Educational Data Mining*, 5(1), 1–8. Retrieved from <http://www.educationaldatamining.org/JEDM/index.php/JEDM/article/view/28>
- Winne, P. H., & Hadwin, A. F. (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). NY: Taylor & Francis.
- Winne, P. H., & Nesbit, J. C. (2009). Supporting self-regulated learning with cognitive tools. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 259–277). New York: Routledge.
- Winne, P. H., & Perry, N. E. (2005). Measuring self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 531–566). San Diego, California: Academic Press.
- Wirth, J. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie*, 23(2), 91–94. doi:10.1024/1010-0652.23.2.91
- Wirth, J., & Leutner, D. (2008). Self-regulated learning as a competence. *Zeitschrift für Psychologie*, 216(2), 102–110. doi:10.1027/0044-3409.216.2.102
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 17(2), 89–100. doi:10.1111/j.1469-7610.1976.tb00381.x
- Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review*, 17(2), 187–202. doi:10.1007/s12564-016-9426-9
- Zhou, M., Xu, Y. X., Nesbit, J. C., & Winne, P. H. (2010). Sequential pattern analysis of learning logs: Methodology and applications. In C. Robero, S. Ventura, M. Pechenizkiy, & R. Baker (Eds.), *Handbook of Educational Data Mining* (pp. 107–121). Boca Raton, FL: Chapman&Hall/CRC.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183. doi:10.3102/0002831207312909

10. Appendix

Appendix 1. Original scheme for coding the students' think-aloud data based on Bannert's (2007) framework for self-regulated hypermedia learning (in German)

Oberkategorie: Metakognition (Regulatorischer Prozess)	
Name der Subkategorie	Beschreibung der Subkategorie und Beispiele
Orientierung	<p>Sammeln von Informationen (Aufgabe, Situation, Person), die für die Ausführung von strategischem Verhalten relevant sind („Wo bin ich?“, „Was muss ich tun?“)</p> <ul style="list-style-type: none"> - Liest Instruktion oder Teile daraus - Erneutes Lesen der Instruktion oder Teile (Wiederholung) - Vergegenwärtigt sich Aufgabenstellung und spezifische Vorgaben der schriftlichen Instruktion („Kommentiert Instruktion, Aufgabe“) - Verschafft sich Überblick mit Seite Gliederung („Ich schau mir jetzt erst mal die Gliederung genauer an“) - Verschafft sich Überblick mit Seite Lernziele („Ich lese mal die Lernziele, um zu sehen, um was es geht“) - Verschafft sich auf andere Art Überblick (mit Überfliegen aller Webseiten) - Verschafft sich Überblick über Webseite (Überfliegen einer Webseite) - Beachtet Länge der Information („Das ist aber so viel“) - Betrachtet Art der Information („Das ist ja hauptsächlich Text“, „Abbildungen sind auch dabei“) - Schaut nach bekanntem bzw. äußert bekanntes Wissen; Wissensaktualisierung („Das habe ich in der Vorlesung schon gehört“, „Das kommt mir bekannt vor“, „Ich überlege mir, was ich zu diesem Gebiet schon weiß“, „Bandura, das ist mit den Lernexperimenten“) - Schaut nach unbekanntem bzw. äußert unbekanntes Wissen; Wissensaktualisierung („Das ist völlig neu für mich“, „Ich überlege mir, was ich zu diesem Gebiet noch nicht weiß“, „Löschen.....hm, habe nie was davon gehört“) - Äußert sich über eigene stabile Eigenschaften („Solche Aufgaben liegen mir nicht“ - nur in Orientierungsphase, wenn im weiteren Lernverlauf wieder solche Äußerungen gemacht werden, dann mit einer Motivationskategorie codieren) - Äußert sich über eigene situative Eigenschaften („Heute kann ich mich schlecht konzentrieren“) - nur in Orientierungsphase, wenn im weiteren Lernverlauf wieder solche Äußerungen gemacht werden, dann mit einer Motivationskategorie codieren
Planung	<p>Strategische Handlungsvorbereitung (beinhaltet Zieldefinition, Spezifikation bestimmter Lernstrategien, deren Eigenschaften und Konsequenzen, zieladaptive Strategieauswahl, Zeitplanung) („Wie gehe ich vor?“, „Was ist zu beachten?“)</p> <ul style="list-style-type: none"> - Plant Vorgehen („Ich überfliege erst mal alle Seiten, danach vertiefe ich“, „Ich überlege erst mal, welche Seiten (evtl. in welcher Reihenfolge) ich lernen muss“, „Jetzt werde ich erst mal schauen, was alles noch dazu gehört und dann schaue ich noch mal, was meine Aufgabe ist“) - Überlegt Kontrollmaßnahmen („Am Ende nehme ich mir noch Zeit, um das Wichtigste zu wiederholen“) - Teilt Zeit auf; berücksichtigt Zeit („Insgesamt habe ich 30 Minuten, davon werde ich in den ersten 5 Minuten“, „Da die Zeit knapp bemessen wird, gehe ich direkt auf Punkt 3: ich schaue erst mal, ob ich mich hier orientieren kann und versuche dann zum Verstärkerplan zu kommen“)
Ziel-spezifikation	<p>Formulierung und Spezifikation eines Lernziels, wobei zwischen globalen, allgemeinen Zielen und Teilzielen unterschieden wird. Das Ziel muss nicht (gleich) verfolgt werden („Was will ich?“)</p> <ul style="list-style-type: none"> - Formuliert/spezifiziert neue globale Ziele („Ich muss die Grundlagen zum operanten Konditionieren lernen“, Person stellt sich allgemeine Lernfragen) - Formuliert/spezifiziert neue <u>Subziele</u> („Ich muss wissen, was Kontingenz bedeutet“, Person stellt sich detaillierte Lernfragen) - Früher formulierte/spezifizierte globale Ziele werden erneut spezifiziert („Jetzt muss ich noch mal schauen, was Kognitivismus bedeutet“) - Früher formulierte/spezifizierte <u>Subziele</u> werden erneut spezifiziert („Jetzt muss ich doch mal schauen, was Kontingenz bedeutet“)

<p>Suche der Information</p>	<p>Informationen werden gesucht <i>(“Wo finde ich die nötige Information?“, “Wie komme ich dahin?“)</i></p> <ul style="list-style-type: none"> - Sucht Information (<i>“Jetzt suche ich mal die Seite mit den Verstärkerplänen“, “Ich suche so lange durch, bis ich das finde, ah hier“</i>) - Findet gesuchte Informationen (<i>“Da ist es ja“, “Endlich“, “Da hab ich ‘s ja“, “Wo ist hier Skinnerbox, ach hier“</i>) - Weiß nicht was genau suchen (<i>“Was suche ich eigentlich?“</i>) - Weiß nicht wo/wie suchen (<i>“Wo steht das denn?“, “Wie komme ich dahin?“</i>) - Klicken ohne genaue Suche geäußert zu haben
<p>Bewertung der Information</p>	<p>Auf- bzw. vorgefundene Informationen werden mit Blick auf ein Ziel bewertet <i>(“Ist die Information für Zielerreichung relevant“)</i></p> <ul style="list-style-type: none"> - Bewertet Information als zielrelevant (<i>“Das ist auch wichtig“</i>) - Bewertet Information als zielirrelevant (<i>“Die Personenbeschreibung von Skinner ist für die Aufgabenstellung nicht weiter wichtig“</i>) - Bewertet Information als irrelevant wegen der <i>“grauen Seiten“</i> (<i>“Ah, das ist eine dunkle Seite, das ist nicht wichtig“, “Obwohl das eine graue Seite ist, lese ich da mal rein“</i>) - Weiß nicht, ob Information relevant ist
<p>Evaluation</p>	<p>Kontrolle des Lernfortschritts während des Lernens und Lernerfolg am Ende <i>(“Habe ich mein Teil-/Ziel erreicht?“)</i></p> <ul style="list-style-type: none"> - Überprüft Verständnis/Lernerfolg (<i>“Jetzt schau ich mir mal die Beispiele an, ob ich das verstanden habe“</i>) - Rekapituliert Inhalte ohne Unterlagen (Webseiten) zur Hilfe zu nehmen (überwiegend am Ende des Lernprozesses) - Überprüft, ob Lernziele (auf Webseite 1) erreicht sind /Aufgabe erfolgreich ausgeführt wurde (<i>“Habe ich alle Punkte erarbeitet?“, “Fehlt noch was?“</i>) - Gleicht mit Aufgabenstellung ab (checkt Lernpunkte der Instruktion) - Checkt Mitschrift - Gleicht Mitschrift mit Instruktion ab
<p>Monitoring und Regulation</p>	<p>Überwachen und Steuern des Lernprozesses, betrifft alle Stufen des Lernvorgangs (Orientierung, Planung, Zielspezifikation, Suche und Bewertung, Evaluation und Verarbeitung). Kontrollmaßnahmen. Positive und negative Äußerungen</p> <ul style="list-style-type: none"> - Beachtet Zeit/-druck (<i>“Dafür bleibt keine Zeit“, “Dafür nehme ich mir ein paar Minuten Zeit“</i>) - Lernpunkt/Verarbeitung erfolgreich (<i>“Reizgeneralisierung – das habe ich jetzt“</i>) - <i>“Jetzt schau ich noch mal die Liste durch, das habe ich, das habe ich und das habe ich noch nicht gelernt“</i> - Äußert Erfolge bei Orientierung - Äußert Erfolge bei Planung - Äußert Erfolge bei Zielspezifikation (<i>“Ach ja, ich lerne also, was operante Konditionierung bedeutet und dabei auch die positive Verstärkung“</i>) - Äußert sich positiv zur Suche (<i>“Das werde ich schnell wiederfinden“</i>) - Äußert sich positiv zur Bewertung (<i>“Ich schau mal da rein und werde schnell sehn, ob das wichtig ist“</i>) - Äußert Schwierigkeiten, Probleme bei Orientierung (findet sich nicht mehr zurecht (<i>“Wo bin ich eigentlich?“, “Was muss ich noch/eigentlich tun?“ “Ist das umfassend genug“</i>)) - Äußert Probleme bei Planung (<i>“Weiß nicht, wie ich vorgehen/weitermachen soll“</i>) - Äußert Probleme bei Zielspezifikation (<i>“Weiß nicht, was ich lernen soll“</i>) - Äußert Suchprobleme (<i>“Das finde ich nie“</i>) - Äußert Bewertungsprobleme (<i>“Ist das überhaupt relevant?“</i>) - Äußert Behaltensprobleme (<i>“Das behalte ich nicht“, “Es ist alles so schwer zu merken“</i>) - <i>“Das kann ich mir merken“</i> - Äußert Elaborationsprobleme (<i>“Ich bringe das nicht zusammen“</i>) - <i>“Das hängt damit zusammen“</i> - Verstehensprobleme (<i>“Das verstehe ich nicht“, “Das ist irgendwie nicht so einleuchtend“</i>) - <i>“Das verstehe ich“, “Jetzt verstehe ich es“</i> - Organisationsprobleme (<i>“Die Struktur des Lernstoffs ist mir unklar“</i>) - <i>“Die Struktur des Lernstoffs ist mir klar“</i>

Oberkategorie: Kognition (Semantische Verarbeitung) Informationsaufnahme und Verarbeitung Strategische Aktivitäten: Dabei handelt es sich um jene Lernaktivitäten, die vom Probanden gezielt zur Aufnahme und Verarbeitung der Informationen (z.B. Webseiten, Tabellen, Bilder) eingesetzt werden.	
Erstmaliges Lesen	Beinhaltet das erstmalige Lesen von Webseiten, Tabellen und der Gliederung
Oberflächliches Wiederholen	Wiederholen im Sinne von Memorieren, Auswendiglernen - Wörtliches Wiederholen, einprägen von Fachbegriffen; Lernt Stoff oder Regeln, Fachbegriffe, Formeln auswendig - Wiederholtes Lesen der Ausschnitte im Sinne von Auswendiglernen
Oberflächliches Aufschreiben	Wortwörtliches Aufschreiben der Information
"Tiefe" Elaboration und Verstehen	Wiedergabe in eigenen Worten (Paraphrasieren) - Stellt Verknüpfungen her (auch mit eigenen Erfahrungen/Wissen), bildet Analogien, stellt Beziehungen zu anderen Aspekten her; versucht Ähnlichkeiten und Unterschiede zwischen Themen herzustellen; formuliert die wichtigsten Ideen - Erklärt sich selbst den Sachverhalt - Formuliert Beispiele (auch aus dem Alltagsleben) - Formuliert praktische Anwendungen - Stellt und beantwortet Fragen an das Lernmaterial - Kommentiert/hinterfragt kritisch den Lerninhalt (d.h. nicht abschätzig!), z.B. überlegt alternative Behauptungen, Schlussfolgerungen, formuliert eigene Ideen dazu, wägt Vor- und Nachteile ab, überprüft Schlüssigkeit der Behauptungen - Aktualisierung des Vorwissens und Verknüpfung des Vorwissens mit dem Lerninhalt
"Tiefe" Organisation	- Fertigt Schaubild, Map, Gliederung an - Schreibt wichtige Inhalte heraus - Fasst Inhalt schriftlich zusammen; Erstellt Zusammenfassung der Hauptideen - Stellt Fachausdrücke und Definitionen in eigener Liste zusammen - Erweitert Mitschrift
Oberkategorie: Motivation Motivationale-affektive Ausrichtung der Selbstregulation, die eine förderliche oder hemmende affektiv-motivationale Einbettung der kognitiven Aktivität erkennen lassen	
Motivation	Person äußert positive oder negative Bemerkungen bzgl. der Aufgabe - <i>„Aufgabe ist zu leicht“</i> - <i>„Aufgabe ist spannend/interessant“, „Ich bin gespannt, was noch kommt“</i> - <i>„Aufgabe ist relevant fürs Studium“</i> - <i>„Aufgabe ist zu schwierig“</i> - <i>„Aufgabe ist langweilig“</i> - <i>„Aufgabe ist irrelevant fürs Studium“</i> Person äußert sich positiv oder negativ über eigene Fähigkeiten, Erfolgszuversicht und Selbstvertrauen - Selbstvertrauen, erfolgszuversichtlich, erfolgsmotiviert (<i>„Das kann ich gut“</i>) - <i>„Ich muss mich mehr anstrengen, konzentrieren“</i> - Mangelnde eigene Fähigkeit für solche Aufgaben, geringes Selbstvertrauen, misserfolgsmotiviert - <i>„Das kann ich sowieso nicht“</i> - <i>„Ich kann mich nicht mehr konzentrieren“</i> - <i>„Ich schweife ab, das lenkt mich ab“</i>

	<p>Person macht positive oder negative Äußerungen über gegenwärtige Lernsituation, Auswirkungen des lauten Denkens und/oder Gestaltung der Webseiten</p> <ul style="list-style-type: none"> - „Lernen/ Lernsituation ist authentisch“ - „Lautes Denken erleichtert“ - „Webseiten sind gut gestaltet/strukturiert“ - „Lernen/ Lernsituation ist artifiziell“ - „Lautes Denken behindert“ - „Webseiten sind schlecht gestaltet/strukturiert“, „Gliederung ist unübersichtlich“
Oberkategorie: Restkategorie	
Interaktion mit Versuchsleiter	<p>Person stellt Fragen bezüglich der Durchführung oder Versuchsleiter erinnert an lautes Denken und Zeitlimitierung</p> <ul style="list-style-type: none"> - Proband fragt nach; „Ist das richtig“, „Darf ich Notizen machen“; Versuchsleiter antwortet - Versuchsleiter erinnert an lautes Denken; Proband antwortet - Versuchsleiter erinnert an Zeitvorgabe; „Noch 5 Minuten“, „Noch 1 Minute“; Proband antwortet
Verbesserungsvorschläge für Webseiten	<p>Person äußert Verbesserungsmöglichkeiten in Bezug auf den Aufbau und die Organisation der Webseiten</p> <ul style="list-style-type: none"> - „Eine Suchfunktion wäre gut“ - „Mehr Farbe wäre besser“
Handhabung des Programms	<p>Person bewertet allgemein den Umgang mit dem Lernprogramm</p>
Kommentar zur Navigation	<p>Kommentar zur Navigation ohne inhaltlichen Bezug und gleichzeitige Ausführung des Gesagten</p> <ul style="list-style-type: none"> - „Ich klicke jetzt darauf“, „Ich fahre mal runter“ - „Ich klick das jetzt an“
Sprechpause	<p>Sprechpause die länger als 5 Sekunden dauert</p>

Note. English translations of the modified coding schemes are displayed in the journal articles.

Appendix 2. Screenshot of the learning environment including a metacognitive prompt

The screenshot shows a web-based learning environment titled "Lerntheorien" with a "Linear Navigation" header. On the left is a "Navigation Menu" with items: Lernziele, Einführung, Theoretische Grundlagen des Lernens, Pädagogische Anwendung der Lerntheorien: Verhaltensmodifikation, Zusammenfassung, Bezüge, Literatur und Links, and Glossar. The main content area is titled "Gründe für meinen Seitenwechsel" (Reasons for my page change) and contains a list of seven reasons with checkboxes: Orientierung, Zielsetzung, Planung, Kontrolle des Verständnisses, Überwachung des Lernens, Steuerung des Lernens, and Überprüfung der Lernzielerreichung. Below the list is a "Senden" button. A yellow "Prompt" label is placed over the bottom right of this window. The background content includes sections for "Verstärkungspläne" (Reinforcement plans) and "Quotenplan" (Quota plan), with text discussing interval and quota plans. A "Content" label is placed at the bottom right of the main content area.

Note. The main elements are labeled here for a better understanding of the environment, however it was not labeled in the learning sessions. Participants were asked to select at least one reason for node selection by choosing among a list of strategic reasons presented in the prompt window. The list comprised *orientation*, *goal-setting*, *planning*, *checking of understanding*, *monitoring of learning*, *control of learning*, and *evaluation of goal attainment*.

Appendix 3. Learning task and instructions of the first learning session (in German)

Lernen in netzbasierten Lernumgebungen

Jetzt beginnt der eigentliche Lernteil. Bitte betrachten Sie hierzu folgenden Fall:

Ein Lehrer klagt über massive Disziplinprobleme seiner Klasse. Die Kinder rufen durcheinander, fallen sich gegenseitig ins Wort, bleiben nicht auf ihren Plätzen sitzen, arbeiten im Unterricht nicht mit und vernachlässigen ihre Hausarbeiten. Seine Kollegin hat mit ihrer Klasse keine Probleme, zumindest nicht mehr. Sie hat vor ein paar Wochen einen Kurs zur pädagogischen Verhaltensmodifikation besucht und dabei gelernt, wie auf der Grundlage der behavioristischen Lerntheorien solche Disziplinprobleme reduziert werden können.



Über solche Themen handelt unser web-basiertes Lernmaterial. Es behandelt die behavioristischen Lerntheorien und deren Anwendung in pädagogischen Situationen.

Ihre Aufgabe ist es nun, das web-basierte Lernmaterial sorgfältig durchzuarbeiten, so dass Sie in der Lage sind, einem Studienkollegen oder -kollegin die grundlegenden Konzepte der operanten Konditionierung zu beschreiben und zu erklären, wie beispielsweise:

- Verstärkerplan
- Premack-Prinzip
- Verstärkung und Bestrafung
- Experimente mit der Skinnerbox
- Das Prinzip der operanten Konditionierung
- Reizgeneralisierung und -diskrimination
- Kontingenz

Sie haben hierfür insgesamt 40 Minuten Zeit.

WICHTIG:

Die Zeit ist knapp bemessen. Sie reicht aus, um die genannten Konzepte zu lernen und zu verstehen. Sie reicht allerdings nicht aus, um die gesamte Lektion erschöpfend durchzuarbeiten. Hierzu benötigt man mindestens 2 Stunden. Aus Zeitgründen konzentrieren Sie sich also bitte hauptsächlich auf die Seiten, welche die theoretischen Grundlagen der operanten Konditionierung behandeln.

Wenn Sie wollen, können Sie sich während des Lernens Notizen machen, die Sie jedoch später nicht mehr verwenden dürfen.

Bitte sprechen Sie nun alles laut aus, was Ihnen durch den Kopf geht. Wenn Sie Textstellen lesen, lesen Sie sie bitte laut vor. Sollten Sie eine Weile schweigen, werden wir Sie auffordern, Ihre Gedanken wieder laut zu äußern.

Appendix 4. Learning task and instructions of the second learning session (in German)

Lernen in netzbasierten Lernumgebungen

Jetzt beginnt der eigentliche Lernteil. Bitte betrachten Sie hierzu folgenden Fall:



Ein Lehrer klagt über massive Motivationsprobleme seiner Klasse. Die Kinder arbeiten im Unterricht nicht mit und vernachlässigen ihre Hausarbeiten. Seine Kollegin hat mit ihrer Klasse keine Probleme, zumindest nicht mehr. Sie hat vor ein paar Wochen einen Kurs zu pädagogischen Motivationsförderprogrammen besucht und dabei gelernt, wie auf der Grundlage der klassischen Motivationspsychologie solche Motivationsprobleme reduziert werden können.

Über solche Themen handelt unser web-basiertes Lernmaterial. Es behandelt die behavioristischen Lerntheorien und deren Anwendung in pädagogischen Situationen.

Ihre Aufgabe ist es nun, das web-basierte Lernmaterial sorgfältig durchzuarbeiten, so dass Sie in der Lage sind, einem Studienkollegen oder -kollegin die grundlegenden Konzepte der Motivationspsychologie zu beschreiben und zu erklären, wie beispielsweise:

- **Risikowahlmodell**
- **Leistungsmotiv**
- **Motive und Motivation**
- **Anschlussmotiv**
- **Bedürfnispyramide**
- **Extrinsische und intrinsische Motivation**

Sie haben hierfür insgesamt 40 Minuten Zeit.

WICHTIG:

Die Zeit ist knapp bemessen. Sie reicht aus, um die genannten Konzepte zu lernen und zu verstehen. Sie reicht allerdings nicht aus, um die gesamte Lektion erschöpfend durchzuarbeiten. Hierzu benötigt man mindestens 2 Stunden. Aus Zeitgründen konzentrieren Sie sich also bitte hauptsächlich auf die Seiten, welche die theoretischen Grundlagen der Motivationspsychologie behandeln.

Wenn Sie wollen, können Sie sich während des Lernens Notizen machen, die Sie jedoch später nicht mehr verwenden dürfen.

Bitte sprechen Sie nun alles laut aus, was Ihnen durch den Kopf geht. Wenn Sie Textstellen lesen, lesen Sie sie bitte laut vor. Sollten Sie eine Weile schweigen, werden wir Sie auffordern, Ihre Gedanken wieder laut zu äußern.

Part 2: Journal Articles

Contents

- Analysis 1: Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques
- Analysis 2: Evaluating the Impact of Instructional Support Using Data Mining and Process Mining: A Micro-Level Analysis of the Effectiveness of Metacognitive Prompts
- Analysis 3: Using Process Mining to Examine the Sustainability of Instructional Support: How Stable are the Effects of Metacognitive Prompting on Self-Regulatory Behavior?

Analysis 1:

Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques

Published in the Journal of Learning Analytics:

Sonnenberg, C., & Bannert, M. (2015). Discovering the effects of metacognitive prompts on the sequential structure of SRL-processes using process mining techniques. *Journal of Learning Analytics*, 2(1), 72–100.

Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques

Christoph Sonnenberg and Maria Bannert

University of Wuerzburg, Instructional Media, Germany

christoph.sonnenberg@uni-wuerzburg.de

ABSTRACT: According to research examining self-regulated learning (SRL), we regard individual regulation as a specific sequence of regulatory activities. Ideally, students perform various learning activities, such as analyzing, monitoring, and evaluating cognitive and motivational aspects during learning. Metacognitive prompts can foster SRL by inducing regulatory activities, which, in turn, improve the learning outcome. However, the specific effects of metacognitive support on the dynamic characteristics of SRL are not understood. Therefore, the aim of our study was to analyze the effects of metacognitive prompts on learning processes and outcomes during a computer-based learning task. Participants of the experimental group (EG, n=35) were supported by metacognitive prompts, whereas participants of the control group (CG, n=35) received no support. Data regarding learning processes were obtained by concurrent think-aloud protocols. The EG exhibited significantly more metacognitive learning events than did the CG. Furthermore, these regulatory activities correspond positively with learning outcomes. Process mining techniques were used to analyze sequential patterns. Our findings indicate differences in the process models of the EG and CG and demonstrate the added value of taking the order of learning activities into account by discovering regulatory patterns.

KEYWORDS: self-regulated learning, metacognitive prompting, process analysis, process mining, think-aloud data, HeuristicsMiner algorithm

1 INTRODUCTION

Recent research in the field of self-regulated learning (SRL) has moved to a process-orientated or event-based view to investigate how learning processes unfold over time and how scaffolds influence the dynamic nature of regulatory activities. Two recent special issues indicate the importance of investigating sequential and temporal patterns in learning processes and present new methodological contributions for the analysis of time and order in learning activities (Martin & Sherin, 2013; Molenaar & Järvelä, 2014). Technical advances allow the recording of learning-related behaviour on a very detailed level and largely unobtrusively for learners (e.g., Azevedo et al., 2013; Winne & Nesbit, 2009). As such, researchers have focused more on behavioural process data and less on measures of aptitude (Azevedo, 2009; Bannert, 2009; Veenman, van Hout-Wolters, & Afflerbach, 2006). When focusing on process data, differences among learners are explained on the event level with respect to regularities and patterns (Winne & Perry, 2000), allowing researchers to gain new insights into the process of learning.

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

Process analysis methods beyond the variable-centred *coding and counting* approach (Kapur, 2011) can provide valuable information on the specific effects of scaffolds (e.g., metacognitive prompts) and are able to inform researchers about how to optimize an applied supporting strategy further (e.g., Jeong et al., 2008; Johnson, Azevedo, & D’Mello, 2011; Molenaar & Chiu, 2014). Moreover, findings on the sequential and temporal structure of SRL processes can provide knowledge for the development of SRL theories on the micro-level (Molenaar & Järvelä, 2014).

Our approach applies the techniques of process mining (Trčka, Pechenizkiy, & van der Aalst, 2010) on process data obtained by concurrent think-aloud protocols (Ericsson & Simon, 1993). For example, we have compared process patterns of students with high versus low learning performance in a recent study (Bannert, Reimann, & Sonnenberg, 2014) and demonstrated that process mining techniques can reveal differences in the sequential patterns of regulatory processes. Now, we are investigating the effects of *metacognitive prompts* (Bannert, 2009) by means of an in-depth analysis using process mining techniques. An analysis of differences in the process models between students supported by metacognitive prompts and students without prompts can provide information on how to promote beneficial regulatory patterns and thereby improve learning.

The paper is structured as follows: First, we introduce research focusing on the support of SRL through metacognitive prompts. Second, we describe SRL models that emphasize the importance of different learning events and event patterns. Third, some of the foundations of analyzing learning processes with process mining are introduced. Fourth, we analyze process data from coded think-aloud protocols of an experimental study. In addition to the traditional frequency-based approach, the relative arrangement of learning activities is taken into account using process mining techniques. Finally, the results of these analyses are compared, and the effects of metacognitive support on the sequential structure of SRL processes are discussed.

2 THEORETICAL BACKGROUND

2.1 Metacognitive Support through Prompts

Current research in metacognition and SRL shows that learners often do not spontaneously use metacognitive skills during learning, which in turn leads to poorer learning outcomes (e.g., Azevedo, 2009; Bannert & Mengelkamp, 2013; Greene, Dellinger, Tüysüzoglu, & Costa, 2013; Winne & Hadwin, 2008; Zimmerman, 2008). The students’ awareness and control of their own manner of learning is important, especially in technology-enhanced and open-ended learning settings (Azevedo, 2005; Lin, 2001; Lin, Hmelo, Kinzer, & Secules, 1999). In most open-ended learning environments, it is constantly necessary to make decisions on what to do and where to go next and to evaluate the retrieved information with respect to current learning goals (Schnotz, 1998). Therefore, the general purpose of our research is to provide metacognitive support for hypermedia learning through metacognitive prompts.

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

Instructional prompts are scaffolds that induce and stimulate students' cognitive, metacognitive, and motivational activities during learning (Bannert, 2009). The underlying assumption is that students have already acquired these processes, but they do not recall or execute them spontaneously in a specific learning situation (*production deficit*; Veenman et al., 2006; Veenman, 2007). Metacognitive prompts aim at inducing regulatory activities such as orientation, goal specification, planning, monitoring and control, and evaluation strategies (Bannert, 2007; Veenman, 1993) by asking students to reflect upon, monitor, and control their own learning process.

Previous research has demonstrated beneficial effects from metacognitive prompting (e.g., Azevedo, Cromley, Moos, Greene, & Winters, 2011; Ge, 2013; Johnson et al., 2011; Lin & Lehman, 1999; Veenman, 1993; Winne & Hadwin, 2013). For example, Lin and Lehman (1999) prompted students to give reasons for their actions to increase the awareness of their own strategies by utilizing a pop-up window at certain times in a computer-based simulation environment (e.g., "What is your plan for solving the problem?"). Their findings showed significantly higher performance on contextually dissimilar problems (i.e., far transfer performance) for the students supported by prompts. Based on an analysis of think-aloud data, Johnson et al. (2011) showed that prompts given by a human tutor during learning in a hypermedia learning environment influenced the deployment of regulatory processes and temporal dependencies. Compared to a control group, the externally assisted condition also achieved a better learning outcome.

In previous experiments, we investigated the effects of different types of metacognitive prompts during hypermedia learning (Bannert & Mengelkamp, 2013; Bannert & Reimann, 2012). The prompts stimulated or even suggested appropriate metacognitive learning activities for university students during a hypermedia learning session lasting approximately 40 minutes. For example, in one of our experiments, students were prompted after each navigational step in a learning environment to verbalize the reasons why they had chosen the next step (so-called *reflection prompts*; Bannert, 2006). Overall, the findings confirm the positive effects of all investigated types of metacognitive prompts on transfer performance and the use of learning strategies during learning.

Our most recent work (Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015) investigates the effects of a new type of metacognitive prompt (so-called *self-directed metacognitive prompts*) on navigation behaviour and learning outcomes. In summary, the findings show that such prompts enhance strategic navigation behaviour (i.e., students visited relevant webpages significantly more often and spent more time on them) and transfer performance (i.e., students performed better at applying knowledge of basic concepts to solve prototypical problems compared with a control group). In addition, learner characteristics (e.g., prior domain knowledge or verbal abilities) were obtained by questionnaires, but they had no effects as covariates in our analyses. The present study extends this contribution by focusing on the sequential analysis of coded think-aloud data obtained during learning.

Despite the findings about the general effectiveness of metacognitive prompts, the specific effects of prompts on learning processes remain unexplained. More precisely, a closer look at the effects of

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

prompts on the sequential and temporal structure of SRL processes is necessary (e.g., Jeong et al., 2008; Johnson et al., 2011). Understanding this process at the micro-level would allow researchers to better design metacognitive support. For example, regulatory patterns associated with successful learning but that could not be fostered by metacognitive prompts could be identified. Subsequently, the metacognitive support could be adapted by taking information about these patterns into consideration. Therefore, we focus on analyzing the sequential order of learning activities obtained by concurrent think-aloud protocols during learning.

2.2 Regulatory Patterns in SRL

Boekaerts (1997) describes SRL as a complex interaction of cognitive, metacognitive, and motivational regulatory components. With respect to assumptions in SRL models (e.g., Winne & Hadwin, 2008; Zimmerman, 2008), successful studying corresponds with an active performance of different regulatory activities during learning. These regulatory activities include employing orientation to obtain an overview of the learning task and resources, planning the course of learning, monitoring and controlling all learning steps, and evaluating the learning product. Research in SRL has confirmed that successful learning is associated with the active deployment of these regulatory activities (e.g., Azevedo, Guthrie, & Seibert, 2004; Bannert, 2009; Johnson et al., 2011; Moos & Azevedo, 2009).

Most SRL models share the common assumption of a time-ordered sequence of regulatory activities, although they do not imply a strict order (Azevedo, 2009). Usually, three cyclic phases of forethought, performance, and reflection (Zimmerman, 2000) are distinguished. The forethought phase comprises task analysis, goal setting, and strategic planning. During the performance phase, self-observations for adaptations (monitoring) and control strategies (self-instruction or time management) are deployed. Finally, the reflection phase includes self-judgments and self-reactions, which, in turn, can inform the next forethought phase. The COPES model (Winne & Hadwin, 2008) represents a more elaborate description of regulatory processes in terms of an information-processing model. Here, learning occurs in three phases, namely, task definition, goal setting and planning, and studying tactics, and a fourth optional phase, adaptations to metacognition. In addition, monitoring and control are crucial elements in the COPES model. Monitoring is used to detect differences between current conditions (e.g., learning progress) and standards (e.g., predefined learning goals), which, in turn, activates control processes to reduce discrepancies (e.g., engaging more intensively in a certain topic).

2.3 Microanalysis Using Process Mining Techniques

In a recent study (Bannert et al., 2014), we suggested process mining (PM) as a promising method in SRL research. PM allows researchers to describe and test process models of learning that incorporate an event-based view and that are at the high end of process granularity. These process models are able to represent the workflow of activities (van der Aalst, Weijters, & Maruster, 2004). Therefore, we argue that PM is adequate for investigating regulatory patterns based on process assumptions conceptualized in SRL research, as described in the previous section. For example, PM or data-mining techniques can

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

extract patterns by analyzing process data (e.g., think-aloud protocols or log files), and the resulting patterns can be compared to assumptions of SRL models (e.g., the assumption of a time-ordered sequence of regulatory activities in successful learning or the concept of dynamic and cyclic patterns). Therefore, the observed behaviour in process data could be aligned with SRL models.

PM is an approach that can be used in the context of *Educational Data Mining* (Romero, Ventura, Pechenizkiy, & Baker, 2010). In this context, PM represents student activities as a process model derived from their log traces while using a computer-based learning environment. In general, PM methods allow researchers to discover process models inductively from activity sequences stored in an event log, test process models through conformance checking with additional event data, or the extension of existing models (Trčka et al., 2010). Especially in the context of computer-supported learning research, PM techniques are increasingly used to study learning from a process-oriented perspective (Reimann & Yacef, 2013; Schoor & Bannert, 2012). For example, PM techniques can be applied to modelling sequences of learning activities that have been recorded in log files or coded think-aloud data.

By using PM techniques to discover process patterns in SRL activities, we assume that the present process data — comprising temporally ordered event sequences — is directed by one or more mental processes, with each set of processes corresponding to a process model. Hence, a process model represents a system of states and transitions that produced the sequence of learning events. Usually, the performance of this system is driven by a plan for action. In the context of SRL, this plan can be a learning strategy or an external resource provided to the learner (e.g., prompts). A process model is able to express a holistic view of a process by modelling a system comprising states and transitions rather than a process-as-sequence perspective (Reimann, 2009).

With respect to related approaches, hidden Markov models (e.g., Jeong et al., 2008) also allow for expressing the holistic nature of a process by taking into account the entire sample of behaviour. However, this approach uses time-consuming iterative procedures; generally, the researcher has to pre-define the appropriate number of states, and the interpretation of the output model is often difficult (van der Aalst, 2011). There are, however, approaches for automatically selecting the appropriate number of states using the Bayesian Information Criterion (e.g., Li & Biswas, 2002). Additionally, hidden Markov models, as well as simple transition graphs and other low-level models, represent a lower abstraction level than the PM notation language (e.g., inability to represent concurrency, which typically results in more complex models). Finally, PM techniques have the advantage of explicitly dealing with noise (i.e., exceptional or infrequent behaviour), which is necessary when analyzing real-life event traces. For these reasons, we recommend PM techniques for analyzing sequences of learning activities (see Bannert et al., 2014 for more information regarding the comparison to other process analysis methods).

2.4 Research Questions and Hypotheses

Metacognitive prompts ask students explicitly to reflect, monitor, and control their own learning

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

process. They focus students' attention on their own thoughts and on understanding the activities in which they are engaged during learning (e.g., Bannert, 2006; Hoffman & Spatariu, 2011). Hence, it is assumed that prompting students to monitor and evaluate their own manner of learning will allow them to activate their repertoire of metacognitive knowledge and learning strategies, which will consequently enhance their learning process and learning outcome. However, according to previous work and research on metacognitive prompting, the use of metacognitive prompts has to be explained and practiced in advance to guarantee an adequate application during learning (e.g., Bannert, 2007; Veenman, 2007). Based on the findings of studies investigating the effects of metacognitive prompts (e.g., Azevedo et al., 2004; Bannert, 2009), we expect that students supported by metacognitive prompts will engage in more regulatory activities, as obtained by coded think-aloud protocols. Moreover, scaffolded SRL processes should result in better learning performance; that is, a positive effect on learning outcomes mediated by improved regulatory behaviour. Whereas these two hypotheses are based on a variable-centred view of learning processes, we assume that an event-centred analysis that takes into account the relative arrangements of multiple learning activities can provide additional information about the sequential structure of the regulatory behaviour induced by the prompts (e.g., a sequence of orientation activities, searching for relevant information, cognitive processing, and evaluation of progress are typically executed). Therefore, the effectiveness of metacognitive prompts can be analyzed on a micro-level, and the results can be used to derive implications for the improvement of metacognitive support. In detail, the following research questions are addressed in the present study:

1. Does metacognitive prompting during learning influence SRL processes by engaging students in more metacognitive learning events?
2. Does the number of metacognitive learning events mediate the effect of metacognitive prompting on learning outcomes?
3. Which sequential patterns of SRL activities are induced by metacognitive prompting compared to a control group without support?

2.5 Process Mining Using the HeuristicsMiner Algorithm

To analyze the relative arrangement of learning activities, we employed the PM approach (Trčka et al., 2010). The basic idea of PM is to use an event log to generate a process model describing this log inductively (process discovery). Furthermore, theoretical models or empirically mined models can be compared to event logs (conformance checking), and existing models can be extended (model extension). Fluxicon Disco Version 1.7.2 (2014) software was used for data preparation. Next, the event log was imported into the ProM framework Version 5.2 (2008), and PM was conducted. The ProM framework comprises a variety of PM algorithms that can be assigned to the functions of discovery, model checking, or model extension. For our analysis, we used the HeuristicsMiner algorithm (Weijters, van der Aalst, & de Medeiros, 2006) for process discovery.

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

We selected the HeuristicsMiner algorithm based on a comparison of seven state-of-the-art process discovery algorithms on the dimensions of accuracy and comprehensibility, provided by de Weerd, de Backer, Vanthienen, and Baesens (2012). Accuracy is defined as the capability of a sound capturing of behaviour in an event log, omitting over- and underfitting (i.e., a process model should balance between generality and precision). Comprehensibility comprises simplicity and structuredness of the resulting process models, and thereby determines the complexity and ease of interpretation of the output. For the first time, real-life event logs containing log data from different information systems were used for benchmarking PM algorithms. Among the seven algorithms, the HeuristicsMiner was the best technique for the real-life logs used and the authors conclude that “HeuristicsMiner seems the most appropriate and robust technique in a real-life context in terms of accuracy, comprehensibility, and scalability” (De Weerd et al., 2012, p. 671). In the following, we explain the general principle and functionality of this algorithm in more detail.

2.6 General Principle of the HeuristicsMiner

The general principle of the HeuristicsMiner algorithm is to take into account the sequential order of events for mining a process model that represents the control flow of an event log (Weijters et al., 2006). The event log containing case IDs, time stamps, and activities represents the data input. Based on this input, the algorithm searches for causal dependencies between activities by computing a dependency graph that indicates the certainty of a relation between two activities (e.g., event *a* is followed by event *b* with a certainty of 0.90). Finally, a so-called *heuristic net* is generated as an output model that constitutes a visual representation of the dependencies among all activity classes in the event log. The resulting process model can be adjusted by setting thresholds for the inclusion of relations in the heuristic net (for more details on parameter settings, see below).

In addition, the HeuristicsMiner is based on two main assumptions. First, each non-initial activity has at least one other activity that triggers its performance, and each non-final activity is followed by at least one dependent activity. This assumption is used in the so-called *all activities connected heuristic* (Weijters et al., 2006). Second, the event log contains a representative sample of the observed behaviour, which usually contains a certain amount of noise, especially if traces of human behaviour are stored in the event log. For example, in our study, a perfect trace of verbal utterances for all performed learning steps is unlikely. Therefore, the event log contains noise caused, for example, by a missing learning step that was not uttered or by disagreement among the raters during the coding procedure. It must be noted that there is also noise in other types of data (e.g., log file data). Consequently, an analysis method is needed that can abstract from noise and that can concentrate on the main relations among learning activities. It is a specific feature of the HeuristicsMiner to be robust to noise in the data. This is the main reason for the appropriateness of applying this PM algorithm to our event log.

An additional advantage of the HeuristicsMiner algorithm is that the mined model (heuristic net) can be converted into a formal *petri net*. A petri net can be described as a bipartite directed graph with a finite set of places, a finite set of transitions, and two sets of directed arcs, from places to transitions and from

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

transitions to places (Reisig, 1985). Thus, the resulting process model can be used as input for other PM algorithms, and it can be utilized in subsequent analyses (e.g., conformance checking between the model and a new event log). In contrast, the output model of another promising process discovery algorithm within the ProM framework that we used in previous process analyses (Bannert et al., 2014; Schoor & Bannert, 2012), called the *Fuzzy Miner* (Günther & van der Aalst, 2007), cannot be converted into a petri net (De Weerd et al., 2012). Therefore, the HeuristicsMiner was the first choice for our present analysis.

2.7 Functionality and Application of the HeuristicsMiner

Considering its functionality, the HeuristicsMiner algorithm uses several parameters that guide the creation of a process model and that can be adjusted to set the level of abstraction from noise and low-frequency behaviour. First, a frequency-based metric is used to determine the degree of certainty of a relation between two events, A and B , based on an event log. The dependency values, ranging between -1 and 1 , between all possible combinations of events are computed using the following formula (Weijters et al., 2006, p. 7):

$$A \Rightarrow_w B = \left(\frac{|a \succ_w b| - |b \succ_w a|}{|a \succ_w b| + |b \succ_w a| + 1} \right)$$

Based on an event log W , the certainty of a dependency relation between two events, $A \Rightarrow_w B$, is computed using the number of times event a is followed by event b , subtracted from the number of times event b is followed by event a , and divided by the number of occurrences of these two relations, plus 1. The number of correct (a follows b) and incorrect (b follows a) event sequences influences the dependency value by the $+1$ in the denominator. For example, an event log containing only correct sequences (a is always followed by b , but never vice versa), but with a low frequency of five observations, results in a certainty of $5/6 = 0.83$, whereas in the case of a high frequency of 50 observations, the certainty of a dependency relation between a and b would be $50/51 = 0.98$.

Moreover, the computed dependency values are used to construct a heuristic net (i.e., the output model). However, not all dependency relations are kept in the process model. Instead, the HeuristicsMiner algorithm concentrates on the main causal dependencies and abstracts from noise and low-frequency behaviour. At first, the all activities connected heuristic is applied. Therefore, only the best candidates (with the highest $A \Rightarrow_w B$ values) regarding the dependency values are kept in the output model. Second, three threshold parameters are used for the selection of further dependency relations. The *dependency threshold* determines the cut-off value for the inclusion of dependency relations in the output model. Furthermore, the *positive observation threshold* defines the minimum number of necessary observed sequences. Finally, the *relative to best threshold* determines that only additional dependency relations with a lower difference to the best candidate are included in the output model. We refer to Weijters et al. (2006) for more information about these threshold parameters.

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

In our analysis, the threshold parameters were kept at their default values of dependency threshold = 0.9, positive observation behaviour = 10, and relative-to-best-threshold = 0.05. As explained above, these threshold parameters can be used to adjust the level of abstraction of the output model. For example, reducing the cutoff-values would result in additional dependency relations in the model and thus increase the complexity. However, there were no reasons for changing the default-values in our case.

Furthermore, the HeuristicsMiner algorithm can also address *short loops* of lengths one (e.g., ACCB) and two (e.g., ACDCDB) as well as *long distance dependencies*; that is, a dependency based on choices made in other parts of the process model. Moreover, the algorithm considers *AND-relations* (two events are executed concurrently) and *OR-Relations* (e.g., either event *b* or event *c* can be executed after event *a*) to construct the heuristic net.

In general, searching for an optimal process model based on a present event log can be challenging, especially if there is a certain amount of noise and less-frequent behaviour in the data. Therefore, it is possible to compare the resulting process model with the event log using a *fitness* value (Rozinat & van der Aalst, 2008). The fitness indicates the gap between the observed behaviour, that is, the set of event sequences in the log, and the mined process model.

By applying the HeuristicsMiner algorithm to our event log, we assume that the present set of sequences of learning events is caused by one or multiple underlying processes. However, it might be possible that there is a high variety in SRL activities within the sample. In this case, using very robust algorithms such as the HeuristicsMiner can result in over-generalization (underfitting); that is, the mined model allows for much more behaviour than what is actually observed (De Medeiros et al., 2008). Therefore, the event log could be modelled more precisely by generating different process models for subsets of participants instead of a single model for all cases. This approach is called *trace clustering*, which can improve the discovery of process models (De Weerdt, vanden Broucke, Vanthienen, & Baesens, 2013; Greco, Guzzo, Pontieri, & Saccà, 2006). A plug-in has been implemented in the ProM framework that combines the HeuristicsMiner algorithm with a trace clustering procedure, namely, *DWS mining* (Disjunctive Workflow Schema; De Medeiros et al., 2008). The basic idea of DWS mining is to split the log into clusters iteratively until the mined process model for each cluster reaches high precision. A process model has a high precision if it only allows for behaviour that was observed in the event log. Consequently, a cluster is further partitioned if the mined model allows for more behaviour than is expressed by the cases within this cluster. For more information on the DWS mining plugin, refer to De Medeiros et al. (2008). In our analysis, we kept the default parameter settings for clustering the log traces.

3 METHOD

The present study extends a previous contribution (Bannert et al., 2015) that investigates the effects of metacognitive prompting on navigation behaviour and learning outcome referring to the same

participants, but to different research questions and to mostly different data.

3.1 Sample and Research Design

A total of n=70 undergraduate students from a German university participated in the study (mean age = 20.07, SD = 1.88, 82.9% female). All participants were either majoring in media communications or in human–computer systems. Participants were recruited via an online recruitment system administered by our institute, and each student received 40 Euros (approximately \$47 USD) for participating.

Altogether, the experimental study was based on a between-subject design and comprised two sessions. In the first session, learner characteristics were obtained as potential covariates (e.g., prior domain knowledge), especially in the case of an unbalanced distribution of characteristics among the groups by randomization (which is possible for the relatively small sample size). Approximately one week later, the participants were randomly assigned to either the experimental group (n=35) or the control group (n=35) and individually participated in hypermedia learning. The experimental group learned with metacognitive prompts, whereas the control group learned without prompts. Figure 1 presents an overview of the research design.

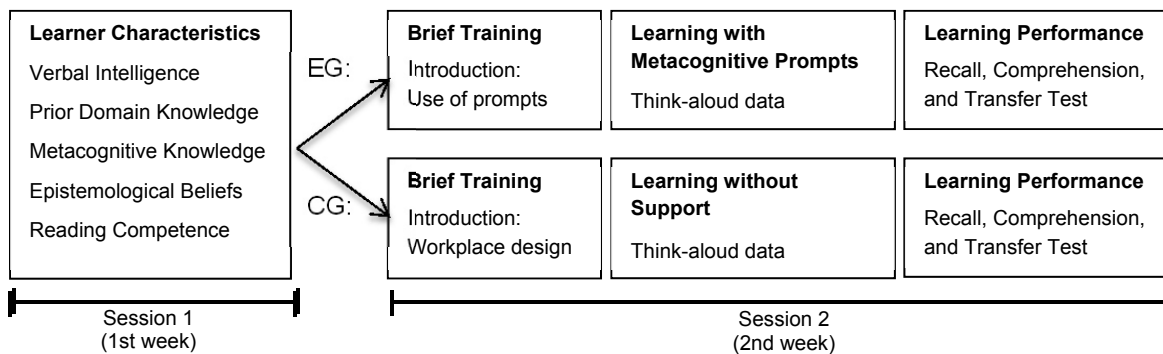


Figure 1. Research design

3.2 Learning Material and Performance Measurement

3.2.1 Learning Environment and Metacognitive Prompts

The learning material comprised a chapter on the topic of learning theories (classical conditioning, operant conditioning, and observational learning) presented in a hypermedia learning environment. For example, the content of one node included a description of the Skinner-box with reference to the concept of operant conditioning, and illustrated with a picture. In total, the material comprised 50 nodes with 13,000 words, 20 pictures and tables, and 300 hyperlinks. Within this chapter, the material relevant for the learning task comprised 10 nodes with 2,300 words, 5 pictures and tables, and 60 hyperlinks. The remaining pages were not relevant for the learning task. These pages included overviews, summaries, and pages with information on concepts not relevant for the learning goals. The Flesch-Kincaid grade-level score of the complete learning material was 19.01.

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

Navigation in the learning environment was possible in four different ways: 1) a hierarchical navigation menu, 2) a next-page and previous-page button on top of each page, 3) the backward- and forward-button of the browser, and 4) hyperlinks embedded in the text.



Figure 2: Learning environment with metacognitive prompt. Students are asked to select one or more reasons for node selection in a hypermedia learning environment by choosing among a list of strategic reasons (e.g., orientation, goal specification, planning) eight times during learning

Support via metacognitive prompts was implemented in the learning environment. A prompt appeared in the form of a pop-up window placed in the middle of the screen eight times during learning. Each prompt contained a list of strategic reasons for node selection. At least one reason had to be selected before continuing with learning. Figure 2 shows the hypermedia learning environment with a metacognitive prompt.

3.2.2 Knowledge Tests

Learning performance was measured with three knowledge tests on different levels based on Bloom's taxonomy of cognitive learning (Bloom, 1956). The measurement comprised a free recall test, a comprehension test, and a transfer test. In the free recall test, students were instructed to write down all basic concepts they could remember. The comprehension test, which assessed knowledge of facts, comprised 22 multiple-choice items, each with one correct and three incorrect answers. Transfer was

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

measured by asking students to apply basic concepts and knowledge of facts to solve eight prototypical problems in educational settings that were not explicitly addressed in the learning material (maximum score: 40 points). For example, students were asked to explain how a teacher should behave in response to a described classroom discipline problem based on the principles of operant conditioning. The answers of the participants were rated on a researcher-developed rating scale by two research assistants (Cohen's Kappa = .84). In case of disagreement among the raters, one of the authors determined the final score. More information on the learning material and knowledge tests used in our prompting studies is provided by Bannert and Reimann (2012) and by Bannert and Mengelkamp (2013).

3.3 Procedure

Approximately one week before the learning session, the learner characteristics *verbal intelligence*, *prior domain knowledge*, *metacognitive strategy knowledge*, *epistemological beliefs*, and *reading competency* were measured by questionnaires. More information on the instruments is provided in Bannert et al. (2015).

The learning session started with an introduction phase. First, the navigation in the hypermedia learning environment was explained by the experimenter. Then, the participant was asked to practice all possible ways of navigating the learning environment by using a practice lesson. After that, a series of exercises had to be performed in the practice lesson using concurrent thinking aloud during the task. The experimenter provided feedback and, if necessary, additional exercises until the participant firmly mastered the think-aloud technique.

Subsequently, the students in the experimental group received an introduction (approximately 10 minutes) to the use of metacognitive prompts, which included a description of the importance of reflecting on one's own learning steps, an explanation of the reasons for strategic node selection listed in the prompts, and the correct use of the prompts. It is necessary to explain the use of metacognitive prompts to the students to guarantee adequate application during learning (e.g., Veenman, 2007). After that, they were instructed to configure the prompts by arranging the list with reasons for node selection and by defining eight time stamps when the prompts should be presented during learning. To keep the workload for both groups equivalent, participants in the control group received an introduction to workplace design, which is not relevant for the stimulation of metacognitive learning activities. Instead of prompt configuration, they were asked to arrange their workplace before learning. Both introductions were realized by the experimenter using a sheet of instructions and advice visible to the participant.

Following this, the learning phase started. All participants received a sheet with their learning task, which instructed them to learn the basic concepts of operant conditioning within 40 minutes. Moreover, they were provided with a list of seven example concepts that had to be learned (e.g., Skinner Box, Positive Reinforcement). Students in the experimental group received metacognitive support by prompts, whereas the control group learned without prompts. All participants were completely free to navigate in the learning environment and to use their learning strategies. During learning, notes could

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

be taken on a blank sheet of paper (e.g., for summarizing or structuring information), but the participants were not allowed to use their notes to work on the knowledge tests. The participants were instructed to read and think aloud during the whole learning phase as practiced before, and these activities were videotaped. If a participant stopped thinking aloud for more than five seconds, the experimenter reminded her or him by saying, “Please think aloud.”

Directly after learning, the students worked on the three knowledge tests described above. Overall, the duration of the session was approximately two hours.

3.4 Coding Scheme

A coding scheme based on our theoretical framework of self-regulated hypermedia learning (Bannert, 2007) was used for segmenting and coding the students’ verbal protocols. Our theoretical framework characterizes hypermedia learning into the major categories *Metacognition*, *Cognition*, and *Motivation*. In addition, it distinguishes several sub-categories within the categories Metacognition and Cognition, as further described below.

Table 1 presents the coding categories and provides descriptions and examples. The coding scheme comprises the main categories Metacognition, Cognition, Motivation, and Other. Metacognition includes the sub-categories *Orientation*, *Goal specification*, *Planning*, *Searching for information*, *Judgment of its relevance*, *Evaluating goal attainment*, and finally *Monitoring and regulation*. Cognition contains *Reading*, *Repeating information*, and deeper processing, that is, *Elaboration* and *Organization of information*. The main category of Motivation includes all positive and negative utterances on the task, the situation, or oneself. Finally, all task-irrelevant utterances, non-classifiable utterances, and the handling of the prompts for the experimental group were assigned to the category Other.

The coding was conducted based on the procedure presented by Chi (1997). Segmentation of the verbal protocols was based on meaning. A segment was assigned for every definable learning activity. Multiple or nested codes were not allowed. Four trained research assistants coded the verbal protocols of all 70 participants. A random sample of three participants from each of the experimental group and the control group was selected to compute the interrater’s reliability. The reliability, based on 1,385 segments, showed substantial agreement: Cohen’s Kappa = .78, which is seen as sufficient for the following analysis.

3.5 Analysis

An example of the coded data used for the process analysis is presented in Table 2. The data comprise three types of information: 1) a Case ID that clearly distinguishes the participants, 2) a time stamp that indicates the beginning of an event, and 3) a learning activity — that is, the assigned category of the coding scheme (CODE). Using this information, it is possible to compute not only the frequency of events but also to determine the relative arrangement of multiple events. For example, in the short section of

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

Table 2, MONITOR is the most frequent activity (3 occurrences). Furthermore, MONITOR is directly followed by READ twice and directly followed by ORGANIZATION once.

Table 1. Coding scheme for analyzing students’ learning activities

Code	Coding Category	Description and Examples
Metacognition		
ORIENT	Orientation	Task clarification, overview of material. <i>I will sketch the menu first.</i>
SETGOAL	Goal specification	Goal setting and sub-goaling <i>I have to learn the basic concepts of operant conditioning.</i>
PLAN	Planning	Planning how to proceed <i>First I will decide in which sequence I have to learn and which pages to read.</i>
SEARCH	Search	Searching for information <i>Where is the page with the information about plans of reinforcement?</i>
EVALUATE	Judgment	Judgments about the relevance of information <i>Skinner’s Vita is not relevant for my learning task.</i>
EVAL	Evaluation	Checking and evaluating <i>Did I process all the topics?</i>
MONITOR	Monitoring	Monitoring one’s own learning <i>Ah, now I understand the principle.</i>
Cognition		
READ	Reading	Reading out loud
REPEAT	Repeating	Repeating
ELABORATE	Elaboration	Deeper processing, <i>paraphrasing, connecting, inferring</i>
ORGANIZATION	Organization	Organization <i>drawing a map, writing down major concepts</i>
Motivation		
MOT	Motivation	Positive, negative, neutral motivational utterances regarding a task, person, or situation <i>The task is very interesting.</i>
Other		
REST	Other	Off-topic statements, comments on technique, not interpretable statements, pauses <i>May I make notes? The mouse doesn’t work well.</i>

Table 2. Section of an event log

<i>Case ID</i>	<i>Timestamp</i>	<i>Learning Activity (CODE)</i>
		...
Case 1	04:14	ORIENT
Case 1	04:19	MONITOR
Case 1	04:24	READ
Case 1	04:37	REPEAT
Case 1	04:43	ORIENT
Case 1	05:01	MONITOR
Case 1	05:03	READ
Case 1	06:51	MONITOR
Case 1	06:56	ORGANIZATION
Case 1	07:30	REPEAT
		...

In the first step of our analysis, we took the frequencies of coded learning activities into account (frequency analysis) and used the frequency of metacognitive events to examine the expected mediation effect on transfer performance (mediation analysis). Subsequently, the sequential order of the coded learning activities is analyzed using a PM algorithm to discover differences in the process models of the experimental and control groups (process mining). Therefore, our view on temporality corresponds to the relative arrangement of multiple events.

4 RESULTS

A preliminary analysis showed that the randomized assignment of participants resulted in two subsamples with similar learner characteristics. With the exception of one subscale of reading competency — namely, *text comprehension* (measured by ELVES; Richter & van Holt, 2005) — no significant differences were found. In the case of the subscale text comprehension, students in the control group scored significantly better than those in the experimental group ($t(69) = 2.97, p = .004, d = 0.72$; two-tailed testing). In conclusion, this analysis indicates that the following results are not caused by unbalanced subsamples.

4.1 Frequency Analysis

Table 3 presents the descriptive and test statistics of all coded events for the experimental group and the control group. In addition to the minimum and maximum occurrence of each category, absolute frequencies, means, and standard deviations are listed. A total of 8,743 events were coded for the students in the experimental group, and a total of 8,087 events were coded for the students in the control group. For the experimental group, there were, on average, approximately 250 events coded in

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

40 minutes of learning time, with 116 metacognitive, 107 cognitive, 2 motivational, and 25 other utterances. Participants in the control group showed a mean of approximately 231 events, with 98 metacognitive, 106 cognitive, 3 motivational, and 24 other utterances.

A one-tailed t-test for independent samples showed that the experimental and control groups significantly differ in the number of metacognitive utterances ($t(69) = 1.80, p = .038, d = 0.44$). As expected, students in the experimental group who had been supported through metacognitive prompts showed a higher number of metacognitive learning activities ($M = 116.43, SD = 45.97$) than students in the control group without prompts ($M = 98.49, SD = 36.72$). Moreover, both groups showed a similar number of utterances in the remaining main categories Cognition ($M_{EG} = 107.43, SD_{EG} = 36.01; M_{CG} = 106.31, SD_{CG} = 44.65$), Motivation ($M_{EG} = 2.06, SD_{EG} = 4.14; M_{CG} = 2.54, SD_{CG} = 4.40$), and Other ($M_{EG} = 25.34, SD_{EG} = 15.08; M_{CG} = 23.71, SD_{CG} = 12.19$). For these three categories, the t-tests for independent samples were not significant.

Concerning the descriptive statistics of the subcategories of Metacognition, the experimental group showed more Monitoring ($M_{EG} = 71.17, SD_{EG} = 37.62; M_{CG} = 58.00, SD_{CG} = 27.46$), Orientation ($M_{EG} = 14.31, SD_{EG} = 7.23; M_{CG} = 11.14, SD_{CG} = 5.80$), Evaluation ($M_{EG} = 3.63, SD_{EG} = 3.08; M_{CG} = 2.49, SD_{CG} = 3.03$), and Planning ($M_{EG} = 1.74, SD_{EG} = 1.65; M_{CG} = 0.74, SD_{CG} = 1.09$) compared to the control group. In both groups, the highest frequency occurred for Monitoring, followed by Orientation, Searching, and Judgment, whereas Planning and Goal specification were rarely executed by students. As reported in Table 3, on the right side, differences between the experimental and control groups are significant for Orientation, Planning, and Monitoring.

Within the main category Cognition, participants of the control group showed more reading activities ($M_{EG} = 40.66, SD_{EG} = 15.75; M_{CG} = 44.49, SD_{CG} = 18.19$) but less Elaboration ($M_{EG} = 21.91, SD_{EG} = 12.92; M_{CG} = 18.40, SD_{CG} = 17.54$) and less Organization ($M_{EG} = 26.37, SD_{EG} = 13.87; M_{CG} = 24.60, SD_{CG} = 12.51$) than participants of the experimental group. However, all differences regarding these categories are non-significant. Finally, motivational events seldom occurred in both groups and with non-significant differences.

4.2 Mediation Analysis

A mediation analysis was conducted to investigate whether the observed relationship between the treatment group and learning performance (outcome variable) is mediated by the number of metacognitive events during learning. Regarding measurements of learning outcome, only transfer performance (i.e., a post-test score) differed significantly between the experimental and control groups ($M_{EG} = 20.61, SD_{EG} = 3.97; M_{CG} = 18.79, SD_{CG} = 4.30; t(69) = 1.85, p = .035, d = 0.45$; for more details on learning outcomes, see Bannert et al., 2015). Furthermore, both the number of metacognitive events and its sub-category Monitoring significantly correlate with transfer performance (Metacognitive events: $r = .22, p = .033$; Monitoring: $r = .32, p = .003$). Therefore, these two variables are regarded as possible mediators, and only transfer performance is included as an outcome variable.

Table 3: Absolute frequencies, means, and test statistics of all coded learning events for the experimental group and the control group

	Experimental Group (n=35)					Control Group (n=35)					<i>t</i>	<i>p</i>	<i>d</i>
	Min	Max	Absolute Frequency	<i>M</i>	<i>SD</i>	Min	Max	Absolute Frequency	<i>M</i>	<i>SD</i>			
Metacognition	34	242	4075	116.43	45.97	34	173	3447	98.49	36.72	1.804	.038	0.44
Orientation	5	30	501	14.31	7.23	2	28	390	11.14	5.80	2.025	.024	0.49
Planning	0	5	61	1.74	1.65	0	4	26	0.74	1.09	2.987	.002	0.73
Goal specification	0	10	72	2.06	2.36	0	8	67	1.91	1.84	0.282	.309	0.07
Search	1	32	414	11.83	7.54	1	57	453	12.94	10.64	-0.506	.308	-0.12
Judgment	2	23	409	11.69	5.70	0	33	394	11.26	7.35	0.273	.393	0.07
Evaluation	0	15	127	3.63	3.08	0	15	87	2.49	3.03	1.565	.061	0.38
Monitoring	11	203	2491	71.17	37.62	9	124	2030	58.00	27.46	1.673	.049	0.41
Cognition	47	201	3760	107.43	36.01	30	193	3721	106.31	44.65	0.115	.909	0.03
Reading	20	84	1423	40.66	15.75	19	89	1557	44.49	18.19	-0.941	.350	-0.03
Repeating	2	45	647	18.49	11.04	2	59	659	18.83	12.25	-0.123	.902	-0.03
Elaboration	3	55	767	21.91	12.92	0	56	644	18.40	17.54	0.955	.343	0.23
Organization	3	61	923	26.37	13.87	0	58	861	24.60	12.51	0.561	.576	0.14
Motivation	0	18	72	2.06	4.14	0	22	89	2.54	4.40	-0.476	.636	0.11
Other	8	69	887	25.34	15.08	6	58	830	23.71	12.19	0.497	.621	0.12
Sum of all coded events	126	473	8743	249.80	76.64	112	390	8087	231.06	76.27	1.023	.310	0.25

Note: Since we expected metacognitive prompting to increase the number of metacognitive utterances, we conducted one-tailed testing for metacognitive categories; elsewhere we conducted two-tailed testing; $p < .05$.

We used the PROCESS custom dialog box for SPSS based on the regression-based approach of Hayes (2013) to run the mediation analysis, which calculates bootstrapped confidence intervals (BCa CI) for the indirect effect and the measurement of effect size. There was a significant indirect effect of the treatment on transfer performance through the number of metacognitive events, $b = 0.33$, BCa CI [-0.01, 1.09]. Kappa-squared (Preacher & Kelley, 2011) was used to measure the effect size. The detected effect is relatively small, $K^2 = .039$, 95% BCa CI [.004, .127]. Furthermore, there was a significant indirect effect through the number of Monitoring events, $b = 0.48$, BCa CI [0.01, 1.24]. Again, this represents a small effect, $K^2 = .058$, 95% BCa CI [.006, .142].

In summary, the number of metacognitive events and of its sub-category Monitoring could be identified as mediator variables. Metacognitive prompting increased the occurrence of metacognitive events, especially of Monitoring, which in turn enhanced the transfer performance. Due to the mediation effect of the sub-category Monitoring being even slightly larger than the effect of all metacognitive events, we conclude that the mediation is mainly driven by Monitoring. Figure 3 presents the mediation model, including Monitoring as mediator variable.

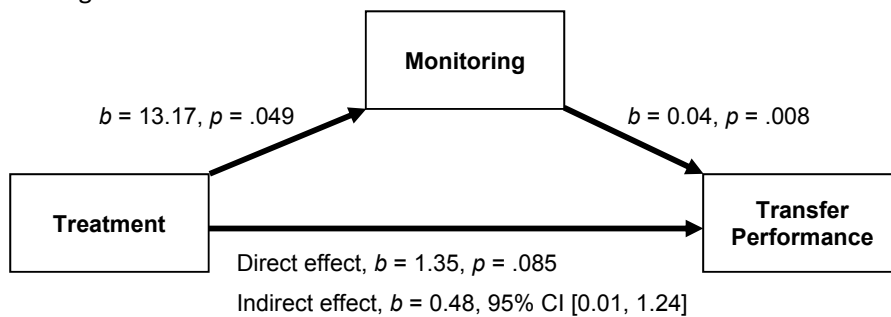


Figure 3: Mediation through the number of monitoring events

4.3 Process Analysis Using the HeuristicsMiner Algorithm

To apply the HeuristicsMiner, we had to simplify the categories of the coding scheme (see Table 1) for three reasons. First, the event classes Planning, Goal specification, and Motivation showed a very low frequency in our event log. Second, the HeuristicsMiner algorithm should preferably be used on data without too many different event classes (Rozinat, 2010). Finally, theoretical SRL models describe regulation processes mainly with the three phases of forethought, performance, and reflection (e.g., Zimmerman, 2000). The simplification was conducted as follows: We aggregated the metacognitive events Orientation, Planning, and Goal specification into a new event class called *Analyze*. Furthermore, the event class Judgment was added to Monitoring. In addition, the cognitive events Elaboration and Organization were combined to form a new event class called *Process*, that is, deeper processing. Finally, the event classes Motivation and Other were excluded from the process analysis. Altogether, seven event classes, listed in Table 4, were used for the analysis with PM techniques. With respect to the mean number of events, there was only a significant difference between both groups for the category Analyze ($t(69) = 2.36, p = .011, d = 0.57$).

Table 4: Absolute frequencies, means, and test statistics of aggregated categories for the experimental group and the control group

	Experimental Group (n=35)					Control Group (n=35)					<i>t</i>	<i>p</i>	<i>d</i>
	Min	Max	Absolute Frequency	<i>M</i>	<i>SD</i>	Min	Max	Absolute Frequency	<i>M</i>	<i>SD</i>			
Analyze	5	37	634	18.11	8.58	3	33	483	13.80	6.61	2.356	.011	0.57
Search	1	32	414	11.83	7.54	1	57	453	12.94	10.64	-0.506	.310	-0.12
Evaluation	0	15	127	3.63	3.08	0	15	87	2.49	3.03	1.565	.061	0.38
Monitoring	20	211	2900	82.86	39.37	12	131	2424	69.26	30.71	1.612	.056	0.39
Reading	20	84	1423	40.66	15.75	19	89	1557	44.49	18.19	-0.941	.350	-0.03
Repeating	2	45	647	18.49	11.04	2	59	659	18.83	12.25	-0.123	.902	-0.03
Process	9	93	1690	48.29	21.54	3	84	1505	43.00	20.73	1.046	.299	0.25

Note: Since we expected metacognitive prompting to increase the number of metacognitive utterances, we conducted one-tailed testing for metacognitive categories; elsewhere we conducted two-tailed testing; $p < .05$.

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

The relative arrangement of learning activities was analyzed by applying the HeuristicsMiner algorithm in combination with the DWS mining plugin. The trace clustering did not split the cases into clusters for the participants of the experimental group (n=35) or the control group (n=35). This means a single process model can already express the event log with sufficient precision for both groups.

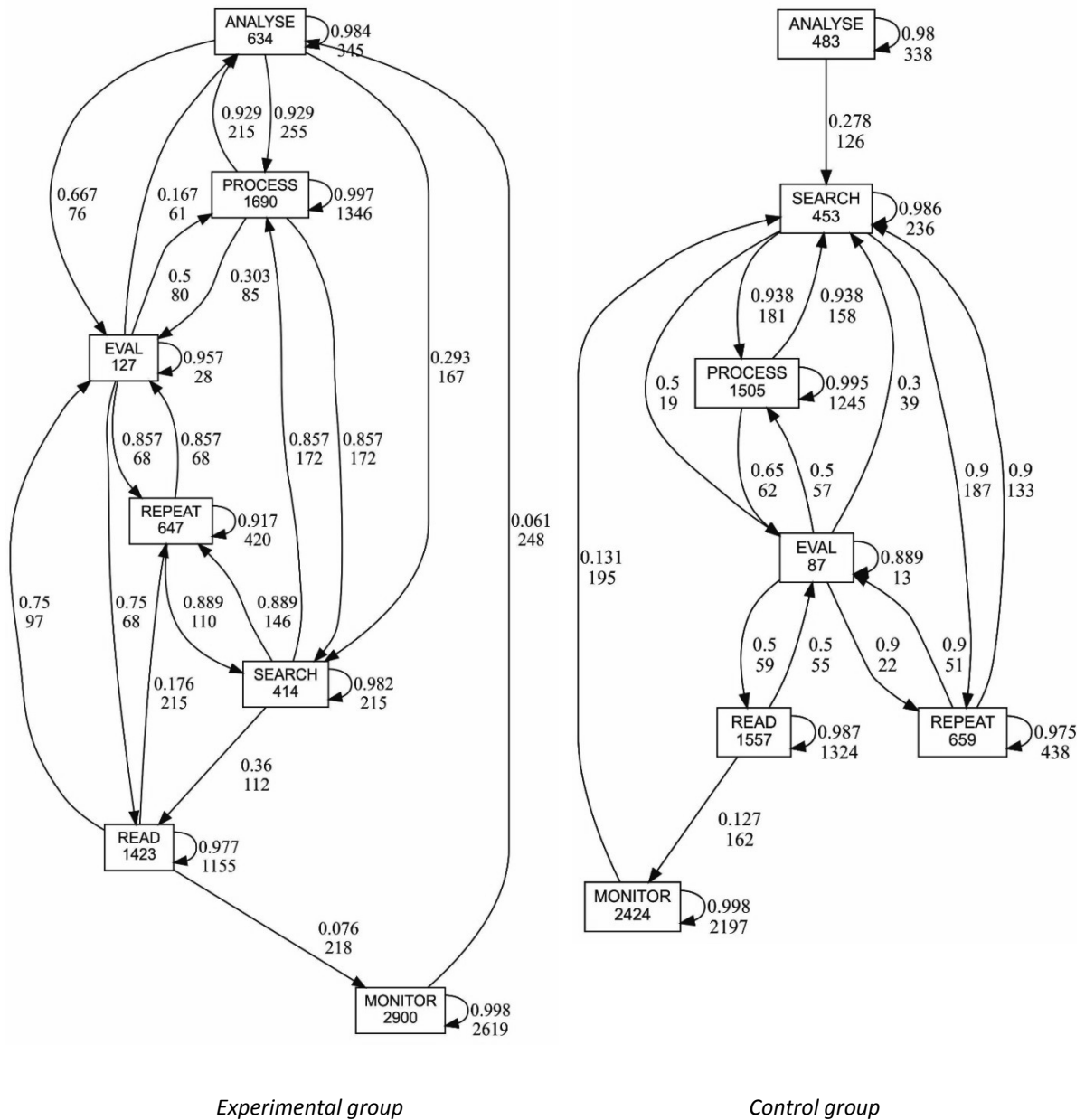


Figure 4. Process models for the experimental group (n=35) and for the control group (n=35) represented as a heuristic net. Metacognitive Activities: ANALYSE = Orientation, Planning, and Goal specification; EVAL = Evaluation; MONITOR = Monitoring and Judgment. Cognitive Activities: READ = Reading; REPEAT = Repeating; PROCESS = Elaborate and Organization

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

The resulting process models are displayed in Figure 4, where they are represented as heuristic nets for the experimental and control groups. These visual representations of the process models comprise square boxes that represent the event classes and arcs between these boxes that indicate the dependency between two event classes. The number in the event box represents the occurrence of an event class in the log. The arcs are labelled with two types of information. The upper number displays the dependency measure, which indicates the certainty of a dependency relation between two activities. A value close to 1.0 indicates a high certainty that a dependency relation exists. The lower number shows the number of times this transition is used, that is, how often event *a* is followed by event *b*. An arc pointing back at the same box indicates a self-loop, meaning that an event class often occurred multiple times in a row in loops of length one or length two (e.g., ACCB, ACDCDB).

The fitness between the mined model and the event log used for generating this model was measured using the so-called *Improved Continuous Semantic Fitness* (De Medeiros, 2006; range: $-\infty$ to 1.0). This fitness measure indicates the number of correct parsed event sequences, whereas a punishment for allowed extra behaviour in the model is subtracted from this number. The idea of this measure is to favour a process model that allows for less extra behaviour if several models can correctly parse the same number of event sequences. Both process models show a substantial fitness value: the experimental group model = 0.53, and the control group model = 0.62.

4.3.1 Process model of the experimental group

For the experimental group, a common pattern — that is, a path of transitions with high certainty — is ANALYZE → PROCESS → SEARCH → REPEAT → EVAL → READ → MONITOR → ANALYZE. Moreover, the process model comprises a number of loops with high certainty between two activities. Participants circle between ANALYZE and PROCESS, EVAL and REPEAT, SEARCH and PROCESS, EVAL and READING, and SEARCH and REPEAT. Apparently, these loops always occur between metacognitive and cognitive learning activities but never between two cognitive or two metacognitive events. Furthermore, it is interesting that EVAL is connected with several other learning events, meaning it takes an important position in the structure of the process, although this event class has a relatively low frequency. MONITOR only shows a weak connection in the process model. This event class follows READ and is followed by ANALYZE. Finally, the model shows self-loops for all event classes, indicating that an activity can be performed multiple times in a row.

4.3.2 Process model of the control group

The model of the control group shows the most common path of transitions for SEARCH → PROCESS → EVAL → REPEAT → READ → MONITOR → SEARCH. In contrast to the model of the students in the experimental group, ANALYZE is only weakly connected with SEARCH, and therefore, it is quite isolated. Similar to the experimental group, the low-frequency event class EVAL is also connected with several other learning activities. MONITOR is only weakly connected, whereas this event class follows READ, as in the model of the experimental group, but is followed by SEARCH instead of ANALYZE. In comparison with the experimental group, this process model shows fewer loops with high certainty between two activities (only between SEARCH and PROCESS, SEARCH and REPEAT, and EVAL and REPEAT), but again

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

these loops only occur between metacognitive and cognitive events. Again, all event classes show self-loops.

Overall, the process models of the experimental and control groups especially differ in two points. First, ANALYZE (including the activities Orientation, Planning, and Goal specification) is hardly connected in the model of the control group, but this event class is well embedded in the process of the experimental group. Second, students in the experimental group show more loops between metacognitive and cognitive events, which can be interpreted as “regulation circles.” For example, they circle with high certainty between ANALYZE and PROCESS and between EVAL and READ. Despite these differences, both models have in common that EVAL takes an important position in the described process. The frequency analysis could not reveal the importance of this event class, even showing that EVAL is one of the least-frequent categories. Here, the analysis of the sequential order provides additional information. Moreover, in both models, MONITOR, the metacognitive category with the highest frequency, is hardly connected with other learning activities. Based on recent SRL models, we argue that MONITOR does not have a clear position but can follow each learning activity (e.g., A → MONITORING → B → MONITORING → C → MONITORING). The HeuristicMiner algorithm could have failed to position this activity in the process model because its modelling notation does not allow for so-called *duplicate tasks* (i.e., an activity that has more than one label in the process model).

5 DISCUSSION AND IMPLICATIONS FOR FUTURE RESEARCH

In this study, we analyzed think-aloud data from an experimental study to investigate the effects of metacognitive prompts during learning on SRL processes. In addition to an analysis of frequencies of learning events, we focused on exploring the sequential structure of regulation activities using PM techniques.

As expected, the analysis of coded think-aloud data provides deeper insights into the effects of metacognitive prompts on students’ regulatory processes during hypermedia learning. The findings of a frequency analysis indicate differences in the number of metacognitive utterances between students in the experimental group, who were prompted by metacognitive prompts, and those in the control group, who learned without prompts. Participants supported by metacognitive prompts articulated significantly more metacognitive activities and achieved better transfer performance. In addition, a mediation analysis revealed that prompting increased the number of metacognitive activities, especially Monitoring, which, in turn, increased the transfer performance. Both results are in line with findings of research on metacognitive prompting (e.g., Azevedo et al., 2004; Bannert, 2009) and the assumed effect mechanism of this type of metacognitive support (e.g., Bannert & Mengelkamp, 2013).

A microanalysis of the relative arrangement of learning activities was conducted by means of PM techniques to discover specific sequential patterns in the learning process of the experimental group versus the control group. This process analysis provided additional information on the effects of metacognitive prompts that could not be revealed by a simple analysis of frequencies of occurring

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

learning events. A comparison of the process models of students in the experimental and the control group showed two striking differences. First, activities of orientation, planning, and goal specification (aggregated as ANALYZE) are much better integrated in the process model of the experimental group, whereas this event class was quite isolated in the process model of the control group. Second, more loops between cognitive and metacognitive learning activities were identified in the process model of the experimental group, indicating that more regulation steps occurred. In conclusion, these differences indicate that the use of metacognitive prompts resulted in a better integration of ANALYZE events and a higher number of regulation loops. SRL models (e.g., Winne & Hadwin, 2008; Zimmerman, 2008) emphasize both the importance of orientation phases and an active regulation for successful learning. Following this, the fostered process patterns in the model of the experimental group are in line with current theoretical assumptions. We conclude that these process patterns could be successfully scaffolded through the application of metacognitive prompts. However, the evaluation of learning progress (EVAL) is similarly integrated in both process models, meaning no different effect on this event category could be detected. The findings of the process patterns could be used to optimize our metacognitive prompts further. For example, the design of prompts could be optimized by aiming at scaffolding the sequential deployment of evaluating activities in more detail. SRL models suggest that evaluation activities are followed by an update of the orientation phase. This transition was not represented in the process model of the experimental group. Here, an optimization process could be used.

With respect to the metacognitive support used in this study — that is, an introduction about what metacognitive prompts are, why they are important, and how to use them in combination with metacognitive prompts during learning — it is necessary to discuss which components of support have contributed to the findings. Based on our experience with metacognitive prompts and research on metacognitive prompting, at least a brief training or an introduction to the concept of metacognitive prompts is necessary in advance to guarantee an adequate application of prompts during learning (e.g., Bannert & Mengelkamp, 2013; Bannert & Reimann, 2012). Therefore, it is challenging to determine the individual effect of both the introduction and prompting components. To our knowledge, there is no empirical study that systematically compares the impact of training of prompt use, metacognitive prompting, and their combination. Consequently, this research question should be addressed in future work.

As an inductive approach, the validity of PM depends on the representativity and quality of the data stored in the event log (Reimann, Frerejean, & Thompson, 2009). It is possible that SRL processes, for example, those obtained by think-aloud protocols or log files, comprise a high variety of regulatory behaviour (Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007; Winne, 2014). Therefore, we applied a technique for trace clustering in combination with a PM algorithm to check whether a single process model for the whole event log is appropriate. The applied trace clustering did not split the cases into subsets of participants. However, new approaches of trace clustering are currently rising in the PM domain (e.g., de Weerd et al., 2013). These approaches could possibly improve the detection of

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

homogenous subsamples, which in turn could enhance the quality of the mined process models. There are approaches to clustering students according to their interactions and activities in computer-based learning environments based on a set of variables (e.g., Biswas, Jeong, Kinnebrew, Sulcer, & Roscoe, 2010; Bouchet, Harley, Trevors, & Azevedo, 2013). However, these approaches do not explicitly include an event-centred perspective and a timing aspect, but are based on frequencies of interactions with the learning environment. Furthermore, a subset of participants can also be selected for process analysis based on learner characteristics (e.g., high vs. low prior knowledge) or learning outcomes (e.g., high vs. low achieving students; see Bannert et al., 2014 for an example).

Regarding further limitations of our analysis, it has to be noted that the resulting process models are dependent on the learning setting (learning environment, learning material, and instructions on the learning task). In addition, they are descriptive models. Moreover, findings depend on the underlying coding scheme and its level of granularity. In general, more research on PM techniques in the field of SRL and metacognition is needed; for example, for deriving guidelines for parameter settings aiming to improve the quality of a mined process model. Therefore, we encourage other researchers to use PM techniques to analyze their process data with respect to the sequential and temporal characteristics of learning events. In addition, a comparison of different methods for sequential and temporal analyses on the same data would be beneficial for discovering the advantages and disadvantages of recent process analysis methods. For example, a comparison could be made of different approaches presented in a special issue on the sequential and temporal characteristics of self- and socially regulated learning (Molenaar & Järvelä, 2014).

The resulting process models of our analysis represent a description of the underlying learning processes in our sample of students. In future studies, the validity of the discovered process patterns should be investigated by checking the conformance of these models to new data sets. For this purpose, the mined process models of the HeuristicsMiner algorithm can be converted into petri nets, and then methods for conformance checking can be applied within the ProM framework (Rozinat & van der Aalst, 2008). In this way, the conformance — that is, the differences between a discovered process model and a new event log — can be determined. Another possible scenario for the application of conformance checking would be the derivation of a system of event sequences on the micro-level based on the theoretical assumptions of SRL models. An illustration of this approach is presented in Bannert et al. (2014). However, more micro-level theories would be needed for this approach. At the moment, only the COPES model (Winne & Hadwin, 2008) provides a detailed level of granularity regarding information processing, but even this model is far from the level of elaboration needed to correspond directly to the granularity of our event data.

Finally, an advantage of PM techniques is the representation of sequential characteristics as visual process models. Primarily, this helps the researcher to grasp easily the course of learning activities and regulatory patterns. However, this type of visual representation can also be used to give process feedback to the students and, thereby, be a resource for learners as well (Reimann et al., 2009). Winne

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

(2014) recommends supporting students with information on their past learning processes through displays of traces they can interpret. Following this direction in future research on SRL, process models generated by PM techniques could make a substantial contribution in providing feedback to learners.

ACKNOWLEDGMENTS

This research was supported by funds from the German Research Foundation (DFG: BA 2044/7-1). We thank Ferdinand Guggenberger, Valentina Haspert, Michaela Hofbauer and Lisa Müsse, the student research assistants of this project, for their valuable help for coding the think-aloud data.

REFERENCES

- Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, 40(4), 199–209.
doi:10.1207/s15326985ep4004_2
- Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning*, 4(1), 87–95. doi:10.1007/s11409-009-9035-7
- Azevedo, R., Cromley, J. G., Moos, D. C., Greene, J. A., & Winters, F. I. (2011). Adaptive content and process scaffolding: A key to facilitating students' self-regulated learning with hypermedia. *Psychological Testing and Assessment Modeling*, 53(1), 106–140.
- Azevedo, R., Guthrie, J., & 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(1–2), 87–111. doi:10.2190/DVWX-GM1T-6THQ-5WC7
- Azevedo, R., Trevors, G., Duffy, M., Feyzi-Behnagh, M., Bouchet, F., & Landis, R. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 427–449). New York, NY: Springer. doi:10.1007/978-1-4419-5546-3_28
- Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research*, 35(4), 359–375. doi:10.2190/94V6-R58H-3367-G388
- Bannert, M. (2007). *Metakognition beim Lernen mit Hypermedia. Erfassung, Beschreibung und Vermittlung wirksamer metakognitiver Lernstrategien und Regulationsaktivitäten [Metacognition and hypermedia learning]*. Münster: Waxmann.
- Bannert, M. (2009). Promoting self-regulated learning through prompts: A discussion. *Zeitschrift für Pädagogische Psychologie / German Journal of Educational Psychology*, 23(2), 139–145. doi:10.1024/1010-0652.23.2.139
- Bannert, M., & Mengelkamp, C. (2013). Scaffolding hypermedia learning through metacognitive prompts. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 171–186). New York, NY: Springer. doi:10.1007/978-1-4419-5546-3_12

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

- Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science*, 40(1), 193–211. doi:10.1007/s11251-011-9167-4
- Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning*, 9(2), 161–185. doi:10.1007/s11409-013-9107-6
- Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior*, 52, 293–306. doi:10.1016/j.chb.2015.05.038
- Biswas, G., Jeong, H., Kinnebrew, J. S., Sulcer, B., & Roscoe, R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology Enhanced Learning*, 5(2), 123–152. doi:10.1142/S1793206810000839
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York, NY: D. McKay.
- Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction*, 7(2), 161–186. doi:10.1016/S0959-4752(96)00015-1
- Bouchet, F., Harley, J. M., Trevors, G. J., & Azevedo, R. (2013). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *Journal of Educational Data Mining*, 5(1), 104–146.
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, 6(3), 271–315. doi:10.1207/s15327809jls0603_1
- De Medeiros, A. K. A. (2006). *Genetic Process Mining*. Eindhoven University of Technology, Eindhoven.
- De Medeiros, A. K. A., Guzzo, A., Greco, G., van der Aalst, W. M. P., Weijters, A. J. M. M., van Dongen, B. F., & Saccà, D. (2008). Process mining based on clustering: A quest for precision. *Lecture Notes in Computer Science*, 4928, 17–29. doi:10.1007/978-3-540-78238-4_4
- De Weerd, J., de Backer, M., Vanthienen, J., & Baesens, B. (2012). A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real-life event logs. *Information Systems*, 37(7), 654–676. doi:10.1016/j.is.2012.02.004
- De Weerd, J., vanden Broucke, S., Vanthienen, J., & Baesens, B. (2013). Active trace clustering for improved process discovery. *IEEE Transactions on Knowledge and Data Engineering*, 25(10), 2708–2720. doi:10.1109/TKDE.2013.64
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Fluxicon Disco (Version 1.7.2) [Computer Software]. (2014). Retrieved from <http://fluxicon.com/disco/>
- Ge, X. (2013). Designing learning technologies to support self-regulation during ill-structured problem-solving processes. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 213–228). New York, NY: Springer. doi:10.1007/978-1-4419-5546-3_15

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

- Greco, G., Guzzo, A., Pontieri, L., & Saccà, D. (2006). Discovering expressive process models by clustering log traces. *IEEE Transactions on Knowledge and Data Engineering*, 18(8), 1010–1027. doi:10.1109/TKDE.2006.123
- Greene, J. A., Dellinger, K. R., Tüysüzoglu, B. B., & Costa, L.-J. (2013). A two-tiered approach to analyzing self-regulated learning data to inform the design of hypermedia learning environments. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 117–128). New York, NY: Springer. doi:10.1007/978-1-4419-5546-3_8
- Günther, C. W., & van der Aalst, W. M. P. (2007). Fuzzy mining: Adaptive process simplification based on multi-perspective metrics. *Business Process Management: Lecture Notes in Computer Science*, 4714, 328–343. doi:10.1007/978-3-540-75183-0
- Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2(2), 107–124. doi:10.1007/s11409-007-9016-7
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.
- Hoffman, B., & Spatariu, A. (2011). Metacognitive prompts and mental multiplication: Analyzing strategies with a qualitative lens. *Journal of Interactive Learning Research*, 22(4), 607–635.
- Jeong, H., Gupta, A., Roscoe, R., Wagster, J., Biswas, G., & Schwartz, D. (2008). Using hidden Markov models to characterize student behaviors in learning-by-teaching environments. *Lecture Notes in Computer Science*, 5091, 614–625. doi:10.1007/978-3-540-69132-7_64
- Johnson, A. M., Azevedo, R., & D’Mello, S. K. (2011). The temporal and dynamic nature of self-regulatory processes during independent and externally assisted hypermedia learning. *Cognition and Instruction*, 29(4), 471–504. doi:10.1080/07370008.2011.610244
- Kapur, M. (2011). Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of Computer-Supported Collaborative Learning*, 6(1), 39–56. doi:10.1007/s11412-011-9109-9
- Li, C., & Biswas, G. (2002). A Bayesian approach for structural learning with hidden Markov models. *Scientific Programming*, 10(3), 201–219.
- Lin, X. (2001). Designing metacognitive activities. *Educational Technology Research and Development*, 49(2), 23–40. doi:10.1007/BF02504926
- Lin, X., Hmelo, C., Kinzer, C. K., & Secules, T. J. (1999). Designing technology to support reflection. *Educational Technology Research and Development*, 47(3), 43–62. doi:10.1007/BF02299633
- Lin, X., & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching*, 36(7), 837–858. doi:10.1002/(SICI)1098-2736(199909)36:7<837::AID-TEA6>3.3.CO;2-L
- Martin, T., & Sherin, B. (2013). Learning analytics and computational techniques for detecting and evaluating patterns in learning: An introduction to the special issue. *Journal of the Learning Sciences*, 22(4), 511–520. doi:10.1080/10508406.2013.840466

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

- Molenaar, I., & Chiu, M. M. (2014). Dissecting sequences of regulation and cognition: Statistical discourse analysis of primary school children's collaborative learning. *Metacognition and Learning*, 9(2), 137–160. doi:10.1007/s11409-013-9105-8
- Molenaar, I., & Järvelä, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning*, 9(2), 75–85. doi:10.1007/s11409-014-9114-2
- Moos, D. C., & Azevedo, R. (2009). Self-efficacy and prior domain knowledge: To what extent does monitoring mediate their relationship with hypermedia learning? *Metacognition and Learning*, 4(3), 197–216. doi:10.1007/s11409-009-9045-5
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, 16(2), 93–115. doi:10.1037/a0022658
- ProM (Version 5.2) [Computer Software]. (2008). Retrieved from <http://www.processmining.org>
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 239–257. doi:10.1007/s11412-009-9070-z
- Reimann, P., Frerejean, J., & Thompson, K. (2009). Using process mining to identify models of group decision making processes in chat data. In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Computer-Supported Collaborative Learning Practices: CSCL2009 Conference Proceedings* (pp. 98–107). Rhodes, Greece: International Society for the Learning Sciences.
- Reimann, P., & Yacef, K. (2013). Using process mining for understanding learning. In R. Luckin, P. Puntambekar, P. Goodyear, B. Grabowski, J. Underwood, & N. Winters (Eds.), *Handbook of design in educational technology* (pp. 472–481). New York, NY: Routledge.
- Reisig, W. (1985). *Petri Nets: An introduction*. Berlin: Springer.
- Richter, T., & van Holt, N. (2005). ELVES: Ein computergestütztes Diagnostikum zur Erfassung der Effizienz von Teilprozessen des Leseverstehens [ELVES: Computer-based assessment of the efficiency of processes involved in reading comprehension]. *Diagnostica*, 51(4), 169–182. doi:10.1026/0012-1924.51.4.169
- Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. (2010). *Handbook of educational data mining*. Boca Raton, FL: Chapman & Hall/CRC.
- Rozinat, A. (2010, October 18). ProM tips: Which mining algorithm should you use? [Web log post]. Retrieved from <http://fluxicon.com/blog/2010/10/prom-tips-mining-algorithm/>
- Rozinat, A., & van der Aalst, W. M. P. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems*, 33(1), 64–95. doi:10.1016/j.is.2007.07.001
- Schnotz, W. (1998). Strategy-specific information access in knowledge acquisition from hypertext. In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools, and reasoning: Essays on situated cognition* (pp. 336–358). Berlin: Springer.
- Schoor, C., & Bannert, M. (2012). Exploring regulatory processes during a computer-supported collaborative learning task using process mining. *Computers in Human Behavior*, 28(4), 1321–1331. doi:10.1016/j.chb.2012.02.016

(2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100.

- Trčka, N., Pechenizkiy, M., & van der Aalst, W. M. P. (2010). Process mining from educational data. In C. Romero, S. Ventura, M. Pechenizkiy, & R. Baker (Eds.), *Handbook of educational data mining* (pp. 123–142). Boca Raton, FL: Chapman & Hall/CRC.
- Van der Aalst, W. M. P. (2011). *Process mining: Discovery, conformance and enhancement of business processes*. Berlin/Heidelberg: Springer. doi:10.1007/978-3-642-19345-3
- Van der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, 16(9), 1128–1142. doi:10.1109/TKDE.2004.47
- Veenman, M. (1993). *Metacognitive ability and metacognitive skill: Determinants of discovery learning in computerized learning environments*. Amsterdam: University of Amsterdam.
- 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., 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
- Weijters, A. J. M. M., van der Aalst, W. M. P., & de Medeiros, A. K. A. (2006). Process mining with the HeuristicsMiner algorithm. *BETA Working Paper Series 166*. Eindhoven, Netherlands: Eindhoven University of Technology.
- Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning*, 9(2), 229–237. doi:10.1007/s11409-014-9113-3
- Winne, P. H., & Hadwin, A. F. (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, NY: Taylor & Francis.
- Winne, P. H., & Hadwin, A. F. (2013). nStudy: Tracing and supporting self-regulated learning in the internet. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 293–308). New York, NY: Springer. doi:10.1007/978-1-4419-5546-3_20
- Winne, P. H., & Nesbit, J. C. (2009). Supporting self-regulated learning with cognitive tools. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 259–277). New York, NY: Routledge.
- Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). San Diego, CA: Academic Press.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego, CA: Academic Press.
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183. doi:10.3102/000283120731290

Analysis 2:

Evaluating the Impact of Instructional Support Using Data Mining and Process Mining: A Micro-Level Analysis of the Effectiveness of Metacognitive Prompts

Published in the Journal of Educational Data Mining:

Sonnenberg, C., & Bannert, M. (2016). Evaluating the impact of instructional support using data mining and process mining: A micro-level analysis of the effectiveness of metacognitive prompts. *Journal of Educational Data Mining*, 8(2), 51-83.

Evaluating the Impact of Instructional Support Using Data Mining and Process Mining: A Micro-Level Analysis of the Effectiveness of Metacognitive Prompts

Christoph Sonnenberg

University of Wuerzburg, Instructional Media, Germany
christoph.sonnenberg@uni-wuerzburg.de

Maria Bannert

Technical University of Munich, Teaching and Learning with Digital Media, Germany
maria.bannert@tum.de

In computer-supported learning environments, the deployment of self-regulatory skills represents an essential prerequisite for successful learning. Metacognitive prompts are a promising type of instructional support to activate students' strategic learning activities. However, despite positive effects in previous studies, there are still a large number of students who do not benefit from provided support. Therefore, it may be necessary to consider explicitly the conditions under which a prompt is beneficial for a student, i.e., so-called adaptive scaffolding. The current study aims to (i) classify the effectiveness of prompts on regulatory behavior, (ii) investigate the correspondence of the classification with learning outcome, and (iii) discover the conditions under which prompts induce regulatory activities (i.e., the proper temporal positioning of prompts). The think-aloud data of an experiment in which metacognitive prompts supported the experimental group ($n = 35$) was used to distinguish between effective and non-effective prompts. Students' activities preceding the prompt presentation were analyzed using data mining and process mining techniques. The results indicate that approximately half of the presented prompts induced metacognitive learning activities as expected. Moreover, the number of induced monitoring activities correlates positively with transfer performance. Finally, the occurrence of orientation and monitoring activities, which are not well-embedded in the course of learning, increases the effectiveness of a presented prompt. In general, our findings demonstrate the benefits of investigating metacognitive support using process data, which can provide implications for the design of effective instructional support.

Keywords: self-regulated learning, instructional support, micro-level analysis, metacognitive prompting, think-aloud data, process mining

1. INTRODUCTION

The research in self-regulated learning (SRL) indicates that many learners have difficulties in spontaneously deploying regulatory activities, which results in lower learning performance (e.g., Azevedo, 2009; Bannert & Mengelkamp, 2013; Greene, Dellinger, Tüysüzoglu, & Costa, 2013;

Winne & Hadwin, 2008; Zimmerman, 2008). Therefore, our main objective is to provide effective instructional support for hypermedia learning. In this context, metacognitive prompting is a promising approach that affects the learning process by inducing regulatory activities as well as learning performance (e.g., Bannert & Mengelkamp, 2013; Bannert & Reimann, 2012). The purpose of these prompts is to foster SRL activities such as orientation, planning, monitoring, and evaluation strategies by asking students to monitor and to control their learning process (Bannert, 2009; Veenman, 1993). However, post-hoc analyses of students' prompt use revealed that approximately half of the sample demonstrated poor compliance with the provided support (Bannert & Mengelkamp, 2013). Thus the students did not benefit from the metacognitive prompting as intended. To improve the provided instructional support, it is our aim to investigate the conditions that influence the effectiveness of metacognitive prompts by taking into account fine-grained process data.

Referring to an event-based view of SRL that describes regulatory activities as dynamically unfolding over time during a learning task (e.g., Azevedo, 2009; Winne, 2014), the current research is increasingly interested in analyzing sequences of learning activities that are measured online during learning (e.g., measured by log files or think-aloud data). Moreover, the development and application of new methods that can take into account the sequential and temporal order of learning events (Martin & Sherin, 2013; Molenaar & Järvelä, 2014), as well as the dynamic relationship between SRL processes (Ben-Eliyahu & Bernacki, 2015), accompany the increasing importance of analyzing process data. In particular, the techniques in the field of educational data mining (EDM) have the potential to support the discovery of event patterns in SRL (Winne & Baker, 2013), for example, by modeling events that are crucial for an understanding of learning through the use of process mining (Reimann & Yacef, 2013; Trčka, Pechenizkiy, & van der Aalst, 2010). As a consequence, these recent developments provide new opportunities for the evaluation of instructional support on the micro level. More precisely, an in-depth process analysis contributes to the investigation of scaffolding effects (e.g., metacognitive prompting) on learning activities, and the evaluation results can inform researchers about how to develop their supporting strategies (e.g., Jeong et al., 2008; Johnson, Azevedo, & D'Mello, 2011; Molenaar & Chiu, 2014; Sonnenberg & Bannert, 2015). In general, we argue that the analysis of fine-grained process data is necessary for the development and the design of effective instructional support (e.g., Bannert & Mengelkamp, 2013; Sonnenberg & Bannert, 2015).

Hence, the aim of the current contribution is to demonstrate the potential of evaluating metacognitive prompting by analyzing process data (i.e., concurrent think-aloud protocols) by using data mining and process mining techniques. Because metacognitive prompts do not always optimally support students (e.g., Bannert & Mengelkamp, 2013; Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015), we investigate the conditions of effectiveness to derive implications for an improved prompt design. According to the concept of adaptive hypermedia systems and adaptive scaffolding (Bouchet, Harley, Trevors, & Azevedo, 2013; Brusilovsky, 2001, 2007; Molenaar & Roda, 2008), the system and the provided support should be tailored to the student's requirements, for example, to prior knowledge and to performed SRL activities (Azevedo, Cromley, & Seibert, 2004; Azevedo, Cromley, Winters, Moos, & Greene, 2005). Moreover, Molenaar and Roda (2008) note that the evaluation of scaffolds needs to be contingent on the learner's current activities and his or her goals. Therefore, it is crucial to examine the conditions that contribute to the effectiveness of a scaffold when its presentation successfully supports a student.

The present paper is organized as follows. First, we outline the effects of scaffolding hyper-

media learning through metacognitive prompts. We focus on the challenges regarding prompt design and the concept of adaptive scaffolding, including the results of the related research. Second, we introduce the analysis of fine-grained process data for the evaluation of instructional support. Third, we investigate the effects of prompting using the coded think-aloud data of an experimental study. To that end, we first classify the prompts by considering the increase of metacognitive utterances following the prompt presentation. Then we explore the conditions of effectiveness applying data mining and process mining techniques to the learning activities that precede each prompt. Finally, the significance of the findings for the design of the prompts and the development of a micro-level theory of SRL processes are discussed.

2. SCAFFOLDING HYPERMEDIA LEARNING THROUGH INSTRUCTIONAL SUPPORT

Especially in open-ended learning settings, the use of self-regulatory skills represents a predictor of learning success (e.g., Azevedo, 2005; Lin, Hmelo, Kinzer, & Secules, 1999). However, students often do not perform regulatory activities spontaneously, which in general results in lower learning outcomes (e.g., Azevedo, 2009; Bannert & Mengelkamp, 2013). Consequently, instructional support aims to counteract this deficiency by promoting the activation of strategic learning processes. For example, the research on metacognitive prompting has provided evidence for its beneficial effects on learning process and outcome (e.g., Azevedo, Cromley, Moos, Greene, & Winters, 2011; Bannert, 2009; Bannert et al., 2015). In a series of experiments comparing students supported by different types of metacognitive prompts to a control group without support, we found medium effects on metacognitive processes and transfer performance (Bannert & Mengelkamp, 2013). This magnitude is in line with meta-analyses on metacognitive instruction ($d = 0.59$; Hattie, 2009). The research on metacognitive prompting is presented in more detail below.

2.1. EFFECTS OF METACOGNITIVE PROMPTING

Metacognitive prompts have the purpose of inducing regulatory activities, for example, orientation, planning, monitoring, and evaluation strategies (Bannert, 2007, 2009; Veenman, 1993). According to theories of SRL (e.g., Winne & Hadwin, 2008; Zimmerman, 2008), the recurring deployment of these activities during a learning task is crucial for the successful regulation of one's learning process. For example, Zimmerman (2008) describes SRL as a cyclical model that comprises three phases, namely *forethought* (i.e., task analysis, goal setting, and planning), *performance* (i.e., strategy use, monitoring, and control), and *reflection* (i.e., self-evaluation). The current study refers to a framework for successful hypermedia learning (Bannert, 2007), which comprises the learning activities orientation, goal specification, planning, information search and relevance judgment, information processing, monitoring, and evaluation of goal attainment. Although these activities may imply a typical sequence, their performance can vary dynamically, considering the challenges of a given learning task. Additionally, the framework also considers the motivational aspects such as achievement motivation, action control, and self-efficacy, which are supposed to influence the strategy use in the process of learning. Especially students' motivational states with respect to the perception of their current task (e.g., the value or level of difficulty), their competencies (e.g., the ability to successfully use a specific strategy), and the learning situation (e.g., the authenticity of the situation) might affect their learning behavior.

For example, Zimmerman (2008) highlights the importance of motivational variables such as self-efficacy, task interest, value, and goal orientation, which influence the student's proactive pursuit towards his or her learning goals.

According to the expected effect mechanism, the presentation of metacognitive prompts fosters the activation of one's repertoire of metacognitive skills (e.g., Bannert & Mengelkamp, 2013). Metacognitive prompting thereby attempts to remedy the phenomenon of *production deficit* (e.g., Winne, 1996; Wirth, 2009). That is, students possess knowledge about regulatory capacities, but they do not use such skills spontaneously. Consequently, metacognitive prompts can foster the performance of learning activities that are described in the framework above and even the sequential order of their implementation during a learning task.

Several studies on metacognitive prompting have confirmed its positive effect on learning activities and learning outcome (e.g., Azevedo et al., 2011; Ge, 2013; Johnson et al., 2011; Kramarski & Gutman, 2006). However, the research also indicates that its effectiveness may depend on learner characteristics or the features of the prompt design (e.g., Bannert, 2009). For instance, the use of prompts requires additional cognitive capacities, which makes it necessary to have sufficient prior domain knowledge or to offer training in advance on how to use the prompts during learning (Bannert & Reimann, 2012; Veenman, 1993). Moreover, Thillmann, Künsting, Wirth, and Leutner (2009) found that strategy instruction must be embedded in the ongoing course of learning. Similarly, findings from cognitive support strengthen the necessity to integrate strategy instruction into ongoing cognitive learning activities (e.g., Wittwer & Renkl, 1998). However, a study on metacognitive feedback for help-seeking skills using an intelligent tutoring system showed that real-time feedback may be not beneficial to learning gains for all students (Aleven, Roll, McLaren, & Koedinger, 2016; Roll, Aleven, McLaren, & Koedinger, 2011). Possible explanations that were discussed by the authors are a cognitive overload that is caused by the real-time interventions or a lack of awareness of the value of the provided feedback (i.e., a motivational issue). As a result, more research is needed that investigates the conditions under which metacognitive support is beneficial.

In our previous experiments, we investigated the effects of different types of metacognitive prompts during hypermedia learning (Bannert & Mengelkamp, 2013; Bannert et al., 2015). The presentation of prompts took place during a 40-minute learning session, and the prompts stimulated or even suggested appropriate regulatory activities to university students. Overall, the results showed positive effects on learning processes, measured by coded think-aloud protocols ($0.35 < d < 1.17$) as well as on navigation behavior ($0.42 < d < 0.59$), and finally on transfer performance ($0.42 < d < 0.59$). With regard to the learning process, the students that were supported by prompts showed significantly more metacognitive learning activities and significantly better navigation behavior in comparison to the students in a control group who received no support. In addition, a detailed analysis that used process mining revealed differences among the process models of the students who learned with prompts and the students who learned without support (Sonnenberg & Bannert, 2015). These findings indicate that metacognitive prompting affects not only the frequency of regulatory activities but also the deployment of the sequences of learning activities.

In our most recent work (Bannert et al., 2015), students were asked to provide the reasons for their navigational decisions several times during learning (i.e., prompts to reflect on one's behavior). We found beneficial effects on navigation behavior (frequency of relevant pages visited: $p = .004$, $d = 0.65$; time spent on relevant pages: $p = .009$, $d = 0.58$) and on transfer performance ($p = .035$, $d = 0.44$) compared with those of a control group. The current study

extends this work by analyzing think-aloud data to gain a deeper insight into the effectiveness of the presented prompts and to address the challenges of prompt design better.

2.2. CHALLENGES OF PROMPT DESIGN

Despite the reported beneficial effects in the previous section, an optimal prompt design in open-ended learning environments remains challenging (Azevedo & Hadwin, 2005). In our studies, a post-hoc analysis based on videos and verbal protocols showed that in general only half of the sample used the provided prompts in the intended manner (Bannert & Mengelkamp, 2013). For example, students started to read the contents of the learning environment immediately instead of first planning their learning steps as requested by prompts, or they even ignored the prompts by not considering the requests at all. Moreover, some students reported that they felt restricted in their course of learning by the prompts. Consequently, the important question arises as to why many students do not comply with a provided support device (Clarebout & Elen, 2006; Clarebout, Elen, Collazo, Lust, & Jiang, 2013). The recent research has noted that the temporal positioning of a prompt within the learning process can significantly influence compliance and thereby the effectiveness of prompting (Azevedo et al., 2011; Sitzmann, Bell, Kraiger, & Kanar, 2009; Thillmann et al., 2009). As a consequence, more research that concerns the temporal characteristics of SRL is needed to inform the timing of scaffolds (Molenaar & Järvelä, 2014; Sonnenberg & Bannert, 2015).

In general, there are two possible approaches to improve the timing of a scaffold during hypermedia learning. First, researchers can try to involve students in designing their scaffolds. Because students should be the experts in their course of learning, it may be best to provide them with the opportunity to adapt the support to their needs (e.g., Bannert et al., 2015). Second, one can attempt to develop an adaptive hypermedia system that can diagnose the demands of a student and that can present a scaffold when it is needed based on specific embedded rules (e.g., Bouchet et al., 2013). Both are promising approaches for optimizing the positioning of metacognitive prompts during hypermedia learning, but they also present certain challenges, which are described below.

In a recent study (Bannert et al., 2015), we investigated the first approach, that is, the effects of so-called *self-directed metacognitive prompts*. The idea behind these prompts is to involve students in the configuration or even the creation of their prompts before learning, for example, by determining the time when a prompt should appear in their course of learning (e.g., after 5 minutes, after 12 minutes, and so on). Contrary to our expectations, despite the beneficial effects compared to a control group, there was no significant improvement of prompt use. Although the students had been familiarized with the hypermedia learning environment by performing some training tasks, it may be possible that they were overcharged with adapting the scaffolds to their needs. It is possible that they could not correctly determine when they should best be supported by a prompt.

Considering the second approach, the learning environment needs to diagnose the requirements of the learner simultaneously during learning, and it must determine when to position a scaffold, provided, for example, by a pedagogical agent (Bouchet et al., 2013; see the next section on adaptive support). This approach has the advantage of taking into account the current learning progress for presenting adaptive support, but both diagnosis and intervention can be difficult to develop and implement into a learning environment (Azevedo & Hadwin, 2005). Therefore, more research is needed that analyses the data of students' learning activities and

how scaffolds affect these activities to derive the conditions under which a scaffold can optimally support a learner. Moreover, these analyses can contribute to the development of more specific models that describe how to support self-regulated behavior and to develop rules for the implementation of adaptive scaffolds into learning environments. For example, which learning activities should trigger a support device or the appropriate timing to present a scaffold. As described in the following, related studies have already addressed the impact of adaptive support on the enhancement of computer-supported learning.

2.3. EFFECTS OF ADAPTIVE SUPPORT

In general, the research that has investigated adaptive scaffolding has confirmed the beneficial effects of individualized support on the regulation of learning processes and on learning outcome (e.g., Azevedo et al., 2004, 2005, 2011; Lehmann, Hähnlein, & Ifenthaler, 2014; Schwonke, Hauser, Nückles, & Renkl, 2006; Yeh, Chen, Hung, & Hwang, 2010). For example, Azevedo et al. (2005) compared the effects of adaptive scaffolding, fixed scaffolding, and no scaffolding during hypermedia learning. The results showed that adaptive scaffolding, which was provided by a human tutor, facilitated the shift in students' mental models significantly compared to the other two groups, and, further, it improved the deployment of regulatory strategies. Moreover, Lehmann et al. (2014) investigated the effectiveness of so-called pre-reflective and reflective prompts. They found benefits in the pre-reflective prompts, which stimulate to reflect on future events, but only for novice learners. Hence, they concluded that the adaptation of prompting in online research is crucial for future research.

Moreover, computerized settings such as adaptive hypermedia systems (Brusilovsky, 2001, 2007) provide new possibilities for the realization of adaptive support (e.g., Molenaar & Roda, 2008; Walker, Rummel, & Koedinger, 2011). Students' interactions with the system can be analyzed automatically and in real-time, thus informing the intervention type and presentation time of support devices. The current research investigating SRL and metacognitive behaviors uses some intelligent tutoring systems and learning environments such as BioWorld (Lajoie et al., 2013), the Geometry Cognitive Tutor (Aleven, 2013), and Crystal Island (Lester, Mott, Robison, Rowe, & Shores, 2013). These systems incorporate approaches to assess and scaffold SRL-behavior dynamically. A specific example of an existing adaptive hypermedia learning environment is MetaTutor, a system that provides scaffolds and feedback through several pedagogical agents (e.g., Azevedo et al., 2012; Bouchet et al., 2013). Its adaptive presentation is based on a set of system-generated rules that refer to the students' interaction with the system. For example, when a learner begins with a new sub-goal, a pedagogical agent prompts the learner to activate any prior knowledge that may be relevant to the sub-goal before starting to work on the content. Furthermore, a pedagogical agent asks the learners if they have adequately completed the current subgoal after they have spent more than 20 minutes working on it. In summary, MetaTutor uses the students' interactions with the hypermedia system as conditions for triggering the specific actions of pedagogical agents. Based on the analyses of learning activities with MetaTutor and the impact of pedagogical agents on learning processes and learning outcome, the system rules are constantly refined to optimize the provided support. Another example for an adaptive scaffolding environment is the AtgentSchool system (Molenaar & Roda, 2008), which was designed to provide dynamic and adaptive support to school children. In general, AtgentSchool is an attention management system that can diagnose a learner's focus using his or her activities within the system and provide appropriate interventions. The system works

adaptively and dynamically by calibrating the scaffolds to the student's progress and his or her characteristics.

Despite the beneficial impact of adaptive scaffolds and the possible implementation of adaptive rules in hypermedia systems, there are still many issues for future research. For example, Yeh et al. (2010) noted the missing theoretical understanding of prompt formats that are tailored to students' different levels of expertise. Moreover, the questions of how to diagnose the proper time to scaffold, how to calibrate the support for the appropriate phase of SRL, and how to gradually reduce support as students progress in self-regulating their learning (e.g., Azevedo & Hadwin, 2005) still need to be addressed in future studies. In the current contribution, we concentrate on the investigation of the appropriate temporal positioning of scaffolds by exploring the conditions of effective scaffolds using process data.

3. EVALUATION OF INSTRUCTIONAL SUPPORT USING PROCESS DATA

The recent research in SRL that emphasizes the investigation of learning as patterns of events (e.g., Azevedo, 2014; Bannert, Reimann, & Sonnenberg, 2014; Winne, 2014) has begun to concentrate on the microanalysis of process data, which is comprised of different traces of students' behavior during learning (e.g., log files, think-aloud data, or eye movements). The development and application of innovative analysis techniques that are appropriate for this type of data accompany the increasing interest in fine-grained process data. For example, the techniques that address the temporal characteristics and the dynamic relations of regulation activities (Ben-Eliyahu & Bernacki, 2015; Molenaar & Järvelä, 2014) or that support the discovery of event patterns in SRL activities (Winne & Baker, 2013). These approaches provide new potential for the evaluation of the effectiveness of scaffolds on the micro level. Furthermore, evaluation results can contribute to the advancement of a supporting strategy (e.g., Jeong et al., 2008; Johnson et al., 2011; Molenaar & Chiu, 2014; Sonnenberg & Bannert, 2015).

The analysis of process data using EDM techniques can stimulate the development and design of effective instructional support that is based on discovered process patterns. For example, Bouchet et al. (2013) investigated how adaptive versions of their learning environment respond differently to students, clustered by their SRL behavior. The authors drew implications for the implementation of increasingly adaptive, individualized support based on learners' profiles. Moreover, Kinnebrew, Segedy, and Biswas (2014) were able to track students' cognitive skills as well as their use of metacognitive skills in learning environments using log files and sequence mining methods. Additionally, the evaluation of scaffolds using process data can support the development of SRL micro-level models that comprise assumptions on the conditions of effective scaffolding (e.g., positioning, students' level of expertise, type of scaffold).

In our approach, we apply process mining techniques (PM; Reimann & Yacef, 2013; Trčka et al., 2010) to model the sequences of events that are crucial for understanding learning processes as well as the impact of the provided scaffolds on these processes. Those learning activities are measured by concurrent think-aloud protocols (Ericsson & Simon, 1993). In general, PM enables the discovery of process models from event sequences that are stored in an event log, the testing of models through conformance checking with additional data, and the extension of existing models (Trčka et al., 2010). We recommend PM as a promising method in SRL research (Bannert et al., 2014; Sonnenberg & Bannert, 2015) because it allows researchers to describe and to test learning models that incorporate a process-oriented view, and that can represent the workflow of activities (Van der Aalst, Weijters, & Maruster, 2004). These process models build

on the concept of Petri nets, which represent an executable system of places and transitions (Bannert et al., 2014). Using PM algorithms for discovery, a model can be generated based on event data. The functionality of a discovery algorithm is described in more detail in the section titled ‘Analysis techniques’. In the context of computer-supported learning research in particular, PM techniques are increasingly used to study learning from an event-based perspective (e.g., Malmberg, Järvelä, Järvenoja, & Panadero, 2015; Reimann, Markauskaite, & Bannert, 2014; Reimann & Yacef, 2013; Schoor & Bannert, 2012).

In previous analyses that have used PM, we compared the process models of students with high versus low learning performance and demonstrated that PM techniques can reveal differences in the sequential patterns of regulatory activities (Bannert et al., 2014). Furthermore, we investigated the effects of metacognitive prompts on the sequential structure of SRL activities by comparing the process models of students supported by prompts and of students in a control group without support (Sonnenberg & Bannert, 2015). Compared to the traditional frequency-based analysis of learning events, which is not able to take into account the sequential order of activities, additional findings on prompting effects were revealed. Now, we aim to obtain more detailed information on the learning process that precedes the presentation of a prompt to derive the conditions for its effectiveness. For example, discovered patterns could indicate that the performance of certain events or sequences of events is beneficial, or detrimental, for the effectiveness of a subsequent prompt presentation. Again, it is expected that PM can contribute to deeper insights into the sequence of learning activities.

4. RESEARCH QUESTIONS

Metacognitive prompts can stimulate the activation of one’s repertoire of strategic learning activities. Despite the beneficial effects of prompting on learning process and learning outcome, students often do not comply with the provided support optimally (Bannert & Mengelkamp, 2013). Poor compliance is most probably caused by a lack of tailoring of instructional support, that is, the conditions under which a scaffold is needed are often not considered. For instance, the presentation of a prompt may not be necessary, and it can even be disruptive at certain times in the learning process. Therefore, we argue that it is necessary to consider students’ current learning process to provide adequate instructional support. The conditions of the learner and his or her learning progress (e.g., learner characteristics, learning material, or current learning activity) need to be analyzed by process analysis to inform the decisions of when and how often a prompt should be presented, that is, the amount of scaffolding, the timing, and the fading out of support. In the present analysis, we aim to classify the effectiveness of metacognitive prompts by considering their impact on subsequent learning activities. Additionally, we focus on the conditions of effectiveness by analyzing the learning activities that precede each prompt. We address the following research questions in detail:

1. Is it possible to distinguish between metacognitive prompts with high and low effectiveness in terms of the activation of regulatory learning activities using process data?
2. Does the effectiveness of prompts correspond with learning outcome; that is, does a student who activates a higher number of metacognitive learning activities following a prompt presentation show a higher learning performance?

3. What are the conditions under which metacognitive prompts effectively induce regulatory activities?

With respect to the effect mechanism of metacognitive prompts, we expect that the presentation of an effective prompt enhances the deployment of metacognitive learning activities. Moreover, the number of induced metacognitive activities should be associated with the learning outcome. Finally, it should be possible to discover the conditions, that is, the learning activities that precede the prompt presentation that influence its effectiveness.

5. METHOD

The present analysis relates to the data of an experimental study that was reported in Bannert et al. (2015), which examined the general effects of metacognitive prompts on navigation parameters and learning performance. In the following analysis, we proceed by concentrating on a microanalysis using think-aloud data. Both contributions refer to the same participants, but they investigate different research questions and mainly consider different data sources.

5.1. SAMPLE AND RESEARCH DESIGN

The participants were $N = 70$ undergraduate students from a German university (mean age = 20.07, $SD = 1.88$, 82.9% female). The students were randomly assigned to the experimental group ($n = 35$, mean age = 20.29, $SD = 1.89$, 88.6% female), or to the control group ($n = 35$, mean age = 19.86, $SD = 1.87$, 77.1% female). All of the participants majored either in media communication or human-computer interaction, and their recruitment was accomplished through an online system that was administered by our institute. Each student received an incentive of 40 Euros (approximately \$47 USD) for their participation.

All of the students participated in a hypermedia learning session. The students who were assigned to the experimental group received support through metacognitive prompts, whereas the control group received no support during learning. Due to the research questions of the current study, we will only focus on the data of the experimental group. More detailed information about the procedure and the findings of the group comparison is reported in Bannert et al. (2015).

5.2. LEARNING ENVIRONMENT AND PERFORMANCE MEASUREMENT

Both the learning material and metacognitive prompts were presented in a hypermedia learning environment. The learning content comprised a chapter on learning theories (i.e., classical conditioning, operant conditioning, and observational learning). Altogether, this chapter included 50 nodes with approximately 13,000 words, 20 pictures and tables, and 300 hyperlinks. The pages that were relevant to the learning goals were limited to 10 nodes with approximately 2,300 words, five pictures and tables, and 60 hyperlinks. Thus, each node comprised approximately 230 words, and there was a figure on approximately every second page. All of the remaining nodes comprised overviews, summaries, and pages with content that were not relevant to the learning task.

It was possible to navigate within the learning environment using one of the following elements: (i) a hierarchical navigation menu, (ii) a next-page and previous-page button, (iii) the backward- and forward-button of the browser, and (iv) hyperlinks that were embedded in the content. Support through metacognitive prompts appeared as a pop-up window on the screen

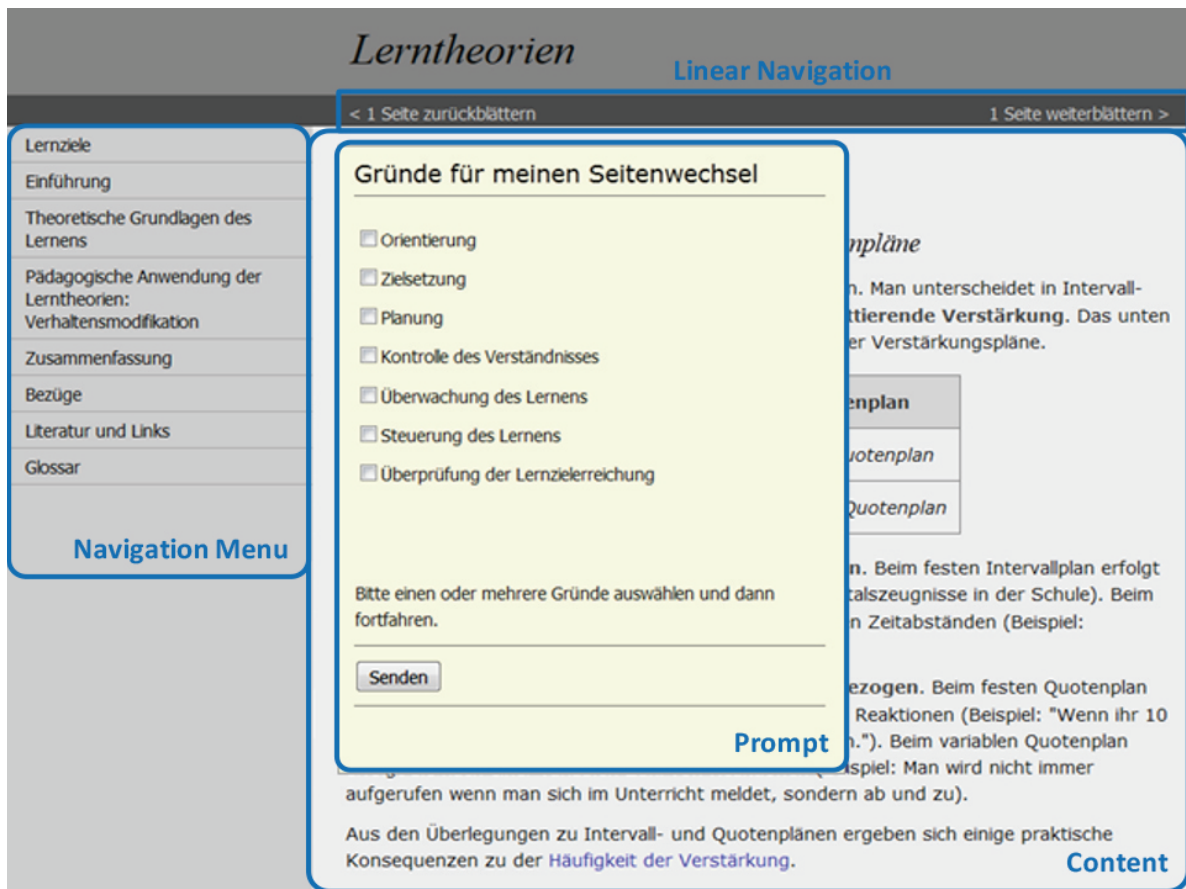


Figure 1: Screenshot of the learning environment including a metacognitive prompt. The main elements are labeled here for a better understanding of the environment, however it was not labeled in the learning sessions. Participants were asked to select at least one reason for node selection by choosing among a list of strategic reasons presented in the prompt window. The list comprised *orientation*, *goal-setting*, *planning*, *checking of understanding*, *monitoring of learning*, *control of learning*, and *evaluation of goal attainment*.

several times during learning. Each pop-up window comprised the same list of strategic reasons for node selection. Examples of these reasons are *orientation*, *goal-specification*, or *evaluation of goal attainment*. The participants had to select at least one reason for node selection before continuing with learning. Figure 1 presents a metacognitive prompt in the form of a pop-up window that was implemented in the learning environment.

Three knowledge tests on different levels based on Bloom's taxonomy of cognitive learning (Bloom, 1956) were used for performance measurement: (i) a free recall test, (ii) a comprehension test, and (iii) a transfer test. During the free recall test, the participants had to write down all of the basic concepts of operant conditioning that they could remember. The comprehension test assessed factual knowledge and comprised 22 multiple-choice items, each with one correct and three incorrect answers (Cronbach's $\alpha = .69$). For example, the students were asked which of the following terms describes a primary reinforcer: money, praising words, food, or any stimulus directly following the behavior. Finally, transfer performance was measured by instructing the participants to apply their knowledge of the basic concepts and facts to eight prototypical

situations in educational settings, which were not explicitly addressed in the learning material. For example, in one task they had to apply the principles of operant conditioning to solve a classroom situation in which a teacher experiences discipline problems. Two research assistants rated the answers to these situations on a researcher-developed rating scale (maximum score = 40 points; Cohen's $\kappa = .84$). In the case of disagreement among the raters, one of the authors determined the final score. More examples of the knowledge tests that were used in our studies are available in Bannert and Reimann (2012).

5.3. PROCEDURE

The hypermedia learning session began with an introductory section. First, the experimenter explained the navigation elements of the learning environment. Then, the participant was instructed to perform a series of exercises using a practice lesson while thinking aloud concurrently during the task. More specifically, he or she was asked to verbalize every thought that came to his or her mind, without any interpretation or justification. These instructions refer to level 2 verbalizations, as specified by Ericsson and Simon (1993). If necessary, the experimenter provided feedback to the participant. Additionally, further exercises could be used until the participant firmly mastered the think-aloud technique.

Next, the participants received a short tutorial regarding the use of metacognitive prompts (approximately 10 minutes). This tutorial comprised information on the importance of reflecting on one's learning activities, an explanation of the reasons for strategic node selection that is listed in the prompts, and the desired usage of the prompts. Such a tutorial is necessary to guarantee the adequate application of metacognitive strategies during learning (e.g., Veenman, 2007). We derived the list of reasons for node selection from categories that had been developed in previous work (Bannert, 2006), in which students were asked to name their reasons freely. Following the tutorial, the participants were instructed to configure the prompts by the following arrangements. First, they could change the order of the items in the list of reasons for node selection. The initial order was randomized. Second, they were asked to define eight time stamps when the prompts should appear during learning.

As the next step, the participants engaged in 40 minutes of learning in our hypermedia learning environment. In the beginning, they were instructed about their learning task, that is, to learn the basic concepts of operant conditioning. During learning, the students were supported by metacognitive prompts. All of the participants were completely free to use the navigation elements of the learning environment and to use their learning strategies. During the whole learning phase, the participants had to read and to think aloud as practiced in advance, and their utterances were recorded using a microphone. If a participant stopped his or her verbalizations for more than five seconds, the experimenter prompted him or her to continue by saying "Please think aloud". Following the learning task, the participants worked on the three knowledge tests that are described above. The total duration of the session was approximately two hours.

5.4. CODING SCHEME

Concurrent think-aloud protocols were used for the online measurement of the learning activities. The participants' recorded verbal protocols were segmented and coded post-hoc according to a coding scheme that was based on our theoretical framework of self-regulated hypermedia learning (Bannert, 2007). This conceptual framework organizes the student's activities into the three major categories *Metacognition*, *Cognition*, and *Motivation*. Additionally, it comprises

several sub-categories, as further described in Table 1. Motivation refers to statements that reflect a beneficial or an obstructive embedding of a metacognitive or cognitive learning activity, particularly according to the concepts of achievement motivation, action control, and self-efficacy. More specifically, we assigned a motivational code when a student made an evaluative statement with respect to the task (e.g., “The task is quite difficult”), to his or her competencies (e.g., “I’m good at finding the relevant information”), and to the situation (e.g., “Thinking aloud isn’t as troublesome as I expected”). These utterances reflect motivational states that might affect the self-regulatory behavior, for example, the use of strategies. According to Zimmerman (2008) especially self-motivation beliefs, such as self-efficacy, task interest, and value, might impact the student’s engagement during learning.

In general, the coding process followed the procedure that was presented by Chi (1997). We segmented the verbal protocols by units of meaning. Thus, we assigned a segment for every definable learning activity. Furthermore, we did not use multiple nor nested codes. Four trained research assistants coded the verbalizations of all of the participants. We selected a random sample of six participants to compute the inter-rater reliability for our coding scheme. Based on 1,385 segments, the reliability showed substantial agreement (Cohen’s $\kappa = 0.78$), which is considered to be sufficient for the following analysis.

5.5. DATA PREPARATION

For the purpose of our analysis, we needed to prepare our data as described below. The starting point for the data preparation was the coded verbal protocols of 35 participants during 40 minutes of hypermedia learning. Each one of the students was supported by metacognitive prompts five to eight times during learning. The absolute frequencies and means of all of the coded learning activities, as well as the mean duration times of the events, are presented in Table 2.

The basic idea of our analysis was to use the coded learning activities, which represent the metacognitive, cognitive, and motivational utterances as described in the previous section, to classify each metacognitive prompt as “effective” or “non-effective” in inducing regulatory activities. We consider an increase in metacognitive utterances following a prompt, in relation to a student’s individual baseline of metacognitive activities, to be an indicator of effectiveness. Therefore, we started our investigation by using a *Dotted Chart Analysis* (Song & van der Aalst, 2007; Van der Aalst, 2011), which is applicable with the ProM framework Version 5.2 (2008), to obtain an overview of the present process data. A dotted chart illustrates a sequence of events by arranging them as dots in a two-dimensional plane. The horizontal axis represents the occurrence of an event in time, and the vertical axis accounts for a case (i.e., the data of one participant). Figure 2 shows the visualization of the coded learning activities during the first ten minutes of four sample cases (i.e., participants CHRO10, KAHA17, KOGE04, and VAIV20). An inspection of the events following a prompt (i.e., the events following the yellow sections) showed that some prompts are followed by a high number of metacognitive events, whereas other prompts are not. Furthermore, the increase in metacognitive events that follow a prompt—if present—usually takes up to two minutes.

Table 1: Coding scheme for analyzing student's learning activities

Code	Category	Description and Examples
Metacognition		
ORIENT	Orientation	Task clarification, overviewing the material to prepare strategic learning behavior <i>At first I read my learning goals to get an overview of my task.</i>
SETGOAL	Goal Specification	Goal setting and sub-goaling <i>I have to learn the basic concepts on operant conditioning.</i>
PLAN	Planning	Planning of proceeding <i>First I will read the introductory text, then I will decide in which sequence I will proceed.</i>
SEARCH	Search	Searching information <i>Now I'm looking for information on reinforcement plans.</i>
JUDGE	Judgement	Judgements of relevance of information <i>Skinner's Vita is not relevant for my learning task.</i>
EVAL	Evaluation	Evaluating the attainment of goals or sub-goals <i>Did I process all topics according to my learning goals?</i>
MONITOR	Monitoring	Monitoring and controlling of one's learning <i>Ah now I understand the principle and I can proceed to my next learning goal.</i>
Cognition		
READ	Reading	Reading out loud
REPEAT	Repeating	Repeating in terms of memorizing <i>Re-reading a paragraph or notes</i>
ELABORATE	Elaboration	Deeper processing: paraphrasing, connecting, inferring <i>I already know the Skinner Box from my biology class.</i>
ORGANIZATION	Organization	Organization of information <i>Drawing a map, writing down major concepts</i>
Motivation		
MOT	Motivation	Evaluative statements regarding the task, the student's competencies, or the situation <i>The task is very interesting and relevant for my studies. I'm good at memorizing this subchapter. This section distracts me from my original goal.</i>
Residual Category		
OTHER	Others	Off-topic statements, comments on technique, not interpretable statements, pauses <i>May I make notes? The mouse doesn't work well.</i>

Table 2: Absolute frequencies and means of all coded learning events, and mean duration times of events during 40 minutes of hypermedia learning (N = 35)

	Absolute Frequency					Duration	
	Min	Max	Frequency	M	SD	M_Dur	SD_Dur
Metacognition	34	242	4075	116.43	45.97	4.54	7.16
Orientation	5	30	501	14.31	7.23	6.10	5.19
Planning	0	5	61	1.74	1.65	5.76	3.21
Goal Specification	0	10	72	2.06	2.36	6.14	3.09
Search	1	32	414	11.83	7.54	8.53	9.18
Judgment	2	23	409	11.69	5.70	4.48	2.61
Evaluation	0	15	127	3.63	3.08	13.41	29.46
Monitoring	11	203	2491	71.17	37.62	3.07	2.74
Cognition	47	201	3760	107.43	36.01	14.72	14.93
Reading	20	84	1423	40.66	15.75	17.56	18.20
Repeating	2	45	647	18.49	11.04	10.82	10.47
Elaborating	3	55	767	21.91	12.92	10.35	11.18
Organizing	3	61	923	26.37	13.87	16.62	13.12
Motivation	0	18	72	2.06	4.14	4.07	2.46
Others	9	76	1124	32.11	14.71	3.25	4.01
Sum of all coded events	127	482	9031	258.03	78.36	8.87	12.08

Note. M(Dur) = mean duration time, SD(Dur) = standard deviation of duration time; duration times in seconds.

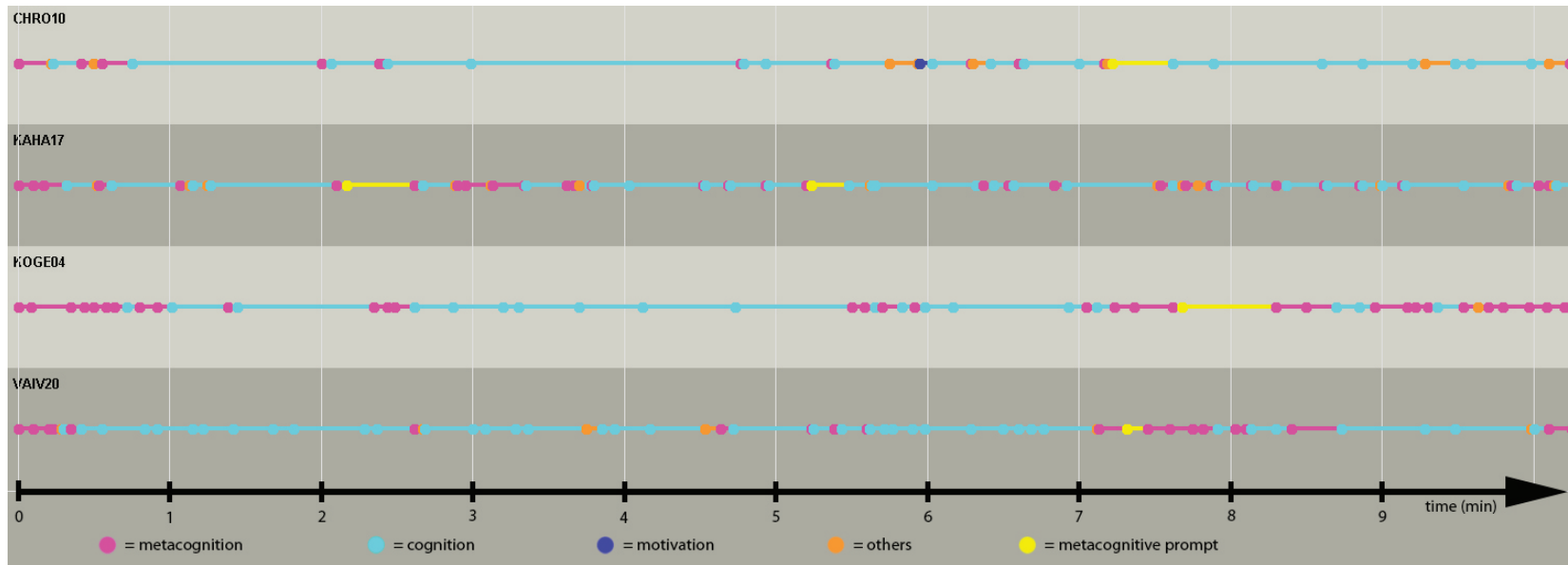


Figure 2: Visualization of four cases using a Dotted Chart Analysis of coded verbal protocols. The horizontal axis represents the first ten minutes of learning with each segment representing one minute. Each dot represents the occurrence of an activity or a prompt presentation. The activities are color-coded and explained at the bottom. The line between two dots represents the duration of an event.

Based on this observation, we decided to use a two-minute time interval following each prompt to determine the successful induction of metacognitive activities. There are additional reasons that support the selection of this measurement unit. A shorter time interval would reduce the sample size to investigate the learning process, that is, the number of investigated events. Here, one must consider that the duration time of learning activities varies from a few seconds (e.g., monitoring one's learning progress) to longer time periods (e.g., reading a paragraph). Table 2 presents the means and standard deviations for the duration times of all of the coded categories. On the other hand, the selection of a longer time interval would be accompanied by a possible overlap of prompt presentations. The participants were instructed to determine eight time stamps for the presentation of prompts, considering a distance of at least two minutes between two presentation times. Each selected time interval comprised all of the coded utterances within this timeframe, including overlapping events, that is, events that started and respectively ended beyond the two-minute interval. Because the duration of events varied from seconds to minutes, the number of events that are included in a time interval also varied ($M = 14.16$, $SD = 5.86$, $Min = 1$, $Max = 31$).

In addition to selecting the two minutes following each prompt, we also computed an *individual baseline* of metacognitive utterances for each participant. This baseline is necessary to determine if an increase of metacognitive events following a prompt occurred. It represents the amount of spontaneous metacognitive utterances within an interval of two minutes. For the computation of this baseline, we excluded the two-minute time interval following each prompt because this time span is supposed to be affected by the prompted requests. The remaining time span of the total learning time was used to compute a baseline for each participant. For example, one participant showed a total of 127 metacognitive utterances within the total learning time of 40 minutes. He or she received eight prompts with 64 metacognitive utterances within the two-minute time intervals following each prompt. For the individual baseline, we considered the number of remaining metacognitive events (i.e., $127 - 64 = 63$) and the remaining time span (i.e., $40 - 16 = 24$ minutes). In this case, the computation yields a baseline of 5.25 metacognitive utterances within a time span of two minutes. Overall, the individual baselines for all of the students ranged from 1.47 to 12.68 metacognitive utterances ($M = 5.32$, $SD = 2.29$).

Although the observed learning time of 40 minutes was quite short, and, therefore, we did not expect the number of metacognitive utterances to vary significantly by time, we conducted additional analyses to account for a possible influence of time or the number of prompts that were received. First, we checked if the number of prompts that are received predicts the number of metacognitive utterances following a prompt (i.e., a possible additivity of prompting effects). The results of a linear regression showed no significant correlation between prompt number and metacognitive utterances, $F(1, 239) = 3.16$, $p = .077$, $R^2 = .013$, $R^2_{adjusted} = .009$. Figure 3 presents the mean numbers of metacognitive utterances for all of the students ($N = 35$) in each two-minute time interval following a prompt. Second, we considered whether there was a general upward drift of metacognitive utterances during the total learning time. Again, a linear regression using the time interval as the predictor and the number of metacognitive utterances as the independent variable revealed no significant influence of time, $F(1, 698) = 2.36$, $p = .125$, $R^2 = .003$, $R^2_{adjusted} = .002$. Additionally, Figure 4 shows that no upward trend is present. The high number of metacognitive events during the first time interval can be explained by necessary orientation activities at the beginning of the learning phase. To summarize, the number of metacognitive utterances is not significantly affected by time or by the number of prompts that are received. Therefore, the computation of the baseline as described above does not need to be

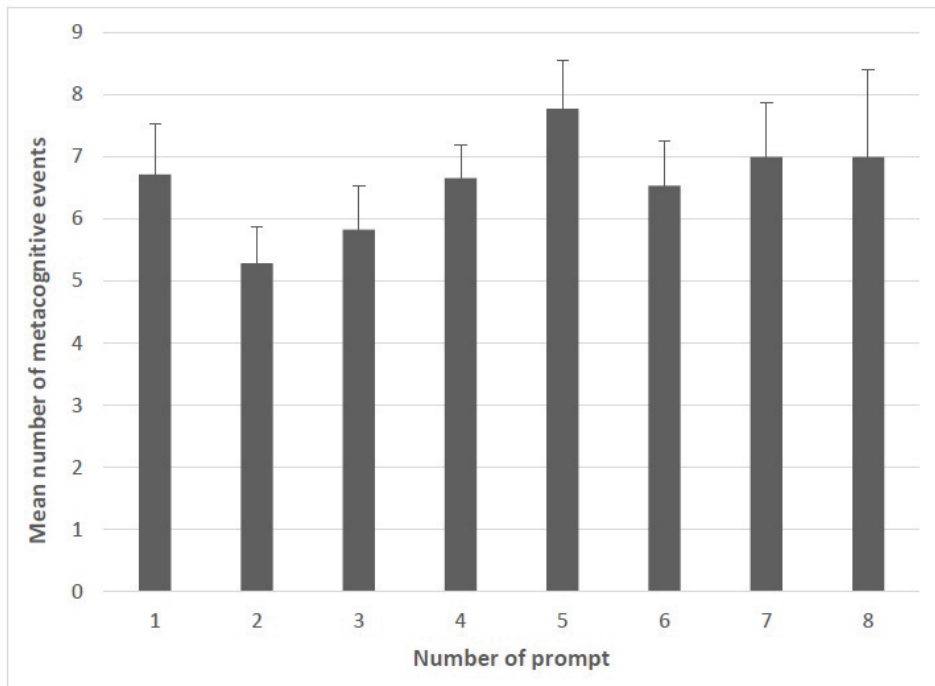


Figure 3: The mean number of metacognitive utterances for all students ($N = 35$) in each two-minute time interval following a prompt. Error bars represent standard errors.

adjusted for these factors.

By comparing the individual baseline with the number of metacognitive utterances in the two-minute time interval following a prompt, it is possible to determine whether there was an increase of metacognitive activities. If the number of metacognitive utterances exceeds the individual baseline, we regard the prompt as effective. Because it may be possible that both values are equal or very close to one another, we decided to set a cutoff value to solve those cases. We selected an absolute value of one metacognitive event because it represents the smallest possible threshold. For example, if the individual baseline of a participant is 5.30, then the number of metacognitive utterances in the time interval following a prompt must be greater than 6.30 to indicate an increase, that is, a successful induction of metacognitive activities based on our criteria.

Finally, we selected two-minute time intervals preceding the presentation of prompts for the purpose of our third research question, that is, an analysis of the conditions of effectiveness. Again, we decided that this time span represents a suitable sample for our analysis. The selection of time intervals that precede and follow a prompt presentation is illustrated in Table 3, using two prompting times of an example case.

5.6. ANALYSIS TECHNIQUES

In the first step of our analysis, we determined the number of metacognitive activities within each two-minute time interval following a prompt. Then, we verified the chosen type of classification by comparing the students' mean baseline of metacognitive utterances with the mean number in the time intervals that follow each prompt: If the prompts induced metacognitive activities

Table 3: The selection of two-minute time intervals preceding and following the presentation of the first two prompts for the case EDM18

<i>Case ID</i>	<i>Activity (Code)</i>	<i>Timestamp</i>	<i>Case ID</i>	<i>Activity (Code)</i>	<i>Timestamp</i>
	
EDMI18	READ	00:04:46	EDMI18	MONITOR	00:11:13
EDMI18	PLAN	00:06:08	EDMI18	ORGANIZATION	00:11:19
EDMI18	MONITOR	00:06:09	EDMI18	OTHER	00:11:53
EDMI18	READ	00:06:14	EDMI18	ELABORATE	00:11:56
EDMI18	MONITOR	00:06:48	EDMI18	ELABORATE	00:12:38
EDMI18	PROMPT1	00:06:56	EDMI18	OTHER	00:12:50
EDMI18	OTHER	00:07:19	EDMI18	ELABORATE	00:12:52
EDMI18	READ	00:07:44	EDMI18	PLAN	00:12:56
EDMI18	MONITOR	00:07:46	EDMI18	MONITOR	00:13:00
EDMI18	ORGANIZATION	00:08:10	EDMI18	EVAL	00:13:02
EDMI18	OTHER	00:08:13	EDMI18	MONITOR	00:13:10
EDMI18	MONITOR	00:08:19	EDMI18	PROMPT2	00:13:13
EDMI18	ELABORATE	00:08:25	EDMI18	OTHER	00:13:24
EDMI18	MONITOR	00:08:27	EDMI18	SETGOAL	00:13:26
EDMI18	ORGANIZATION	00:08:52	EDMI18	ORGANIZATION	00:13:27
EDMI18	MONITOR	00:08:53	EDMI18	READ	00:13:35
EDMI18	READ	00:09:04	EDMI18	ORGANIZATION	00:13:59
EDMI18	MONITOR	00:09:04	EDMI18	READ	00:14:25
EDMI18	ORGANIZATION	00:09:10	EDMI18	MONITOR	00:14:31
	...		EDMI18	ORGANIZATION	00:14:32
			EDMI18	READ	00:14:38
			EDMI18	MONITOR	00:15:09
			EDMI18	ORGANIZATION	00:15:11
			EDMI18	OTHER	00:15:24
				...	

Note. For an explanation of the codes please see the coding scheme presented in Table 1. The timestamp indicates the starting time of an activity, using the format hh:mm:ss.

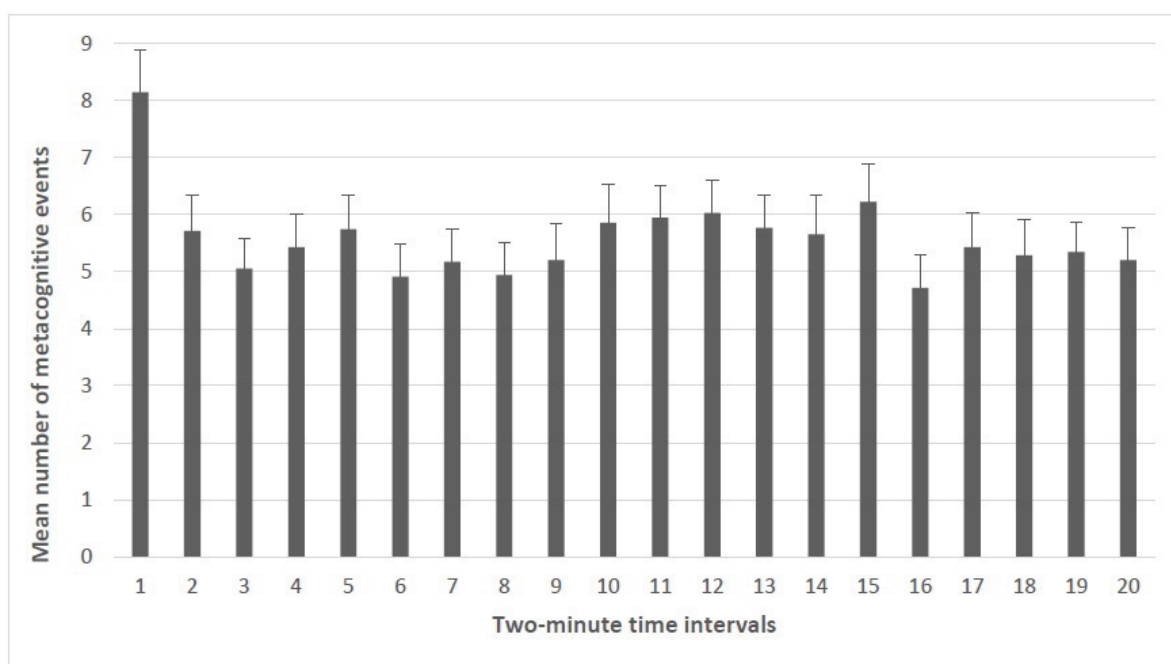


Figure 4: The mean number of metacognitive utterances for all students ($N = 35$) in each two-minute time interval. The total learning time was 40 minutes, resulting in 20 two-minute time intervals. Error bars represent standard errors.

within the determined time interval successfully, the mean number of metacognitive utterances in these intervals should be higher than the mean baseline. Additionally, we examined the correlation between the number of induced metacognitive activities and learning performance. According to our theoretical framework, we assumed that students who show a higher number of induced metacognitive activities should perform better than students who show less induced regulatory activities.

The second step of our analysis considered the conditions of effectiveness. For this purpose, we first used the RapidMiner Studio Version 6.3 (2015) and a linear regression learner to explore the learning activities that occurred before a prompt appeared. A linear regression can be used as a classification method in the case of exclusive numeric attributes, and it has the advantages of working well on small data sets and of generating simple models to enable a natural interpretation (Hämäläinen & Vinni, 2010). As described above, we decided to use two-minute time intervals for our analysis. The coded learning activities that occurred before each prompt were used as the predictors, and the number of metacognitive events that followed each prompt as the outcome variable. We applied cross-validations to avoid over-fitting of the learning algorithm. By using this data mining (DM) approach, we expected to find conditions, that is, states of the learning process, that indicate when a prompt is more likely to be effective and non-effective, respectively, in inducing regulatory activities.

Finally, we applied PM to gain more detailed insights into the learning process preceding the effective versus non-effective prompts by taking into account the sequential order of the learning activities. For this analysis, it was necessary to classify the prompts into the discrete categories of “effective” and “non-effective” in inducing metacognitive activities, as described in the section titled “Data Preparation”. We used the ProM Framework Version 5.2 (2009) and

the Fuzzy Miner algorithm (Günther & van der Aalst, 2007) to generate inductively a process model for the learning events that preceded all of the prompts that were classified as effective and a process model for the activities that preceded all of the prompts that were classified as non-effective. A comparison of these process models provides more detailed information on the conditions that may influence the effectiveness of prompts. In the following section, we provide a short introduction to the principle of the Fuzzy Miner algorithm.

Fuzzy Mining (Günther & van der Aalst, 2007) is an approach that was designed to find underlying processes in data that are less structured in appearance, such as our coded learning activities. The algorithm allows a flexible simplification of output models by distinguishing between the important and the less important details of an input event sequence. Thus, it is possible to generate a model that emphasizes the main features of a process and that is easily understandable. More precisely, the Fuzzy Miner transforms an event sequence into a process model that consists of nodes (event classes or categories) and edges (relations between two event classes). The data input comprises several cases with every case including an event sequence, which is ordered by a timestamp. First, the Fuzzy Miner algorithm uses these data to generate a complete model that comprises all of the observed nodes and edges by taking into account the relative importance and the sequential order of all of the events. Then, the algorithm uses two fundamental metrics, which are referred to as *significance* and *correlation*, to compute a simplified model for the given data set. *Significance* measures the relative importance of the occurrence of event classes and relationships between events. For instance, events that occur more frequently are assessed as being more significant. *Correlation* is calculated for edges. It indicates the closeness of two events, measured by their temporal proximity. The basic concepts of significance and correlation are embedded in a metrics framework that calculates three primary types of metrics: unary significance (event classes), binary significance (relationships between event classes), and binary correlation (relationships between event classes). Günther and van der Aalst (2007) describe these metrics in more detail. As a final step, the model is simplified by making decisions regarding the inclusion of nodes and edges in the final model using the following rules: Events that are highly significant are preserved, events that are less significant, but highly correlated, are aggregated, and events that are less significant and lowly correlated are removed. It is possible to influence the model simplification by parameter setting, for example, by specifying cutoff values. To bring structure to the model, the algorithm uses edge filtering and thereby tries to focus only on the most important relationships between event classes. The utility of edges, which is the weighted sum of significance and correlation of an edge, is calculated, and this weighting is then configured by the utility ratio. Moreover, by setting an edge cutoff, an absolute threshold value for filtering edges can be determined. The higher the value at which the edge cutoff is set, the more likely the Fuzzy Miner is to remove an edge. Finally, there is another important mean to simplify the model: node aggregation and abstraction. Nodes are removed based on a parameter referred to as the node cutoff. If the unary significance of a node is below this cutoff, it will be excluded from the resulting model, or it will be aggregated. The latter happens if it is possible to preserve less significant nodes by merging them into a cluster of highly correlated nodes.

6. RESULTS

This section presents the results of our analyses as follows. First, we report the classification results using the learning activities following each prompt and an individual baseline of

metacognitive utterances. Additionally, we refer to the correlation between our classification and learning performance. Second, the findings of the investigation of conditions for the effectiveness of prompts are reported. We further present the results of a linear regression learner as well as those of a process mining algorithm.

6.1. CLASSIFICATION OF PROMPTS

In the first step of our analysis, we took into account the number of metacognitive utterances in a two-minute time interval following each prompt as well as an individual baseline to determine each prompt's effectiveness in inducing regulatory activities. To verify our classification criteria, we used a Wilcoxon signed-rank test to compare the participants' mean baseline and the mean number of metacognitive utterances following each prompt presentation. The number following the prompts was significantly higher ($Mdn = 6.80$) than the individual baseline ($Mdn = 5.05$), $z = -4.18$, $p < .001$, $d = 2.00$. Thus, in general, the participants showed significantly more metacognitive activities within the two-minute time intervals following each prompt compared to their individual baseline (i.e., the mean number of metacognitive utterances within any two minutes). The results support the selection of a two-minute time span following a prompt to evaluate the successful induction of metacognitive activities.

Next, we analyzed the correlation between the induced number of metacognitive utterances through prompts and learning outcome. The alpha level was adjusted according to the procedure of Benjamini and Hochberg (1995) because we conducted multiple tests of statistical significance. The results showed no significant correlation for any of the performance measurements (recall: $r = -.23$, comprehension: $r = .08$, transfer: $r = .12$). Because the total number of metacognitive activities within 40-minutes learning time also showed no significant correlation with learning outcome, we checked the correlations with the sub-categories. The findings revealed that the sub-category Monitoring shows the highest positive correlation with two measurements of learning performance (recall: $r = -.17$, *n.s.*, comprehension: $r = .25$, *n.s.*, transfer: $r = .27$, *n.s.*). Therefore, we additionally analyzed the correlation between the number of monitoring activities within the two-minute intervals following the prompts and learning performance. The number of monitoring activities corresponds positively with transfer performance ($r = .26$, *n.s.*) and comprehension performance ($r = .22$, *n.s.*), but not with recall performance (recall: $r = -.18$, *n.s.*). Possibly due to the small sample size, all of the correlations are non-significant. A regression equation to predict transfer performance ($y = 18.95 + 0.068 x$) shows that a low number of induced monitoring activities through prompting ($M = 24.54$, $SD = 15.89$; $M - SD = 8.65$) results in a performance score of $y = 19.54$ points and that a high number ($M + SD = 40.43$) results in a score of $y = 21.70$ points. These observations are in line with previous work (Bannert & Mengelkamp, 2013), which indicated that prompting primarily enhances transfer performance. Additionally, a process analysis showed that the positive effect on transfer performance is mainly mediated by inducing monitoring activities (Sonnenberg & Bannert, 2015). In conclusion, the positive correlation between the number of monitoring activities and transfer performance indicates the usefulness of our classification based on a two-minute time interval following each prompt. The students who showed a higher number of induced monitoring activities through prompting performed better at transfer tasks than the students who showed less monitoring events.

Moreover, the classification of prompts into two discrete classes (effective and non-effective, respectively, in inducing metacognitive activities), based on a comparison of the individual base-

line and the number of metacognitive utterances following a prompt, resulted in a total of $n = 113$ effective prompts and $n = 127$ non-effective prompts for our learner sample ($N = 35$). Furthermore, the classification shows that each student received a mean of $M = 3.23$ ($SD = 1.46$, $Min = 1$, $Max = 6$) effective prompts. Except for one student whose prompts were all classified as being effective, each student received both effective and non-effective prompts.

The selection of two-minute time intervals following the presentation of a prompt and the classification of prompts represents the starting point for the following analyses, which aim at exploring conditions for effectiveness. Please note that for the purposes of these analyses, each prompt was regarded as a separate case labeled either as “effective” or “non-effective” ($N = 240$, $n_{effective} = 113$, $n_{non-effective} = 127$).

6.2. CONDITIONS FOR THE EFFECTIVENESS OF PROMPTS

We analyzed the coded events that preceded effective and non-effective prompts by using two approaches. First, we applied a linear regression learner taking into account the absolute frequency of coded events as predictors and the number of metacognitive events following a prompt as the outcome variable. Then, we considered the sequential order of coded learning activities by generating process models for the time interval before a prompt appeared using a PM algorithm.

LINEAR REGRESSION LEARNER. We applied a linear regression learner to our data set using the RapidMiner Studio Version 6.3 (2015). Because we used the same data set for training and for measuring the performance of the linear regression model, we used two cross-validation techniques to avoid over-fitting. First, a ten-fold cross-validation was conducted, randomly splitting the data into a training and a test set at the prompt-level ($N = 240$ prompts). Second, we conducted a student-level cross-validation by defining a subset for each student (in RapidMiner: Batch-X-Validation) to account for the possible influence of student factors. In this procedure, complete cases are removed iteratively before training the algorithm and are then used to estimate the goodness of the model. Moreover, we activated the automatic feature selection using the M5 prime method. The input predictors were all learning activities that occurred in a two-minute time interval preceding each prompt, whereas the outcome variable was the number of metacognitive utterances that followed each prompt. We used the VIF and tolerance statistics to assess potential multicollinearity within our data. Because VIF values were well below 10 (average VIF = 1.182) and tolerance statistics were well below 0.2, there was no cause for concern. The feature selection method selected the categories Orientation, Search, and Monitoring as relevant predictors for the regression model.

Table 4 presents the weight table of the linear regression model that was trained with our data. The results show the influence of the regression coefficients. Based on these coefficients, the findings indicate that a high occurrence of orientation and monitoring activities preceding the prompt presentation are significant predictors for the successful induction of metacognitive events. The occurrence of all other learning activities has no significant influence on prompt effectiveness in the resulting linear model. The performance measurement of the linear regression model using a ten-fold cross-validation showed a root-mean-squared error of 3.64 ± 0.55 and a correlation of $.46 \pm .19$. In case of the student-level cross-validation, the model goodness was slightly lower ($RMSE = 3.60 \pm 1.23$, $R = .22 \pm .26$).

Table 4: Weight table of the learned linear regression model

	B	$SE B$	β	p
Constant	3.40	0.44		< .001
Orientation	0.78	0.20	.24	< .001
Search	0.13	0.22	.01	.550
Monitoring	0.61	0.08	.05	< .001

Note. Prediction of the metacognitive activities following a prompt by considering all activities in a two-minute time interval preceding the prompt presentation. $N = 240$ prompts. All codes were included, but only three codes were selected by feature selection. Estimation of model goodness using cross-validation: prompt-level, $RMSE = 3.64$, $R^2 = .24$; student-level: $RMSE = 3.60$, $R^2 = .12$.

PROCESS MINING USING THE FUZZY MINER ALGORITHM. In a further analysis, we used the approach of PM and an algorithm for process discovery. This algorithm inductively generated process models for the time interval before an effective prompt appeared and before a non-effective prompt appeared. Our analysis was conducted by using the following parameter settings: edge filtering was set with the edge cutoff = 0.20 and the utility ratio = 0.75 (both are default values); the significance cutoff of the node filter was set to 0.20.

Figure 5 shows the resulting models with the event classes and their process relationships. Event classes are represented by the rectangular nodes that include the label and its significance (a value between 0 and 1). The arcs between categories indicate successive events (the upper number displays significance, that is, their relative importance, and the lower number shows correlation, that is, a measure which indicates how closely related two events are). The arcs that point towards a category indicate a repeated occurrence of that category. Less significant and lowly correlated events were discarded from the process model. Thus, the nodes and arcs that fall into this category were not included in the graph.

In Figure 5, the model of the learning process preceding non-effective prompts (left part) comprises nine event classes, and the model of effective prompts (right part) comprises seven categories with two clusters. The model of non-effective prompts does not include the categories SETGOAL and PLAN, whereas in the model of effective prompts, the categories PLAN and SEARCH as well as JUDGE and EVAL were combined into a cluster. This means that these event classes did not reach the significance cutoff value. Here, the main idea of the Fuzzy Miner is realized by abstracting from information that is perceived as being too fuzzy, that is, it does not play a significant role in the process as a whole.

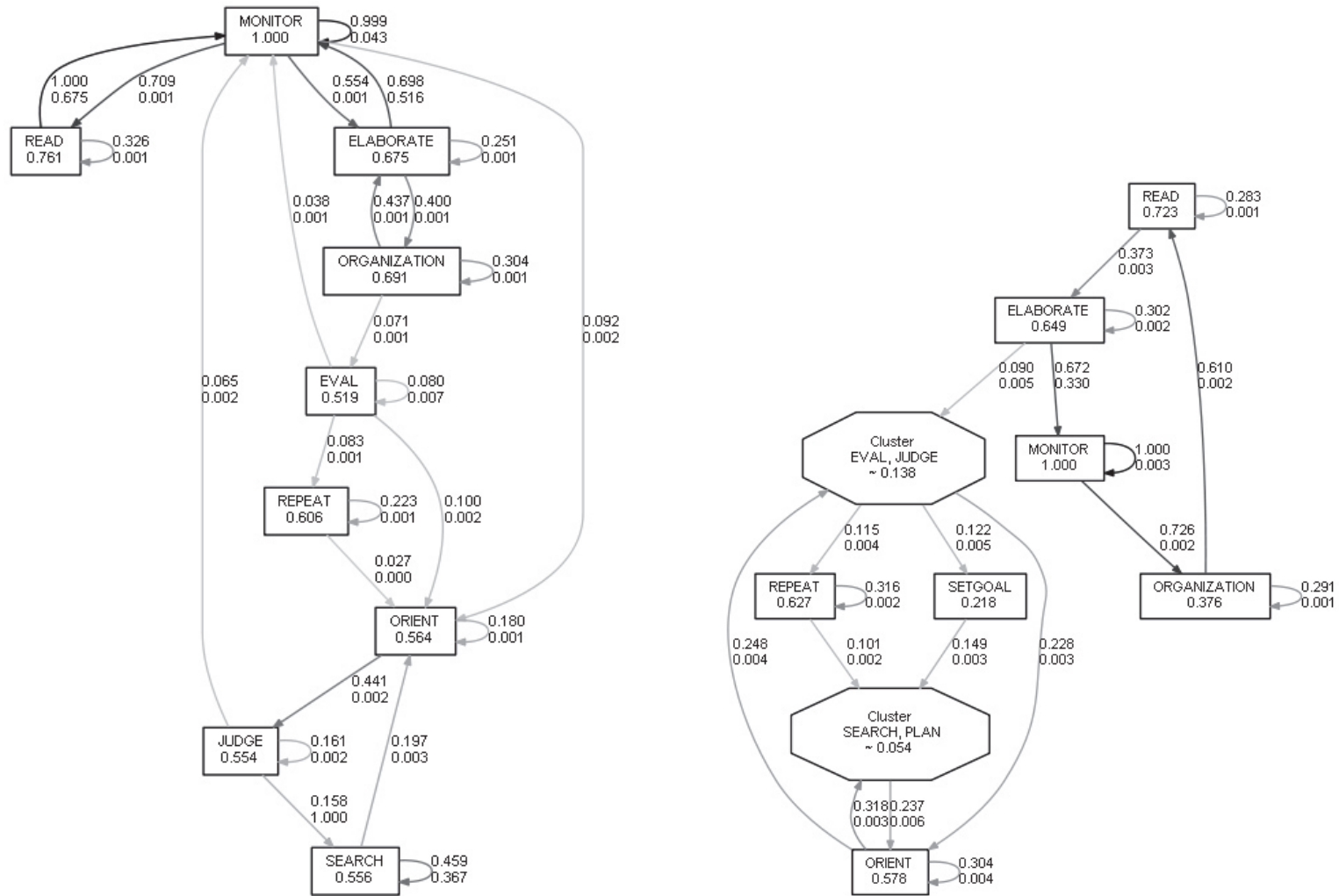


Figure 5: Process models for the time interval preceding non-effective prompts ($n = 127$, left) and effective prompts ($n = 113$, right). For an explanation of the codes please see the coding scheme presented in Table 1.

The process model of non-effective prompts comprises five metacognitive event classes. Furthermore, the most important event classes are MONITOR, READ, ORGANIZATION, and ELABORATION. Referring to the edges in the process model, the most important connections among event classes are between READ and MONITOR, between ELABORATE and MONITOR, between ORIENT and EVAL, as well as between ORGANIZATION and ELABORATE. Moreover, the classes MONITOR and ORIENT show several connections to other events, thus representing significant routing points for the process. Finally, there is a triple loop between READ, MONITOR, ELABORATE, and ORGANIZATION, which indicates a very active processing of information which is constantly monitored. In general, the integration of metacognitive event classes in the model of non-effective prompts can be interpreted as a successful regulation of learning activities with respect to SRL assumptions. According to SRL models (e.g., Winne & Hadwin, 2008; Zimmerman, 2008), students ideally perform various metacognitive activities and run through different regulation phases (forethought, performance, and evaluation).

In contrast, the process model of effective prompts only shows the three metacognitive categories, MONITOR, ORIENT, and SETGOAL, but all of the four cognitive event classes previously described. The most important event classes are MONITOR, READ, ELABORATE, and REPEAT. Furthermore, important edges are between MONITOR and ORGANIZATION, between ELABORATE and MONITOR, as well as between ORGANIZATION and READ. There is a cycle that connects READ, ELABORATE, MONITOR, and ORGANIZATION. This cycle can be interpreted as the processing of information that is monitored. However, compared to the model of non-effective prompts, there is a less active interconnection between these four event classes. Finally, the event classes REPEAT, SETGOAL, and ORIENT represent a second cycle, which shows a weak structure between these events. In general, the performance of cognitive activities, which is connected with MONITOR, represents the only clear structure in this process model. Furthermore, there is poor regulation of learning activities in comparison to the model of non-effective prompts.

In summary, the process model of non-effective prompts represents a highly regulated learning process, including various metacognitive categories, and a very active cycle of information processing. In contrast, the model of effective prompts shows a learning process that comprises few metacognitive event classes, and a cycle of monitored information processing. However, this cycle shows a weaker interconnectedness than the active cycle in the model of non-effective prompts. As a result, in the case of the model of non-effective prompts, an intervention through a metacognitive prompt may not be necessary, or it may even be disturbing considering the already high state of regulation. Moreover, interventions in a process that comprises very active information processing may not be the right context to foster metacognition because the learner has to interrupt his or her cognitive activities to interact with the prompt and follow its request. Consequently, this can lead to ignoring the prompts and to a continuation of one's information processing activities. On the other hand, the presentation of a metacognitive prompt can be more effective if the learning process comprises a weak regulatory behavior and a less intensive cycle of cognitive activities. With respect to the findings of the linear regression, the precursor states that are likely to make a prompt effective may be orientation and monitoring events that are, however, not well embedded in the course of learning, as displayed in the structure of the process model of the effective prompts. Moreover, an effective prompt may foster the learning process by supporting the interconnectedness between the two cycles, especially by stimulating the connections that lead back from the regulatory behavior to the information processing cycle. As is shown above, particularly induced monitoring activities seem to cause better learning.

This may be due to the fact that effective prompts helped to cycle back to monitoring and the cognitive loop.

7. DISCUSSION

In this contribution, we evaluated the effectiveness of metacognitive prompts on the micro level to investigate the conditions for effective scaffolding during hypermedia learning. Because the research on SRL and scaffolding hypermedia learning has noted the importance of tailoring support to the needs of the learner (e.g., Azevedo & Hadwin, 2005), and learning systems offer the possibility of implementing adaptive features (e.g., Brusilovsky, 2001; Molenaar & Roda, 2008), we attempted to explore the reasons why a provided metacognitive prompt is sometimes effective in stimulating regulatory activities and why sometimes it is not. In general, this type of research is important for the improvement of scaffold design because information on the conditions for effectiveness can inform the positioning of support in the course of learning (e.g., Thillmann et al., 2009). For our purpose, we analyzed fine-grained process data, which was measured by coded think-aloud protocols, using DM and PM. With this approach, we sought to classify the effectiveness of prompts, to investigate the correspondence of the classification with learning outcome, and to discover the conditions under which prompts induce regulatory activities (i.e., the proper temporal positioning).

Our findings indicate that it is possible to distinguish between effective and non-effective prompts by considering the increase of metacognitive utterances following the prompt presentation. Overall, approximately half of the prompts that were provided to the students were classified as effective according to our chosen classification criteria. Because we had to make a decision about how to determine the time intervals to investigate the increase of metacognitive utterances, it is possible that a different selection of measurement units would result in a different prompt classification. The determination of measurement units is a general challenge in analyzing fine-grained process data (Johnson et al., 2011; Winne, 2014). Therefore, we validated our decision by showing that the mean number of metacognitive events following the prompts is significantly higher than the baseline of metacognitive utterances. Moreover, we related the selection of time intervals to learning performance. Ideally, the students who showed a higher number of induced metacognitive activities in the two-minute time intervals following a prompt would also show a better learning performance. The results revealed that there is no significant correlation between the total number of induced metacognitive activities and learning performance. A closer look at the sub-categories of metacognition showed that the highest correlation was between monitoring and transfer performance. However, this small effect was not significant, possibly because of the small sample size. Nevertheless, the positive correspondence between monitoring and transfer performance is in line with the previous research (Sonnenberg & Bannert, 2015).

Based on the classification of effective and non-effective prompts, it was our next aim to compare these two types by considering the students' current learning process preceding a prompt. Depending on the learning activities that were performed by the students (e.g., reading, the processing of information, or monitoring and controlling their learning progress), it is possible that the presentation of a scaffold affects the learning process differently. For example, one reason for the non-effectiveness of a provided prompt in our study may be a suboptimal positioning within the student's learning process. By applying DM and PM techniques to the coded learning activities, we tried to discover indicators of the effectiveness of a prompt (i.e.,

the conditions under which it might be more likely to be effective and non-effective, respectively) to improve the positioning of metacognitive support in future scaffold design. First, we analyzed the frequency of learning activities in the time intervals that preceded each prompt by applying a linear regression learner. The findings indicate that a high occurrence of orientation and monitoring activities fosters the desired prompting effects (i.e., the activation of regulatory activities).

These findings were supported by applying a PM algorithm that not only takes into account the frequency of coded learning activities but also their relative positioning (i.e., their sequential order). We compared two process models, both inductively generated by the Fuzzy Miner algorithm, that represent the learning processes (more precisely, the workflow of learning activities) that precede the appearance of an effective and a non-effective prompt. The sequence of learning activities in the process model of non-effective prompts resembles the performance of successful regulation patterns as described in SRL models (e.g., Winne & Hadwin, 2008; Zimmerman, 2008). Additionally, this process model shows a very active sequence of monitored information processing. In contrast, the process model of effective prompts comprises the execution of cognitive activities, which are only connected with one metacognitive activity, namely monitoring. This could be interpreted as a poor regulation of the learning process.

In summary, the findings of the learned linear regression function and the process models indicate that the occurrence of orientation and monitoring activities, if they are not yet well embedded in the course of learning, increase the likelihood of effectiveness. In this case, metacognitive prompting may be successful in fostering further regulatory activities and in structuring the regulating of the learning process. If students already show a highly regulated process and an intense sequence of cognitive processing, the intervention of a metacognitive scaffold may not be necessary, and it may even be disruptive. These results relate to the implications for the design of adaptive systems based on attention management of Molenaar and Roda (2008). The authors note that one has to evaluate the cost of switching attentional focus and that learners should be able to predict interruption times.

Our results can be considered for future scaffold design, for example, for adapting the presentation of a metacognitive prompt to the current learning process. In general, our analysis shows how to use process data to evaluate the effectiveness of provided support on the micro level by applying DM and PM techniques. Furthermore, the findings strengthen the importance of using process data in the investigation of the effectiveness of instructional support, which in turn provides implications for improving its design (e.g., by providing conditions for adaptive scaffolds).

With regard to the limitations of our study, it should be noted that the results depend on our learning setting (e.g., learning material and learning environment) and on our participants. Therefore, our findings may be task-specific, and they probably do not represent general patterns. For different learning settings, metacognitive regulatory processes may look entirely different. Additionally, because the sample group of this study was predominately female, the results may be influenced by gender differences. The question of possible differences between male and female students could be addressed in future studies.

Moreover, our coding scheme, which comprises several metacognitive, cognitive, and motivational activities, determines the level of granularity for the present analysis. Even more importantly, the applied DM and PM techniques are inductive approaches, and thereby the findings rely on the representativeness and quality of the underlying data sources (e.g., Bannert et al., 2014). Therefore, the resulting patterns need to be validated in future research. Furthermore,

the process data were measured by concurrent think-aloud protocols. The results should also be replicated using process data on different levels (e.g., log files or eye tracking), ideally advancing towards a temporal alignment of different data channels (Azevedo, 2014), and towards an integration of findings across several SRL studies (Dent & Hoyle, 2015). However, approaches that use log files to track students' cognitive and metacognitive strategies in learning environments imply a more difficult interpretation of awareness and intent regarding students' actions and behaviors than coded think-aloud protocols (e.g., Kinnebrew et al., 2014). Finally, we did not consider the impact of previously presented prompts in detail, but there seems to be no additive effect in our data. Additivity means that it might be possible that the success of a scaffold depends on the effectiveness of previous scaffolds. For example, Molenaar and Roda (2008) recommend considering the history of a learner's interactions with previously provided scaffolds to calibrate the presentation of support. Moreover, without this type of history, the challenge of fading support when the learner's regulative behavior progresses cannot be addressed.

Future directions should address the following issues. In our study, we investigated the learning process that precedes a prompt by analyzing coded learning activities to discover the conditions for effectiveness. However, there may be further conditions that influence the effectiveness of metacognitive prompts that were not considered in the present analysis. For example, the students' level of expertise could be another important factor that affects scaffolding effects (e.g., Yeh et al., 2010). Moreover, questions with regard to the role of the presentation time (e.g., is a prompt that appears during the first minutes of the learning session more effective than a prompt that appears in the middle of the learning time?) and of the number of prompts (e.g., is the first prompt already effective or are two or more prompts needed to affect the regulatory behavior?) still have to be investigated in future research. Furthermore, the findings on possible conditions for the effectiveness of scaffolds should be related to models of SRL and metacognition. For example, the COPES model (Winne & Hadwin, 2008) could be specified by including positions for intervening the SRL cycle. These findings could contribute to a micro-level theory of SRL and to the impact of scaffolds on learning processes that is needed to implement optimally adaptive support in learning environments, and that is currently missing in the field of SRL (Molenaar & Järvelä, 2014). Additionally, more research on the automatic detection of learning processes (e.g., Cocea & Weibelzahl, 2009) is needed to improve the diagnosis of adaptive learning systems. The next step in our future research will be the validation of findings on the effectiveness of metacognitive prompts by analyzing think-aloud data from another prompting experiment. Additionally, we will aim to derive standards and guidelines for the application of DM and PM techniques in SRL settings.

ACKNOWLEDGEMENTS

This research was supported by funds from the German Science Foundation (DFG: BA 2044/7-1). We would like to thank our student research assistants, Ferdinand Guggenberger, Valentina Haspert, Michaela Hofbauer, and Lisa Müsse, for their assistance with regard to coding the think-aloud data.

Correspondence concerning this manuscript should be sent to Christoph Sonnenberg, University of Wuerzburg, Instructional Media, Oswald-Kuelpe-Weg 82, 97074 Wuerzburg, Germany, email: christoph.sonnenberg@uni-wuerzburg.de; phone: +49931/3189141

REFERENCES

- ALEVEN, V. (2013). Help seeking and intelligent tutoring systems: Theoretical perspectives and a step towards theoretical integration. In R. Azevedo & V. Alevén (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 311–335). New York: Springer. doi:10.1007/978-1-4419-5546-3_21
- ALEVEN, V., ROLL, I., MCLAREN, B. M., & KOEDINGER, K. R. (2016). Help helps, but only so much: Research on help seeking with intelligent tutoring systems. *International Journal of Artificial Intelligence in Education*, 26(1), 205–223. doi:10.1007/s40593-015-0089-1
- AZEVEDO, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The Role of self-regulated learning. *Educational Psychologist*, 40(4), 199–209. doi:10.1207/s15326985ep4004_2
- AZEVEDO, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning*, 4, 87–95. doi:10.1007/s11409-009-9035-7
- AZEVEDO, R. (2014). Issues in dealing with sequential and temporal characteristics of self- and socially-regulated learning. *Metacognition and Learning*, 9, 217–228. doi:10.1007/s11409-014-9123-1
- AZEVEDO, R., CROMLEY, J. G., MOOS, D. C., GREENE, J. A., & WINTERS, F. I. (2011). Adaptive content and process scaffolding: A key to facilitating students' self-regulated learning with hypermedia. *Psychological Testing and Assessment Modeling*, 53(1), 106–140.
- AZEVEDO, R., CROMLEY, J. G., & SEIBERT, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology*, 29, 344–370. doi:10.1016/j.cedpsych.2003.09.002
- AZEVEDO, R., CROMLEY, J. G., WINTERS, F. I., MOOS, D. C., & GREENE, J. A. (2005). Adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instructional Science*, 33, 381–412. doi:10.1007/s11251-005-1273-8
- AZEVEDO, R., & HADWIN, A. F. (2005). Scaffolding self-regulated learning and metacognition - Implications for the design of computer-based scaffolds. *Instructional Science*, 33, 367–379. doi:10.1007/s11251-005-1272-9
- AZEVEDO, R., LANDIS, R. S., FEYZI-BEHNAGH, R., DUFFY, M., TREVORS, G., HARLEY, J. M., . . . HOSSAIN, G. (2012). The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with MetaTutor. *Lecture Notes in Computer Science*, 212–221. doi:10.1007/978-3-642-30950-2_27
- BANNERT, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research*, 35(4), 359–375. doi:10.2190/94V6-R58H-3367-G388
- BANNERT, M. (2007). *Metakognition beim Lernen mit Hypermedia. Erfassung, Beschreibung und Vermittlung wirksamer metakognitiver Lernstrategien und Regulationsaktivitäten [Metacognition and hypermedia learning]*. Münster: Waxmann.
- BANNERT, M. (2009). Promoting self-regulated learning through prompts: A discussion. *Zeitschrift für Pädagogische Psychologie*, 23(2), 139–145. doi:10.1024/1010-0652.23.2.139
- BANNERT, M., & MENGELKAMP, C. (2013). Scaffolding hypermedia learning through metacognitive prompts. In R. Azevedo & V. Alevén (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 171–186). New York: Springer. doi:10.1007/978-1-4419-5546-3_12
- BANNERT, M., & REIMANN, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science*, 40(1), 193–211. doi:10.1007/s11251-011-9167-4

- BANNERT, M., REIMANN, P., & SONNENBERG, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning*, 9(2), 161–185. doi:10.1007/s11409-013-9107-6
- BANNERT, M., SONNENBERG, C., MENGELKAMP, C., & PIEGER, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior*, 52, 293–306. doi:10.1016/j.chb.2015.05.038
- BEN-ELIYAHU, A., & BERNACKI, M. L. (2015). Addressing complexities in self-regulated learning: A focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, 10, 1–13. doi:10.1007/s11409-015-9134-6
- BENJAMINI, Y., & HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289-300.
- BLOOM, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York: D. McKay.
- BOUCHET, F., HARLEY, J. M., TREVORS, G. J., & AZEVEDO, R. (2013). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *Journal of Educational Data Mining*, 5(1), 104–146.
- BRUSILOVSKY, P. (2001). Adaptive hypermedia. *User Modeling and User-Adapted Interaction*, 11, 87–110. doi:10.1023/A:1011143116306
- BRUSILOVSKY, P. (2007). Adaptive navigation support. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The adaptive web. Methods and strategies of web personalization* (pp. 263–290). Berlin Heidelberg: Springer. doi:10.1007/978-3-540-72079-9
- CHI, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, 6(3), 271–315. doi:10.1207/s15327809jls0603_1
- CLAREBOUT, G., & ELEN, J. (2006). Tool use in computer-based learning environments: Towards a research framework. *Computers in Human Behavior*, 22, 389–411. doi:10.1016/j.chb.2004.09.007
- CLAREBOUT, G., ELEN, J., COLLAZO, N. A. J., LUST, G., & JIANG, L. (2013). Metacognition and the use of tools. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 187–195). Springer Science. doi: 10.1007/978-1-4419-5546-3_13
- COCEA, M., & WEIBELZAHN, S. (2009). *Log file analysis for disengagement detection in e-Learning environments*. *User Modelling and User-Adapted Interaction*, 19, 341–385. doi:10.1007/s11257-009-9065-5
- DENT, A. L., & HOYLE, R. H. (2015). A framework for evaluating and enhancing alignment in self-regulated learning research. *Metacognition and Learning*, 10, 165–179. doi:10.1007/s11409-015-9136-4.
- ERICSSON, K. A., & SIMON, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, Mass.: MIT Press.
- GE, X. (2013). Designing learning technologies to support self-regulation during III-structured problem-solving processes. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 213–228). New York: Springer. doi:10.1007/978-1-4419-5546-3_15
- GREENE, J. A., DELLINGER, K. R., TÜYSÜZOĞLU, B. B., & COSTA, L.-J. (2013). A two-tiered approach to analyzing self-regulated learning data to inform the design of hypermedia learning environments. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 117–128). New York: Springer. doi:10.1007/978-1-4419-5546-3_8

- GÜNTHER, C. W., & VAN DER AALST, W. M. P. (2007). Fuzzy mining – Adaptive process simplification based on multi-perspective metrics. *Business Process Management - Lecture Notes in Computer Science*, 4714, 328–343. doi:10.1007/978-3-540-75183-0
- HÄMÄLÄINEN, W., & VINNI, M. (2010). Classifiers for educational data mining. In C. Romero, S. Ventura, M. Pechenizkiy, & R. Baker (Eds.), *Handbook of Educational Data Mining* (pp. 57–74). Boca Raton, FL: Chapman&Hall/CRC.
- HATTIE, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- JEONG, H., GUPTA, A., ROSCOE, R., WAGSTER, J., BISWAS, G., & SCHWARTZ, D. (2008). Using hidden markov models to characterize student behaviors in learning-by-teaching environments. *Lecture Notes in Computer Science*, 5091, 614–625. doi:10.1007/978-3-540-69132-7_64
- JOHNSON, A. M., AZEVEDO, R., & D’MELLO, S. K. (2011). The temporal and dynamic nature of self-regulatory processes during independent and externally assisted hypermedia learning. *Cognition and Instruction*, 29(4), 471–504. doi:10.1080/07370008.2011.610244
- KINNEBREW, J. S., SEGEDY, J. R., & BISWAS, G. (2014). Analyzing the temporal evolution of students’ behaviors in open-ended learning environments. *Metacognition and Learning*, 9, 187–215. doi:10.1007/s11409-014-9112-4
- KRAMARSKI, B., & GUTMAN, M. (2006). How can self-regulated learning be supported in mathematical e-learning environments? *Journal of Computer Assisted Learning*, 22, 24–33. doi:10.1111/j.1365-2729.2006.00157.x
- LAJOIE, S. P., NAISMITH, L., POITRAS, E., HONG, Y.-J., CRUZ-PANESSO, I., RANELLUCCI, J., . . . WISEMAN, J. (2013). Technology-rich tools to support self-regulated learning and performance in medicine. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 229–242). New York: Springer. doi:10.1007/978-1-4419-5546-3_16
- LEHMANN, T., HÄHNLEIN, I., & IFENTHALER, D. (2014). Cognitive, metacognitive and motivational perspectives on prelection in self-regulated online learning. *Computers in Human Behavior*, 32, 313–323. doi:10.1016/j.chb.2013.07.051
- LESTER, J. C., MOTT, B. W., ROBISON, J. L., ROWE, J. P., & SHORES, L. R. (2013). Supporting self-regulated science learning in narrative-centered learning environments. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 471–483). New York: Springer. doi:10.1007/978-1-4419-5546-3_30
- LIN, X., HMELO, C., KINZER, C. K., & SECULES, T. J. (1999). Designing technology to support reflection. *Educational Technology Research and Development*, 47(3), 43–62. doi:10.1007/BF02299633
- MALMBERG, J., JÄRVELÄ, S., JÄRVENOJA, H., & PANADERO, E. (2015). Promoting socially shared regulation of learning in CSCL: Progress of socially shared regulation among high- and low-performing groups. *Computers in Human Behavior*, 52, 562–572. doi:10.1016/j.chb.2015.03.082
- MARTIN, T., & SHERIN, B. (2013). Learning analytics and computational techniques for detecting and evaluating patterns in learning: An introduction to the special issue. *Journal of the Learning Sciences*, 22(4), 511–520. doi:10.1080/10508406.2013.840466
- MOLENAAR, I., & CHIU, M. M. (2014). Dissecting sequences of regulation and cognition: Statistical discourse analysis of primary school children’s collaborative learning. *Metacognition and Learning*, 9(2), 137–160. doi:10.1007/s11409-013-9105-8
- MOLENAAR, I., & JÄRVELÄ, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning*, 9(2), 75–85. doi:10.1007/s11409-014-9114-2

- MOLENAAR, I., & RODA, C. (2008). Attention management for dynamic and adaptive scaffolding. *Pragmatics & Cognition*, 16(2), 224-271. doi: 10.1075/pc.16.2.04mol
- REIMANN, P., MARKAUSKAITE, L., & BANNERT, M. (2014). E-Research and learning theory: What do sequence and process mining methods contribute? *British Journal of Educational Technology*, 45(3), 528-540. doi:10.1111/bjet.12146
- REIMANN, P., & YACEF, K. (2013). Using process mining for understanding learning. In R. Luckin, P. Puntambekar, P. Goodyear, B. Grabowski, J. Underwood, & N. Winters (Eds.), *Handbook of Design in Educational Technology* (pp. 472-481). New York: Routledge.
- ROLL, I., ALEVEN, V., MCLAREN, B. M., & KOEDINGER, K. R. (2011). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction*, 21(2), 267-280. doi:10.1016/j.learninstruc.2010.07.004
- SCHOOR, C., & BANNERT, M. (2012). Exploring regulatory processes during a computer-supported collaborative learning task using process mining. *Computers in Human Behavior*, 28, 1321-1331. doi:10.1016/j.chb.2012.02.016
- SCHWONKE, R., HAUSER, S., NÜCKLES, M., & RENKL, A. (2006). Enhancing computer-supported writing of learning protocols by adaptive prompts. *Computers in Human Behavior*, 22, 77-92. doi:10.1016/j.chb.2005.01.002
- SITZMANN, T., BELL, B. S., KRAIGER, K., & KANAR, A. M. (2009). A multilevel analysis of the effect of prompting self-regulation in technology-delivered instruction. *Personnel Psychology*, 62, 697-734. doi:10.1111/j.1744-6570.2009.01155.x
- SONG, M., & VAN DER AALST, W.M.P. (2007). Supporting process mining by showing events at a glance. In K. Chari & A. Kumar (Eds.), *Proceedings of 17th Annual Workshop on Information Technologies and Systems (WITS 2007)* (pp. 139-145). Montreal, Canada.
- SONNENBERG, C., & BANNERT, M. (2015). Discovering the effects of metacognitive prompts on the sequential structure of SRL-processes using process mining techniques. *Journal of Learning Analytics*, 2(1), 72-100.
- THILLMANN, H., KÜNSTING, J., WIRTH, J., & LEUTNER, D. (2009). Is it merely a question of "what" to prompt or also "when" to prompt? *Zeitschrift für Pädagogische*, 23(2), 105-115. doi:10.1024/1010-0652.23.2.105
- TRČKA, N., PECHENIZKIY, M., & VAN DER AALST, W. M. P. (2010). Process mining from educational data. In C. Romero, S. Ventura, M. Pechenizkiy, & R. Baker (Eds.), *Handbook of Educational Data Mining* (pp. 123-142). Boca Raton, FL: Chapman&Hall/CRC.
- VAN DER AALST, W. (2011). *Process mining: Discovery, conformance and enhancement of business processes*. Berlin, Heidelberg: Springer.
- VAN DER AALST, W., WEIJTERS, T., & MARUSTER, L. (2004). Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, 16(9), 1128-1142. doi:10.1109/TKDE.2004.47
- VEENMAN, M. (1993). *Metacognitive ability and metacognitive skill: Determinants of discovery learning in computerized learning environments*. University of Amsterdam.
- 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
- WALKER, E., RUMMEL, N., & KOEDINGER, K. (2011). Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity. *International Journal of Computer Supported Collaborative Learning*, 6(2), 279-306. doi: 10.1007/s11412-011-9111-2

- WINNE, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences*, 8(4), 327–353.
- WINNE, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning*, 9, 229–237. doi:10.1007/s11409-014-9113-3
- WINNE, P. H., & BAKER, R. S. (2013). The potentials of educational data mining for researching metacognition, motivation and self-regulated learning. *Journal of Educational Data Mining*, 5(1), 1–8.
- WINNE, P. H., & HADWIN, A. F. (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). NY: Taylor & Francis.
- WIRTH, J. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie*, 23(2), 91–94. doi:10.1024/1010-0652.23.2.91
- WITTWER, J., & RENKL, A. (2008). Why instructional explanations often do not work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist*, 43, 49–64. doi:10.1080/00461520701756420
- YEH, Y. F., CHEN, M. C., HUNG, P. H., & HWANG, G. J. (2010). Optimal self-explanation prompt design in dynamic multi-representational learning environments. *Computers and Education*, 54, 1089–1100. doi:10.1016/j.compedu.2009.10.013
- ZIMMERMAN, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183. doi:10.3102/0002831207312909

Analysis 3:

Using Process Mining to Examine the Sustainability of Instructional Support: How Stable are the Effects of Metacognitive Prompting on Self-Regulatory Behavior?

Submitted to *Computers in Human Behavior*:

Sonnenberg, C., & Bannert, M. (submitted). Using process mining to examine the sustainability of instructional support: How stable are the effects of metacognitive prompting on self-regulatory behavior? *Computers in Human Behavior*.

Abstract

The current study investigates the sustainability of metacognitive prompting on self-regulatory behavior using a Process Mining approach. Previous studies confirmed beneficial short-term effects of metacognitive prompts on the learning process and on learning outcomes. However, the question of how stable these effects are for similar tasks in the future so far remains unanswered. Also, the use of online trace methods and the emergence of new analytical approaches allow deeper insights into the sequential structure of learning activities. Therefore, we examined long-term effects of instructional support on the micro level using Process Mining. Data gathered through the think-aloud method from 69 university students was measured during two learning sessions. Metacognitive prompts supported the experimental group ($n = 35$) only during the first session. Based on a process model generated by using the data of the first learning session, we analyzed the sustainability of effects during the second learning session. Results showed significant differences between the experimental and control group regarding the frequency of metacognitive learning activities, which remain stable over time. Additionally, the application of Process Mining indicated which sequences of learning activities were transferred to the second learning session. Our findings demonstrate the benefits of evaluating instructional support using analysis techniques that take into account the sequential structure of learning activities. Moreover, while the results provide initial evidence for sustainable long-term effects on self-regulatory behavior, they have to be replicated in future research.

Keywords: metacognitive prompting, long-term effects, think-aloud protocols, process analysis, process mining, conformance checking

The research in self-regulated learning (SRL) shows that students have difficulties in spontaneously performing regulatory activities, especially in technology-enhanced and open-ended learning environments (e.g., Azevedo, 2009; Bannert & Mengelkamp, 2013; Greene, Dellinger, Tüysüzoglu, & Costa, 2013). To counter this production deficit, instructional support aims to activate strategic learning processes. Therefore, our objective is to investigate the effects of metacognitive prompts (Bannert, 2009) provided during hypermedia learning, in order to design helpful support. Previous studies confirmed beneficial short-term effects on the learning process and on learning outcome (e.g., Azevedo, Cromley, Moos, Greene, & Winters, 2011; Bannert & Mengelkamp, 2013; Lehmann, Hähnlein, & Ifenthaler, 2014). However, few studies have addressed the stability and transfer of strategies, which have been fostered previously, during follow-up learning tasks (e.g., Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015; Nückles, Hübner, Dümer, & Renkl, 2010; Roll, Alevén, McLaren, & Koedinger, 2011). Considering the high relevance of sustainable effects in practice, research needs to examine the long-term impact of metacognitive prompting as well as further scaffolding techniques.

Moreover, current research that considers SRL a dynamic interplay of various learning events highlights the importance of assessing and analyzing fine-grained traces of learners' behavior to understand the impact of instructional support (e.g., Azevedo, 2014; Azevedo et al., 2013; Sonnenberg & Bannert, 2015, in press). The emergence of innovative analytical approaches, such as techniques that take into account the sequential and temporal structure of learning activities (Molenaar & Järvelä, 2014), has accompanied the growing interest in analyzing process data. In previous analyses, we applied *Process Mining* techniques (PM; Bannert, Reimann, & Sonnenberg, 2014; Trčka, Pechenizkiy, & van der Aalst, 2010) to investigate the effect of metacognitive prompts on micro-level processes. PM has the potential to contribute not only to the discovery of dynamic regulatory behavior, but also to the implementation of confirmatory analysis that would validate present models of learning.

The current study contributes to the existing literature by examining long-term effects of metacognitive prompting on SRL in a follow-up learning task without instructional support. The students' activities were measured using concurrent think-aloud protocols. In addition to a frequency analysis of coded learning events, we applied a PM approach, namely *Conformance Checking* (Rozinat & van der Aalst, 2008), to compare the sequential structure of the learning

process between two learning tasks. Through this approach, we attempted to determine if the fostered self-regulatory behavior is stable, and which activities are transferred to a future task.

The paper is structured as follows. First, we present research on metacognitive support, especially focusing on findings regarding the short- and long-term effects of metacognitive prompts. Second, we introduce the assessment of learning activities using PM. Third, a hypermedia learning experiment, comprising a prompted experimental group and a control group, is described. Fourth, we report the findings of a frequency analysis, as well as of PM which compared the students' activities between two learning sessions. Finally, we discuss the significance of our results for the sustainability of metacognitive prompting, as well as the benefits of analyzing learning activities on a micro level using PM.

Metacognitive Support for Scaffolding Technology-Enhanced Learning

According to SRL models (Winne & Hadwin, 2008; Zimmerman, 2008), successful learning corresponds with the active deployment of cognitive and metacognitive learning activities, such as goal setting, activation of prior knowledge, and monitoring of one's learning progress. Moreover, these models describe SRL as a dynamic interplay of various learning activities to achieve one's learning goals. To counter students' deficiency in actively regulating their learning, instructional support is necessary, for example, to stimulate the use of metacognitive learning activities when studying with a non-linear hypermedia system. Because metacognition represents a key role in SRL models, especially monitoring and controlling one's learning, interventions that focus on metacognitive support have the potential to foster students' successful learning (e.g., Bannert & Reimann, 2012; Künsting, Kempf, & Wirth, 2013). A robust body of research has confirmed the beneficial effects of metacognitive support on learning in computer-based learning environments (CBLE; Devolder, van Braak, & Tondeur, 2012; Zheng, 2016).

Moreover, SRL models also comprise the transfer of learning experiences to further tasks in the future, at least if self-evaluation or external feedback has taken place. For example, the COPES model (Winne & Hadwin, 2008) assumes that evaluations influence task and cognitive conditions in future learning tasks (e.g., knowledge of task, study tactics, and strategies). Consequently, successful learning strategies should be transferred to further tasks, especially if the requirements are similar. Furthermore, that means metacognitive support during

a learning phase possibly has beneficial effects not only on the current task but also on similar subsequent tasks without support.

In the following, we report previous findings of one type of instructional support, namely metacognitive prompting, as a scaffolding technique used with hypermedia learning. Also, we discuss related work on the sustainability and transfer of SRL activities to subsequent learning tasks.

Supporting Hypermedia Learning Through Metacognitive Prompts

Metacognitive prompts attempt to activate the students' repertoire of regulation strategies in educational settings. Usually, stimulating questions or advice on execution is used to scaffold the deployment of desired behavior (Bannert, 2009). For instance, metacognitive prompts can ask students to reflect on their navigational decisions while learning with hypermedia (Bannert, 2006; Stark & Krause, 2009).

According to their key mechanism, metacognitive prompts support students' active regulation of learning by stimulating their strategic knowledge (Bannert & Mengelkamp, 2013). The presentation of prompted requests aims to counter the so-called *production deficit* (e.g., Winne, 1996; Wirth, 2009). Consequently, the students' regulation of learning should shift to a better use of metacognitive activities as described in SRL models (Winne & Hadwin, 2008; Zimmerman, 2008).

With respect to the additional cognitive load caused by prompted requests (Berthold, Röder, Knörzer, Kessler, & Renkl, 2011) and the threat of expertise-reversal effects (Nückles et al., 2010), it is desirable to reduce and eventually omit instructional support. This also corresponds with the original concept of scaffolding (Puntambekar & Hübscher, 2005; Wood, Bruner, & Ross, 1976), in which the support of a teacher or educational technology is withdrawn when a student advances to the scaffolded behavior. Ideally, the student will spontaneously use the fostered behavior in the following, even in further contexts. To know when to reduce support, research on the stability of metacognitive prompting effects is needed.

Short- and Long-Term Effects of Metacognitive Prompting

A robust body of research indicates the effective short-term impact of metacognitive prompts on learning processes and outcomes in different domains and educational settings. The investigation of short-term effects refers to the immediate prompting effects for a supported

learning task. In general, the findings show beneficial effects for hypermedia learning (e.g., Azevedo et al., 2011; Bannert & Mengelkamp, 2013), writing learning journals (e.g., Hübner, Nückles, & Renkl, 2010; Nückles et al., 2010), and additional settings (e.g., Künsting et al., 2013; Thillmann, Künsting, Wirth, & Leutner, 2009). Prompting results in an improvement of self-regulatory behavior, often measured using online trace methods such as think-aloud protocols or computer log files (e.g., Greene & Azevedo, 2010). Moreover, several studies—but not all—also report a beneficial impact on learning outcomes, usually measured using performance tests immediately after the learning phase (e.g., Azevedo et al., 2011; Stark & Krause, 2009).

A previous series of experiments (Bannert & Mengelkamp, 2013) that investigated the impact of different types of metacognitive prompts indicates beneficial effects on hypermedia learning. Analyses of concurrent think-aloud protocols demonstrated a quantitative increase of metacognitive learning activities, and also showed that prompting positively affected the sequential structure of regulatory behavior (Sonnenberg & Bannert, 2015). Moreover, prompting also had a beneficial effect on learning outcome, but mainly on transfer performance (i.e., application of knowledge). Overall, related work and our studies support the expected beneficial short-term effect of metacognitive prompting.

Although theoretical assumptions in SRL models (e.g., Winne & Hadwin, 2008; Zimmerman, 2008) explicitly comprise the transfer of learning experiences to future tasks, relatively few studies have examined the long-term effects of metacognitive prompts. Stark and Krause (2009) investigated the sustainability of reflection prompts in a CBLE for statistics education. In a learning test repeated four weeks after the intervention, they found stable effects of prompting on domain knowledge, whereas the test scores of all groups decreased in general. Moreover, Roll et al. (2011) examined the sustainable use of help-seeking behavior in subsequent learning tasks without support. Students learning with the Geometry Cognitive Tutor and initially supported by metacognitive feedback showed a transfer of fostered help-seeking skills to a new task, but showed no improvement in learning outcome. Similarly, the findings of Roll, Yeh, and Briseno (2014) indicate that a scaffolded learning activity can be transferred to a subsequent task within the same learning environment. In contrast to these positive findings, Hilbert et al. (2008) found no beneficial impact of cognitive and metacognitive prompts in a follow-up session one week after the intervention. The authors investigated the short- and long-

term effects of prompts on writing a concept map. They argued that a one-time intervention might not be sufficient to internalize the prompted strategies; therefore, students might not spontaneously use these strategies in a following task without support. Finally, Nückles et al. (2010) provided cognitive and metacognitive prompts to support students' writing of learning journals during a semester. Their results show that constant prompting can lead to an expertise-reversal effect, and that fading out of support is necessary when students have internalized prompted strategies. Furthermore, the use of metacognitive strategies decreased during the semester, even when supported by prompts. In summary, there is initial evidence of beneficial long-term effects of metacognitive prompting; however, research needs to extend the existing small body of literature.

In a recent study (Bannert et al., 2015), we investigated the short- and long-term effects of self-directed metacognitive prompts during hypermedia learning. Students were asked to reflect on their navigational decisions during a learning task; then they learned without support during a second learning session three weeks later. When compared to a control group with no support at all, we found beneficial effects on systematic navigation behavior and transfer performance in both learning sessions. These findings support the assumption that prompted strategies are also transferred to subsequent learning tasks. The current study extends the previous contribution through an in-depth analysis of concurrent think-aloud protocols.

To advance the understanding of sustainability and transfer of SRL processes, researchers agree that more microanalyses using fine-grained process data are necessary (Hilbert et al., 2008; Moos & Miller, 2015; Schunk & Ertmer, 2005; Severiens, Ten Dam, & Van Hout-Wolters, 2001). Online trace methods for SRL assessment (e.g., think-aloud protocols, or log files) combined with the application of techniques from Educational Data Mining (EDM; Winne & Baker, 2013) have the potential to contribute to these issues. For example, EDM allows precise modeling of robust learning and the impact of scaffolds in computer-supported settings (Baker & Corbett, 2014; Sonnenberg & Bannert, 2015). Therefore, models describing the learning activities of a specific learning session or learner sample can be tested and validated with data from future sessions or students. One promising approach used in EDM to analyze the learning process on the micro level is presented as follows.

Using Process Mining to Analyze Students' Learning Activities

Research in metacognition and SRL has started to emphasize the increasing significance of fine-grained process data for analyzing technology-enhanced learning activities. Recent studies use rich data from various online trace methods, for example from verbal reports, eye-tracking, physiological measurement, or computer log files (Azevedo et al., 2013; Trevors, Feyzi-Behnagh, Azevedo, & Bouchet, 2016) to address the complex interplay of SRL processes, including their sequential and temporal dynamics.

Measuring and Analyzing SRL as Patterns of Events

The researchers' perspective on SRL has shifted from SRL being an aptitude to being a dynamic interplay of events during learning (Reimann, 2009; Winne & Perry, 2005). Following the event-based approach, it is possible to analyze the sequential structure of learning activities. For example, a student starts with reading the learning task, then tries to obtain an overview of the learning material, and next, reads the goals again to select the relevant sections. Such analyses require the online measurement of learning events using a granularity that is appropriate for the research question. One measurement approach often used in SRL settings is concurrent think-aloud protocols (Ericsson & Simon, 1993). Although the think-aloud technique is not unobtrusive for the learner, as some other online trace methods are (e.g., log files or eye-tracking), it provides a valuable access to the SRL events performed during learning (Azevedo, Moos, Johnson, & Chauncey, 2010). Overall, coded think-aloud data comprises a detailed trace of learning that is appropriate for examining the dynamics of SRL as well as the impact of instructional support on these processes.

While examining fine-grained trace data for SRL issues is potentially very valuable, researchers will have to face new analytical challenges. For example, advanced analysis approaches might be necessary to address the dynamic sequence of SRL activities as well as their relationship to learning performance. Recent advances in the field of metacognition, learning analytics, and EDM present contributions that rise to these challenges (Ben-Eliyahu & Bernacki, 2015; Molenaar & Järvelä, 2014; Roll & Winne, 2015; Winne & Baker, 2013). However, current analyses are exploratory, usually using process data to assess and model the learning activities in a specific setting or for a specific sample of students. Future research needs to shift toward the validation of resulting process patterns (Winne, 2014). Moreover, analysis

techniques that allow model testing can be applied to investigate the stability of learning activities. Therefore, we made use of an EDM technique, namely PM, which fulfills these requirements. Below, we present the approach of PM in more detail.

PM Techniques for Evaluating Instructional Support on the Micro Level

EDM techniques applied on SRL process data allow assessment and modeling sequences of learning activities (Baker & Corbett, 2014; Jeong et al., 2008; Roll & Winne, 2015). For instance, Baker and Corbett (2014) used Bayesian Knowledge Tracing to determine the probability of a student mastering a skill in a lesson. Moreover, model predictions can be used to evaluate whether a student deploys the skills in other contexts. As a further example, Jeong et al. (2008) applied hidden Markov models to investigate the effect of metacognitive prompting in a learning-by-teaching environment. PM has also shown potential as an approach for evaluating instructional support on the micro level (Sonnenberg & Bannert, 2015, 2016). Compared to hidden Markov models, PM also analyses the sequence of learning activities as a whole, but several functions are applicable in a unified framework (as described in the next paragraph), and a variety of algorithms are designed to meet specific requirements (e.g., dealing with noise in the data).

Within the field of EDM, PM represents an approach that uses event data to model the underlying process (Bannert et al., 2014; Reimann, Markauskaite, & Bannert, 2014, Trčka et al., 2010). In general, PM models build on the concept of Petri nets (Van der Aalst, 2011). These models comprise states (i.e., event classes or activities) and transitions between states. Moreover, the Petri net notation can be used to model achievable behavior and its workflow; that is, a model representing a learning process predicts the potential activities of a student or a sample of students.

Furthermore, PM comprises three main functions: (i) discovery of a process model using an event log, (ii) testing of conformance between a model and new data, and (iii) the extension of a present model. For these functions, several algorithms are implemented within a unified PM framework (ProM Version 5.2, 2008). PM considers the process as a whole by taking into account all recorded events to generate an underlying model. In other words, PM refers to end-to-end processes, and not only to certain reoccurring sequences as, for example, in Sequential Pattern Mining (Zhou, Xu, Nesbit, & Winne, 2010).

PM is increasingly used to model and understand learning activities as well as the impact of instructional support on these activities (Reimann & Yacef, 2013). One advantage of this approach is that it allows analysis of the relative arrangement of activities, and not only the frequencies of certain events. For example, it is possible to evaluate the impact of metacognitive prompting on the sequential structure of a learning process (Sonnenberg & Bannert, 2015). Since the quality of a learning process also predicts learning outcome, and not only the quantity of regulatory activities (e.g., Moos & Azevedo, 2009), it is also crucial to assess the sequence of events. Moreover, these analyses can also advance SRL theories on the micro level (Molenaar & Järvelä, 2014). Several studies in the field of SRL (Bannert et al., 2014; Sonnenberg & Bannert, 2015, 2016), Computer-Supported Collaborative Learning (Malmberg, Järvelä, Järvenoja, & Panadero, 2015; Reimann, Frerejean, & Thompson, 2009; Schoor & Bannert, 2012), and workplace learning (Siadaty, Gašević, & Hatala, 2016a, 2016b) have indicated the added value of applying PM techniques as well as related microanalytic approaches in educational settings.

As described above, PM explicitly comprises the function “testing conformance between a process model and an event log”. Compared to other analysis techniques, implementing Conformance Checking (Rozinat & van der Aalst, 2008) allows researchers to validate process models using future data in a relatively simple way. For instance, as a first step, a PM algorithm discovers the sequential structure of the learners’ traces and generates a process model. Then, as a second step, the conformance between this model and another set of learner traces can be measured. Therefore, Conformance Checking allows comparing the learning process between two or more learning tasks. In conclusion, PM supports the examination of sustainable long-term effects of instructional support on the micro level among several learning sessions.

Research Questions

A robust body of research has demonstrated the beneficial short-term effects of metacognitive prompting on learning processes and outcomes (e.g., Azevedo et al., 2011; Bannert & Mengelkamp, 2013; Hübner et al., 2010). However, only a few studies have investigated the sustainability of previously fostered strategies during learning in follow-up tasks (e.g., Roll et al., 2011). Especially from a practical perspective, it is crucial to know if

pedagogical interventions foster robust learning and the transfer of competencies to new situations (Baker & Corbett, 2014). Moreover, the analysis of fine-grained trace data (e.g., think-aloud protocols) contributes to a deeper understanding of learning activities as well as the impact of instructional support on these activities. Using process data allows researchers to analyze learning on the micro level, and to test assumptions regarding the temporal dynamics of SRL. Furthermore, analysis techniques, such as PM, that take into account the sequential structure of learning activities support a detailed examination of SRL among several tasks, and provide more value than a traditional frequency analysis. Therefore, the current study uses coded think-aloud data from two learning sessions and Conformance Checking to investigate the long-term impact of metacognitive prompts during hypermedia learning. In detail, we address the following research questions:

1. Do metacognitive prompts affect the learning process in a follow-up task by showing sustainable long-term effects on self-regulatory behavior?
2. Is it possible to identify the sequential patterns that are transferred to a follow-up learning task using Conformance Checking?

Referring to the key mechanism of metacognitive prompting, we expected the following results. Providing metacognitive prompts during learning should activate the students' repertoire of regulatory strategies. The activation should lead to an increase of metacognitive activities as well as a better sequence of activities according to SRL models (Winne & Hadwin, 2008; Zimmerman, 2008). In addition to beneficial short-term effects, we also expected a sustainable impact on the self-regulatory behavior in a future similar task without prompting. Again, this meant that students should show a better use of metacognitive activities. Finally, we need to mention that the learning process might also change between two tasks because of learning within the same environment and a task adaptation (Pieschl, Stahl, Murray, & Bromme, 2012). However, there are no previous findings for our learning material that would have helped us to predict expectations regarding an adaptation.

Method

The present study refers to a hypermedia learning experiment already reported in Bannert et al. (2015). However, the previous contribution does not include any analyses using think-aloud data, and it addresses different research questions. To avoid redundancy, we report

only the most important information for understanding the following analyses as well as the interpretation of results. Furthermore, we consider the reporting standards recommended by Dent and Hoyle (2015) to facilitate the evaluation and alignment of SRL research.

Sample and Research Design

A total of 69 undergraduate university students majoring in Media Communication or Human-Computer Systems (mean age = 20.00, $SD = 1.79$, 84.1% female) participated in a hypermedia learning experiment. The experiment comprised two 40-minute learning sessions. In both sessions, we used concurrent think-aloud protocols for the measurement of students' learning activities, and we obtained learning outcomes directly after learning. During the first learning session, metacognitive prompts supported students randomly assigned to the experimental group (EG; $n = 35$). Students in the control group (CG; $n = 34$) learned without instructional support. Approximately three weeks later, all students participated in a second learning session without support, but they had a similar learning task.

Learning Material

For both sessions, the learning material was provided in the same hypermedia learning environment. Navigation within this environment was possible through using a navigation menu, a next page and previous page button, the browser buttons (forward and backward), or hyperlinks. The topic of the first learning session was a chapter on learning theories (e.g., classical and operant conditioning), and the topic of the second learning session was a chapter on motivational psychology (e.g., Maslow's pyramid and achievement motivation). Both chapters represented educational psychology content, and they comprised a comparable amount of pages. One chapter included approximately 50 pages with 13,000 words, 20 pictures and tables, and 300 hyperlinks. We analyzed the text readability of both chapters using the koRpus R package (Michalke, 2015) to compare their level of difficulty. The Flesch-Kincaid Reading Ease score (Amstad for German texts) of the first chapter, learning theories, was 20.03. The reading ease score for the second chapter, motivational psychology, was 18.98. These scores indicated that both chapters had a similar difficulty level and were appropriate for university students (i.e., a score between 0 and 30).

Metacognitive support for students assigned to the EG was implemented within the learning environment through the use of a pop-up window that appeared several times during

the first learning session. The pop-up window comprised a list of strategic reasons for a selection of a particular node, for example, orientation or evaluation of goal attainment.

Procedure

We conducted both learning sessions as individual lab sessions. The first learning session started with an introduction to navigation within the learning environment and the think-aloud technique. The student was asked to verbalize every thought that came to his or her mind, without any justification or interpretation. As specified by Ericsson and Simon (1993), these instructions referred to level 2 verbalizations. A researcher provided exercises until a participant firmly mastered both navigation and concurrent thinking aloud. Next, participants randomly assigned to the EG received an introduction into the use of metacognitive prompts during learning, and they practiced the handling of prompts. To keep the workload equivalent for both groups, students in the CG received an introduction to workplace design. Then, all participants were given a sheet of paper comprising their learning task (i.e., learning the basic concepts and principles of operant conditioning), and they worked through the learning material for 40 minutes. Metacognitive prompts supported students in the EG eight times during learning. After a navigational step (e.g., clicking on a hyperlink), a pop-up window appeared in the middle of the screen, and students were asked to give reasons for their node selection by choosing from a list of strategic reasons. Except for these prompted requests, all participants were completely free in the execution of their learning activities. Additionally, they were allowed to take notes on a blank sheet of paper, but they were not authorized to use their notes during the learning test. During learning, a researcher stayed nearby and reminded the participant to think aloud if he or she remained silent for more than five seconds. All verbal statements were recorded using a microphone. Directly after learning, the participants worked on a learning test comprising different levels of knowledge (i.e., recall, comprehension, and transfer performance).

Approximately three weeks later, the second learning session was conducted. First, a researcher reminded the participants how to navigate within the learning environment and think-aloud during learning. Then, the learning task was given, and the students had to work through the chapter “motivational psychology” for 40 minutes. This time, no participant received any support. As in the first learning session, note-taking was allowed, and all participants were instructed to think aloud while learning. Again, the learning phase was directly followed by a

test comprising three levels of knowledge. More information about the learning tests is provided in Bannert et al. (2015).

Coding Scheme

For both learning sessions, the think-aloud data of all participants were coded posthoc using a scheme based on a framework for self-regulated hypermedia learning (Bannert, 2007). This framework describes hypermedia learning as an interplay of metacognitive, cognitive, and motivational activities. In the current study, we used a modified version of the original coding scheme; that is, we aggregated categories showing a rare occurrence. Additionally, we excluded the motivation category because motivational statements occurred very seldom in our data. Finally, the coding scheme comprised the main categories *metacognition* and *cognition*, and a total of seven sub-categories. Table 1 displays all categories, including descriptions and examples for each category.

Four trained research assistants coded the think-aloud data based on the procedure recommended by Chi (1997). We categorized the students' utterances by meaning, and we assigned a category for every definable learning activity. A total of 26,772 segments were labeled with a code. Furthermore, we selected a random sample of participants to compute the interrater reliability. Based on 2520 segments, this reliability showed a substantial agreement (Cohen's $\kappa = .78$).

Analysis Techniques

To address our research questions, we conducted the following two analysis steps using the coded think-aloud data. Table 2 summarizes the strategy for analyzing the students' learning activities. First, we considered only the frequencies of coded learning activities of both learning sessions. A mixed MANOVA was run using the treatment as an independent variable (between factor) and the coded learning activities as repeated-measures variables (within factor). This analysis compares the means of the learning activities of the first and second learning sessions within each group (main effect of time), between the two groups (main effect of group), and the interaction between these factors.

Table 1. Description of the categories for coding the students' think-aloud data

Code	Category	Description and Examples
Metacognition		
ANALYSE	Task Analysis	Task clarification, overviewing the material, goal setting, sub-goaling, and planning of proceeding <i>At first I read my learning goals to get an overview of my task. I have to learn the basic concepts of operant conditioning. First I will read the introductory text, then I will decide in which sequence I will proceed.</i>
SEARCH	Search	Searching for information <i>Now I'm looking for information on reinforcement plans.</i>
EVAL	Evaluation	Evaluating the attainment of goals or sub-goals <i>Did I process all topics according to my learning goals?</i>
MONITOR	Monitoring	Monitoring and controlling of one's learning, judgements of relevance of information <i>Ah, now I understand the principle. Skinner's Vita is not relevant for my learning task.</i>
Cognition		
READ	Reading	Reading text passages out loud
REPEAT	Repeating	Repeating in terms of memorizing <i>Re-reading a paragraph or notes</i>
PROCESS	Deep Processing	Elaborating and organizing: paraphrasing, connecting, and inferring <i>I already know the Skinner Box from my biology class. Drawing a map, writing down major concepts</i>

Second, we applied a PM technique called Conformance Checking (Rozinat & van der Aalst, 2008) to compare the sequential structure of coded learning activities between the first and the second learning sessions. This technique considers not only the frequencies of events, but also their relative arrangement (e.g., task analysis is followed by monitoring, which is followed by reading). For the first learning session, process models for the EG and the CG were already available and reported in a previous contribution (Sonnenberg & Bannert, 2015; see Appendix). Based on these models, we analyzed the sustainability of regulatory patterns during the second session using Conformance Checking. The analysis was conducted using the ProM framework Version 5.2 (2008).

Table 2. Overview of the strategy for analyzing the coded learning activities

	1 st Learning Session	2 nd Learning Session	Analysis Technique
Data of EG ($n = 35$) and CG ($n = 34$)	Frequencies of coded activities (means of seven categories)	Frequencies of coded activities (means of seven categories)	Mixed MANOVA
	Process models based on the sequence of coded activities	Sequence of learning activities: comparison with models of the 1 st session	Process Mining: Conformance Checking

Note. EG = experimental group, CG = control group. The seven categories are described in Table 1.

In the following, we summarize the basic concepts and output measures of Conformance Checking, based on the work of Rozinat and van der Aalst (2008). More detailed information is provided in their work. The basic idea is to compare the conformance between a process model and a sequence of observed activities, and to quantify the degree of conformance using metrics (e.g., fitness). To apply Conformance Checking, one has to consider two requirements. First, the process model has to be represented as a Petri net. The Petri net representation comprises transitions (i.e., executable tasks or activities), places which can hold tokens, and directed arcs between these elements. If all input places of a transition hold at least one token, it is enabled and it can be executed. Moreover, the execution of a transition produces tokens in all output places of this transition. In a Petri net, the state of a process is defined by the distribution of tokens. The second requirement refers to the observed behavior. It has to be stored in an event log which comprises chronologically ordered activities. Furthermore, a mapping of these activities to the transitions in the Petri net is needed.

Two orthogonal dimensions are the basis of conformance measurement: *fitness* and *appropriateness*. Fitness expresses the compliance of the observed sequence of activities and the control flow specified by the model (“Is it possible to produce the observed behavior by using the paths represented in the Petri net?”). It is computed by replaying the observed activities in the model while mismatches are recorded (e.g., if more tokens are needed or tokens are remaining). The token-based fitness f represents the extent of conformance, which ranges from 0 to 1. Additionally, the output provides more detailed diagnostic information. Places of observed mismatches, the path coverage (used or unused paths), and the number of passed edges (frequency of used paths) are reported.

The second dimension of conformance, appropriateness, refers to the question of whether the model describes the observed process in a suitable way. A process representation

might be too generic or too specific (i.e., allowing too much or, too little behavior), or it might include unnecessary elements (e.g., redundant transitions). Therefore, Conformance Checking also measures appropriateness on two sub-dimensions: *behavioral appropriateness* and *structural appropriateness*.

Behavioral appropriateness refers to the proportion of behavior allowed by the model and the behavior observed in the event log. It is measured by two metrics, *simple* and *advanced behavioral appropriateness*, both ranging from 0 to 1. Simple behavioral appropriateness is determined by considering the mean number of enabled transitions while replaying the observed activities. In addition to this procedure, the metric advanced behavioral appropriateness is computed by comparing follows and precedes relations specified by the model and by the event log. Since the first metric is dependent on the model flexibility, the second metric is needed to make comparisons among several models.

Furthermore, structural appropriateness covers the syntactical representation of behavior within the model. Again, two sub-dimensions are specified: *simple* and *advanced structural appropriateness*. The simple structural appropriateness compares the number of event classes and the graph size of the model. The second metric, advanced structural appropriateness, is independent of graph size and checks compliance with design principles (e.g., the avoidance of redundant elements). Again, both metrics range from 0 to 1.

Results

The findings of our analyses are reported as follows. First, we present the results of a frequency-based analysis using the coded learning activities of the first and second learning sessions. Second, we report the findings of applying a PM technique, namely Conformance Checking, which additionally takes into account the relative arrangement of learning activities. It is the aim of both analyses to examine the stability of fostered regulatory patterns in a follow-up learning task without metacognitive support. The Type I error rate was set to .05 for all statistical analyses.

Analysis Based on Frequencies

The descriptive statistics for the coded learning activities during the first and second learning sessions, separated by group, are reported in Table 3. As described above, the coding

scheme comprised seven sub-categories. Monitoring, Reading, and Deep Processing were the most frequent learning activities in both sessions. The categories Repeating, Task Analysis, Search, and Evaluation followed in decreasing order. Although the ranking of categories remained stable, there was a decreasing trend for all categories from the first to the second session, except for Reading. Considering the absolute difference of means, the decline was highest for the categories Monitoring (-18) and Deep Processing (-8), and lowest for Search (-3.5) and Evaluation (-1). In both sessions, the EG showed a higher mean value for the metacognitive categories Task Analysis, Evaluation, and Monitoring, as well as for the cognitive category Deep Processing, compared to the CG.

Table 3. Means and standard deviations for the number of coded learning activities during the first and second learning session

Category	1 st Session				2 nd Session			
	EG (n = 35)		CG (n = 34)		EG (n = 35)		CG (n = 34)	
Task Analysis								
<i>M (SD)</i>	18.11	(8.58)	13.24	(5.79)	11.51	(5.18)	9.65	(5.96)
Search								
<i>M (SD)</i>	11.83	(7.54)	13.24	(10.66)	11.06	(5.42)	9.09	(4.56)
Evaluation								
<i>M (SD)</i>	3.63	(3.08)	2.47	(3.08)	2.80	(2.91)	1.65	(1.52)
Monitoring								
<i>M (SD)</i>	82.86	(39.37)	67.44	(29.20)	59.80	(28.05)	53.50	(29.37)
Reading								
<i>M (SD)</i>	40.66	(15.75)	44.29	(18.43)	44.91	(16.54)	44.38	(15.89)
Repeating								
<i>M (SD)</i>	18.49	(11.04)	18.50	(12.28)	14.54	(11.28)	14.12	(10.29)
Deep Processing								
<i>M (SD)</i>	48.29	(21.54)	42.06	(20.27)	37.80	(18.99)	35.56	(19.30)

Note. Students in the experimental group (EG) received metacognitive prompts during the first learning session, but not during the second session. Students in the control group (CG) received no support.

We conducted a mixed MANOVA to examine the significance of metacognitive prompting effects on the learning activities deployed during the first and second learning sessions. The treatment (EG vs. CG) was used as an independent variable (between factor), and the coded learning activities of both sessions were used as repeated-measures variables (within factor).

Using Pillai's trace, there was a significant multivariate main effect of session on the seven learning activities, $V = 0.570$, $F(7, 61) = 11.571$, $p < .001$, $\eta_p^2 = .570$. Univariate ANOVAs revealed a significant change of the frequency of coded activities for all categories except Reading (see Table 4 for the statistics of the separate univariate comparisons).

Table 4. Test statistics of univariate follow-up comparisons of the mixed MANOVA

	MS	MS _{error}	F(1, 67)	p	η_p^2
between subjects					
Group					
Task Analysis	196.23	32.61	6.017	.017	.082
Search	1.36	32.85	0.041	.839	.001
Evaluation	23.03	5.14	4.478	.038	.063
Monitoring	2033.28	842.16	2.414	.125	.035
Reading	41.57	235.24	0.177	.676	.003
Repeating	0.73	102.83	0.007	.933	.000
Deep Processing	309.18	331.02	0.934	.337	.014
	MS	MS _{error}	F(1, 67)	p	η_p^2
within subjects					
Session					
Task Analysis	895.088	19.796	45.216	< .001	.403
Search	208.608	44.138	4.726	.033	.066
Evaluation	23.536	4.611	5.104	.027	.071
Monitoring	11804.072	345.983	34.117	< .001	.337
Reading	162.826	85.876	1.896	.173	.028
Repeating	597.666	47.134	12.680	.001	.159
Deep Processing	2487.915	142.397	17.472	< .001	.207
Session x Group					
Task Analysis	78.219	19.796	3.951	.051	.056
Search	98.260	44.138	2.226	.140	.032
Evaluation	0.000	4.611	0.000	.995	.000
Monitoring	716.594	345.983	2.071	.155	.030
Reading	149.869	85.876	1.745	.191	.025
Repeating	1.666	47.134	0.035	.851	.001
Deep Processing	136.987	142.397	0.962	.330	.014

Note. MS = Means squared, MS_{error} = Means squared of residual, η_p^2 = partial eta squared; Group = independent variable (EG vs. CG); Session = repeated-measures variable (first and second learning session); Session x Group = interaction effect.

Moreover, Pillai's trace for the main effect of group indicated no multivariate effect on the seven categories, $V = 0.144$, $F(7, 61) = 1.464$, $p = .197$, $\eta_p^2 = .144$. However, because of

the large effect size we decided to consider the univariate follow-up ANOVAs as well. There was a significant effect for the categories Task Analysis, $F(1, 67) = 6.017, p = .017, \eta_p^2 = .082$, and Evaluation, $F(1, 67) = 4.478, p = .038, \eta_p^2 = .063$. In both cases, students in the EG demonstrated a greater number of learning activities compared to students in the CG. The upper part of Table 4 shows the statistics of all univariate comparisons among subjects.

Finally, there was no significant multivariate interaction effect (session x group) on the seven categories, $V = 0.188, F(7, 61) = 2.020, p = .067, \eta_p^2 = .188$. Again, because of the large effect size, we also calculated univariate ANOVAs. As shown in Table 4, all univariate comparisons were non-significant.

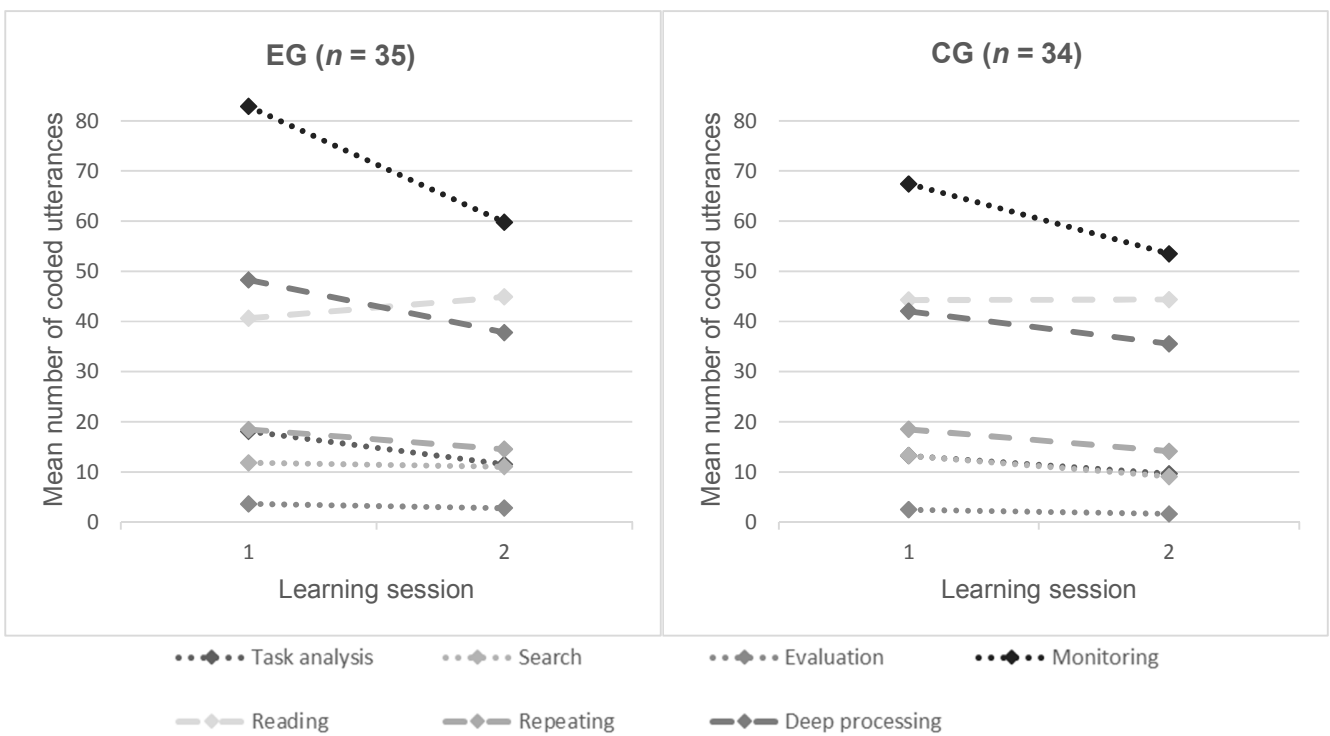


Figure 1. Course of all seven categories from the first to the second learning session. Left part: data of experimental group (EG), right part: data of control group (CG). The y-axis represents the mean absolute frequency of coded learning activities.

To illustrate our findings, Figure 1 presents the course of all seven categories, from the first to the second learning session, separately for the EG and the CG. In summary, our findings indicate that metacognitive prompting had beneficial effects on the quantity of regulatory learning activities, especially on the categories Task Analysis and Evaluation.

Despite a general downward trend of the mean number of coded utterances, the positive effects remained stable over time. Considering the descriptive statistics presented in Table 3, participants in the EG showed a higher decline in Task Analysis, Monitoring, and Deep Processing, but the means still remained above the values of the CG.

Process Mining – Conformance Checking

The second part of our analysis concentrated on the application of a PM technique called Conformance Checking. In general, PM allows the researcher to take into account the sequential structure of learning activities, in addition to their frequency distribution. We examined whether the learning process during the first session corresponded with the learning activities during the second session three weeks later. Specifically, we used process models describing learning during the first session to measure the conformance with the sequence of activities in the second learning session. The process models for the EG and the CG have already been presented in a previous contribution (Sonnenberg & Bannert, 2015), and they are portrayed in the Appendix, Figure A.1. In short, a comparison of the process models showed that Task Analysis was much better integrated in the learning process of the EG, and more loops between cognitive and metacognitive activities were observed for this group. Moreover, evaluation activities were similarly integrated, and monitoring activities were hardly connected with other learning activities in both models.

The following analysis tests the validity of these process models using the data of the second learning session. If metacognitive prompting effects were stable over time, we would expect a high conformance between the process models and the learning sequence during the second session. Furthermore, Conformance Checking indicates which patterns are transferred, and which deviations between models and event logs occurred, respectively.

Table 5 reports the conformance metrics using the dimensions *fitness* and *appropriateness*. These metrics were described in detail in the methods section. The high fitness value f for both groups indicates that the process models of the first learning session and the event logs of the second session corresponded very well. That means, the modeled patterns for both the EG and the CG allowed for the description of the observed learning events during the second learning session, which in general points to the stability of patterns. Regarding the precision of the process models, the simple behavioral appropriateness reached a medium value,

whereas the advanced behavioral appropriateness reached only a very low score. Consequently, these findings indicate a moderate precision, that is, the process models allowed for more behavior than actually observed in the event logs of the second learning session. Finally, the metric simple structural appropriateness reached a low value for both groups, which points to a suboptimal proportion of the number of categories and the graph size of the model. However, the advanced structural appropriateness was very high for both groups, which shows that design guidelines for process models were not violated (e.g., avoidance of redundant tasks).

Table 5. Measurement of conformance between process models and event logs using the dimensions fitness and appropriateness (precision and structure)

Metric	EG ($n = 35$)	CG ($n = 34$)
Fitness		
<i>f</i>	.96	.94
Precision		
<i>saB</i>	.43	.49
<i>aaB</i>	0.00	0.00
Structure		
<i>saS</i>	.14	.16
<i>aaS</i>	1.00	1.00

Note. *f* = token-based fitness metric, *saB* = simple behavioral appropriateness, *aaB* = advanced behavioral appropriateness, *saS* = simple structural appropriateness, *aaS* = advanced behavioral appropriateness. All metrics range from 0 to 1, with 1 being the highest extent of fitness or appropriateness, respectively.

Furthermore, the output of Conformance Checking allowed a more detailed view on the mismatches of the model-log comparison. Each trace of a participant (i.e., his or her sequence of learning activities) was replayed using the process model. Observed activities that could not be executed according to the model are highlighted as failed log events. Figure 2 shows the traces of five sample cases in which the failed events are colored. For instance, referring to the first case a transition from ANALYSE to PROCESS was observed in the event log, but according to the model this transition was not enabled. As described above, the model fitness was very high, but it still might be useful to consider the most common mismatches. In Table 6, we report the seven most frequent failed log events, separately for the EG and the CG. The findings for the EG show that transitions originating from MONITOR, in particular, failed during the model-log replay. Moreover, the transition from ANALYSE to SEARCH, and from

PROCESS to EVAL failed several times. Regarding the CG, transitions originating from MONITOR were also among the most common failed events. Furthermore, transitions directing to SEARCH failed especially frequently.

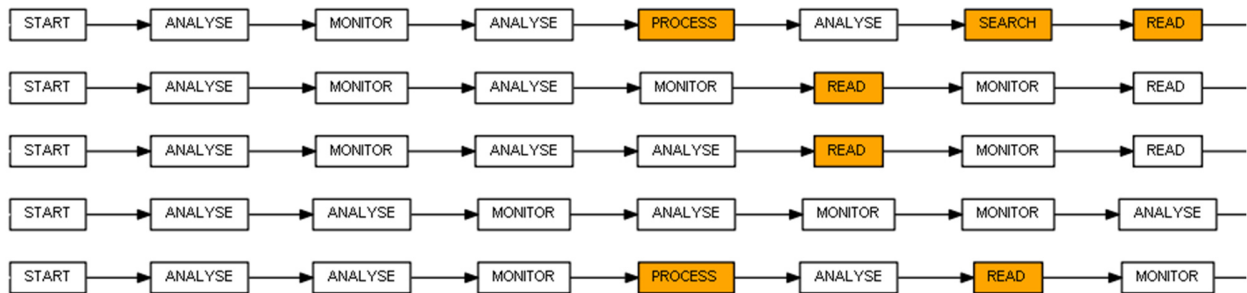


Figure 2. Log perspective of failed log events illustrated using five sample cases. Each line represents the first seven events of a case. Failed events measured by the log replay are colored. Codes are explained by the coding scheme presented in Table 1.

Table 6. Frequency of the seven most common failed log events

Number	Failed Log Event	EG (n = 35)		CG (n = 34)		
		Absolute Frequency	Relative Frequency	Failed Log Event	Absolute Frequency	Relative Frequency
1	MONITOR > EVAL	38	.18	MONITOR > SEARCH	59	.21
2	MONITOR > PROCESS	23	.11	ANALYSE > SEARCH	47	.17
3	MONITOR > READ	22	.10	MONITOR > EVAL	19	.07
4	ANALYSE > SEARCH	16	.08	READ > SEARCH	18	.06
5	PROCESS > EVAL	16	.08	MONITOR > READ	17	.06
6	MONITOR > SEARCH	14	.07	PROCESS > SEARCH	17	.06
7	PROCESS > REPEAT	13	.06	MONITOR > PROCESS	14	.05

Note. Total number of failed log events: experimental group (EG) = 212, control group (CG) = 280. Event_a > event_b means that the transition from event_a to event_b was not enabled in the process model, although observed in the event log.

In summary, the results of the mixed MANOVA using the frequencies of coded learning activities and Conformance Checking, which additionally takes into account the sequential structure of learning activities, indicated that fostered regulatory behavior through metacognitive prompting during the first session remained stable over time and was transferred

to a subsequent similar learning task. Despite a general downward trend of uttered learning activities in both groups, participants in the EG showed more metacognitive events during both sessions compared to the CG. Additionally, beneficial process patterns seemed to be transferred to the second learning task. Please note that a decrease in the frequency of learning activities from the first to the second session does not automatically result in a low conformance between the process model and the event log, because Conformance Checking primarily considers the sequential structure of activities. For example, an observed sequence of learning activities might be executed 10 times or 20 times, with each case resulting in a high conformance. Additionally, findings of short- and long-term effects on learning outcomes showed that metacognitive prompting does not affect only the learning process positively and sustainably as reported here, but also affects the performance in a learning test with transfer items (Bannert et al., 2015).

Discussion

The current study investigated the sustainability of learning activities stimulated by metacognitive prompting. Specifically, we addressed the question of whether fostered strategies are transferred to a similar subsequent task within the same learning environment three weeks later. The measurement of SRL processes using concurrent think-aloud protocols during two hypermedia learning sessions allowed us to observe learning behavior on a very detailed level. In addition to our analysis of the frequencies of learning activities, a PM analysis has allowed us to consider the sequential structure of observed events. The findings of a comparison between two learning sessions on the micro level provided new insights into the long-term impact of metacognitive prompting.

Our first research question addressed the stability of beneficial prompting effects in a follow-up learning task. The results of a mixed MANOVA using the frequencies of coded activities showed that the prompted EG demonstrated more metacognitive learning activities in both sessions. That means, metacognitive prompting also had beneficial effects in a follow-up task three weeks later without support. However, the findings also showed a general downward trend of metacognitive activities in both groups. The high fitness values using Conformance Checking support the assumption of stable effects between the first and second learning sessions. During the second session, a similar learning task within the same environment seemed to be sufficient to activate fostered strategies spontaneously. This finding is in line with Roll et

al. (2011), who found a transfer of fostered help-seeking skills to a subsequent task without support. Referring to SRL models (Winne & Hadwin, 2008; Zimmerman, 2008), the stability of effects indicates that our students perceived the usefulness of metacognitive strategies, because transfer requires the self-evaluation of one's learning process. Pressley et al. (1990) state three requirements for successful transfer of self-regulation skills: (i) the knowledge of strategies, (ii) the awareness that self-regulation is beneficial, and (iii) the competencies to adapt one's regulatory skills to new contexts. Since metacognitive prompting assumes the availability of strategic knowledge and our second task was designed similarly, the second precondition seems crucial for explaining our results. Hilbert et al. (2008) found no sustainable effect of cognitive and metacognitive prompting; however, more intensive prompting or more cues during the follow-up task would have been needed, especially because their participants were younger and probably not experienced with self-evaluation. Still, this shows that the question of how much support is needed, and for how long, until students internalize the fostered strategies is crucial to achieving stable effects.

Furthermore, the observed decrease in the frequency of metacognitive learning activities from the first to the second session, which was also observed for the CG, needs to be addressed. An explanation could be a growing automation of regulatory processes, particularly of monitoring activities, that resulted in a lower awareness of internal processes and therefore a lower frequency of uttered metacognitive events. Moreover, a related reason might be the adaptation to the demands of the task in the second session. Students could have adapted their regulatory behavior through repeated learning in the same learning environment. For example, Eitel (2016) showed that repeated testing affects multimedia learning through adaptation to the task. Similarly, the participants might have perceived the second task as easier because of a carryover effect through repeated learning. Further explanations for the decrease of metacognitive activities might be motivational issues, such as a change in self-efficacy that possibly affects the active regulation of learning (Moos & Miller, 2015), or the avoidance of repeatedly questioning one's learning behavior (Nückles et al., 2010). Other studies also report a decrease of metacognition over time (DiBenedetto & Bembenuddy, 2013; Labuhn, Bögelholz, & Hasselhorn, 2008; Nückles et al., 2010), but they consider different time frames (e.g., changes during a semester). Possibly this might also be a cause of repeatedly measuring metacognition using the same instrument.

Taking into account the findings of a previous contribution that investigates the long-term impact on learning outcome (Bannert et al., 2015), sustainable effects of metacognitive prompting both positively affected the learning process as well as the performance in a transfer test that requires the application of knowledge. The beneficial impact on learning outcomes is not self-evident, even if fostered strategies are transferred to a follow-up task, as observed in a study of Roll et al. (2011).

In our second research question, we aimed to identify sequential patterns that were transferred to the second learning session using Conformance Checking. PM techniques that take into account the relative arrangement of learning activities contributed to a microanalysis of learning, and they allowed conducting confirmatory model testing. Considering the mismatches between the first and second learning sessions, transitions originating from Monitoring frequently failed during the model-log replay. That means monitoring activities were embedded differently between the sessions. However, this might also be a deficiency of the Petri net notation as already discussed in Sonnenberg and Bannert (2015). Moreover, for the EG, transitions from Task Analysis to Search, and from Deep Processing to Evaluation failed several times. Compared to the first learning session with prompts, these patterns were observed less frequent. As discussed above, an explanation might be automation of regulation or adaptation to task demands. For the CG, transitions directing to Search failed frequently. Especially the transitions from Task Analysis to Search and the transitions from Monitoring to Search, which were executed during the second session, but not specified by the process model. An explanation could be that the students improved their systematic search for relevant information, possibly because they adapted to the demands of the task. Therefore, task analysis and monitoring activities are followed more frequently by search activities.

Although Conformance Checking resulted in high fitness values for both the models of the EG and CG, the low behavioral appropriateness indicated that the process models might not have been sufficiently precise. That means the models allowed for more behavior than actually observed during the replay with the data of the second session. Since both fitness and appropriateness represent the main criteria of models that adequately describe an event log, a better precision would have been desirable. The reason for the low precision is probably a high variability of learning sequences between subjects. The higher the variability in the event log, the more challenging is it for a PM algorithm to generate a single model describing the observed

behavior. A possible solution to improve precision could be the application of trace-based clustering (De Weerd, Vanden Broucke, Vanthienen, & Baesens, 2013) to select more homogeneous subgroups and to describe their behavior in separate models. However, the present sample might have been too small for this procedure.

Regarding the limitations of the current study, the findings are dependent on our learning setting (i.e., learning material, tasks, and hypermedia environment) and our sample. Consequently, the resulting patterns might be specific for the setting and the participants, and generalizations for other contexts and other samples still have to be verified in future research. For example, it is possible that a transfer of strategies works for university students, as in our sample, but not for younger school students, because they are not experienced in self-evaluating their learning. Furthermore, the coding scheme used to categorize the students' utterances determines the level of granularity for the observation of learning behavior. Another measurement unit or process data from other channels (e.g., log files or eye-tracking) might result in different findings. Moreover, a pre-determined time frame of three weeks was used to evaluate the sustainability of metacognitive prompting. Therefore, it is not possible to determine for how long the beneficial effects lasted, whether it was for weeks or even months. More specific information would be needed to better control the fading out of support when a student progresses in spontaneously performing the desired activities. In addition, the concept of booster sessions (e.g., Souvignier & Trenk-Hinterberger, 2010) might be useful to refresh the fostered strategies and to guarantee the stability of beneficial effects over a longer period. Finally, we did not assess motivational constructs like self-efficacy before learning, which might affect the regulation behavior (e.g., Bernacki, Nokes-Malach, & Alevan, 2015; Moos & Miller, 2015). Therefore, it is not possible to determine if a change in motivation caused the decrease of metacognitive activities from the first to the second session.

Future research needs to replicate findings on the sustainability of metacognitive prompting. In general, more studies addressing the stability of SRL processes and the transfer of strategies into subsequent similar tasks and other domains are needed. Using process data like think-aloud protocols provides deeper insights into the learning behavior; but replications using different data channels, as well as a triangulation of channels, could advance the validity of findings (Azevedo, 2014). With respect to the discovery and testing of process models, future analyses should investigate the benefits of trace-based clustering to improve their precision (De

Medeiros et al., 2008; De Weerd et al., 2013). Moreover, Conformance Checking also allows for the comparison of a theoretical-build model and an event log. Therefore, it would be possible to test theoretical assumptions of SRL models with observed behavior from various learning sessions.

In conclusion, the current study provides initial evidence for sustainable long-term effects of metacognitive prompts on hypermedia learning. From an analytical point of view, our analysis indicates the benefits of evaluating instructional support on the micro level using PM techniques, which take into account the sequential structure of learning activities. SRL process data and analysis techniques that consider the dynamics of SRL contribute to the advanced understanding of regulatory processes and the impact of support for them. In addition, our findings can support the refinement of current SRL models by providing new insight into micro-level processes.

References

- Azevedo, R. (2009). Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: A discussion. *Metacognition and Learning, 4*, 87–95. doi:10.1007/s11409-009-9035-7
- Azevedo, R. (2014). Issues in dealing with sequential and temporal characteristics of self- and socially-regulated learning. *Metacognition and Learning, 9*, 217–228. doi:10.1007/s11409-014-9123-1
- Azevedo, R., Cromley, J. G., Moos, D. C., Greene, J. A., & Winters, F. I. (2011). Adaptive content and process scaffolding: A key to facilitating students' self-regulated learning with hypermedia. *Psychological Testing and Assessment Modeling, 53*(1), 106–140.
- Azevedo, R., Harley, J. M., Trevors, G., Duffy, M., Feyzi-Behnagh, M., Bouchet, F., & Landis, R. (2013). Using trace data to examine the complex roles of cognitive, metacognitive, and emotional self-regulatory processes during learning with multi-agent systems. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 427–449). New York: Springer. doi:10.1007/978-1-4419-5546-3_28
- Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist, 45*, 210–223. doi:10.1080/00461520.2010.515934
- Baker, R., & Corbett, A. (2014). Assessment of robust learning with educational data mining. *Research & Practice in Assessment, 9*, 38–50.

- Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research*, 35(4), 359–375. doi:10.2190/94V6-R58H-3367-G388
- Bannert, M. (2007). *Metakognition beim Lernen mit Hypermedia. Erfassung, Beschreibung und Vermittlung wirksamer metakognitiver Lernstrategien und Regulationsaktivitäten [Metacognition and Hypermedia Learning]*. Münster: Waxmann.
- Bannert, M. (2009). Promoting self-regulated learning through prompts: A discussion. *Zeitschrift für Pädagogische Psychologie*, 23(2), 139–145. doi:10.1024/1010-0652.23.2.139
- Bannert, M., & Mengelkamp, C. (2013). Scaffolding hypermedia learning through metacognitive prompts. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 171–186). New York: Springer. doi:10.1007/978-1-4419-5546-3_12
- Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science*, 40(1), 193–211. doi:10.1007/s11251-011-9167-4
- Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning*, 9, 161–185. doi:10.1007/s11409-013-9107-6
- Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior*, 52, 293–306. doi:10.1016/j.chb.2015.05.038
- Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: A focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, 10, 1–13. doi:10.1007/s11409-015-9134-6
- Bernacki, M. L., Nokes-Malach, T. J., & Aleven, V. (2015). Examining self-efficacy during learning: Variability and relations to behavior, performance, and learning. *Metacognition and Learning*, 10, 99–117. doi:10.1007/s11409-014-9127-x
- Berthold, K., Röder, H., Knörzer, D., Kessler, W., & Renkl, A. (2011). The double-edged effects of explanation prompts. *Computers in Human Behavior*, 27, 69–75. doi:10.1016/j.chb.2010.05.025
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, 6(3), 271–315. doi:10.1207/s15327809jls0603_1
- De Medeiros, A. K. A., Guzzo, A., Greco, G., van der Aalst, W. M. P., Weijters, A. J. M. M., van Dongen, B. F., & Saccà, D. (2008). Process mining based on clustering: A quest for precision. *Lecture Notes in Computer Science*, 4928, 17–29. doi:10.1007/978-3-540-

78238-4_4

- De Weerd, J., Vanden Broucke, S., Vanthienen, J., & Baesens, B. (2013). Active trace clustering for improved process discovery. *IEEE Transactions on Knowledge and Data Engineering*, *25*(10), 2708–2720. doi:10.1109/TKDE.2013.64
- Dent, A. L., & Hoyle, R. H. (2015). A framework for evaluating and enhancing alignment in self-regulated learning research. *Metacognition and Learning*, *10*, 165–179. doi:10.1007/s11409-015-9136-4
- Devolder, A., van Braak, J., & Tondeur, J. (2012). Supporting self-regulated learning in computer-based learning environments: Systematic review of effects of scaffolding in the domain of science education. *Journal of Computer Assisted Learning*, *28*, 557–573. doi:10.1111/j.1365-2729.2011.00476.x
- DiBenedetto, M. K., & Bembenuy, H. (2013). Within the pipeline: Self-regulated learning, self-efficacy, and socialization among college students in science courses. *Learning and Individual Differences*, *23*(1), 218–224. doi:10.1016/j.lindif.2012.09.015
- Eitel, A. (2016). How repeated studying and testing affects multimedia learning: Evidence for adaptation to task demands. *Learning and Instruction*, *41*, 70–84. doi:10.1016/j.learninstruc.2015.10.003
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, Mass.: MIT Press.
- Greene, J. A., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist*, *45*, 203–209. doi:10.1080/00461520.2010.515935
- Greene, J. A., Dellinger, K. R., Tüysüzoglu, B. B., & Costa, L.-J. (2013). A two-tiered approach to analyzing self-regulated learning data to inform the design of hypermedia learning environments. In R. Azevedo & V. Aleven (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 117–128). New York: Springer. doi:10.1007/978-1-4419-5546-3_8
- Hilbert, T. S., Nückles, M., Renkl, A., Minarik, C., Reich, A., & Ruhe, K. (2008). Concept mapping zum Lernen aus Texten: Können Prompts den Wissens- und Strategieerwerb fördern? [Concept mapping as a follow-up strategy for learning from texts: Can the acquisition of knowledge and skills be fostered by prompts?]. *Zeitschrift für Pädagogische Psychologie*, *22*(2), 119–125. doi:10.1024/1010-0652.22.2.119
- Hübner, S., Nückles, M., & Renkl, A. (2010). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction*, *20*, 18–29. doi:10.1016/j.learninstruc.2008.12.001

- Jeong, H., Gupta, A., Roscoe, R., Wagster, J., Biswas, G., & Schwartz, D. (2008). Using hidden markov models to characterize student behaviors in learning-by-teaching environments. *Lecture Notes in Computer Science*, 5091, 614–625. doi:10.1007/978-3-540-69132-7_64
- Künsting, J., Kempf, J., & Wirth, J. (2013). Enhancing scientific discovery learning through metacognitive support. *Contemporary Educational Psychology*, 38(4), 349–360. doi:10.1016/j.cedpsych.2013.07.001
- Labuhn, A. S., Bögelholz, S., & Hasselhorn, M. (2008). Selbstregulationsförderung in einer Biologie-Unterrichtseinheit. Langfristige und Differenzielle Wirksamkeit [Fostering self-regulated learning in science education. Long-term and subgroup effects]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40(4), 167–178. doi:10.1026/0049-8637.40.4.167
- Lehmann, T., Hähnlein, I., & Ifenthaler, D. (2014). Cognitive, metacognitive and motivational perspectives on preflexion in self-regulated online learning. *Computers in Human Behavior*, 32, 313–323. doi:10.1016/j.chb.2013.07.051
- Malmberg, J., Järvelä, S., Järvenoja, H., & Panadero, E. (2015). Promoting socially shared regulation of learning in CSCL: Progress of socially shared regulation among high- and low-performing groups. *Computers in Human Behavior*, 52, 562–572. doi:10.1016/j.chb.2015.03.082
- Michalke, M. (2015). koRpus: An R package for text analysis (Version 0.06-3). Retrieved from <http://reaktanz.de/?c=hacking&s=koRpus>
- Molenaar, I., & Järvelä, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning*, 9, 75–85. doi:10.1007/s11409-014-9114-2
- Moos, D. C., & Azevedo, R. (2009). Self-efficacy and prior domain knowledge: To what extent does monitoring mediate their relationship with hypermedia learning? *Metacognition and Learning*, 4, 197–216. doi:10.1007/s11409-009-9045-5
- Moos, D. C., & Miller, A. (2015). The cyclical nature of self-regulated learning phases: Stable between learning tasks? *Journal of Cognitive Education and Psychology*, 14(2), 199–218. doi:10.1891/1945-8959.14.2.199
- Nückles, M., Hübner, S., Dümer, S., & Renkl, A. (2010). Expertise reversal effects in writing-to-learn. *Instructional Science*, 38(3), 237–258. doi:10.1007/s11251-009-9106-9
- Pieschl, S., Stahl, E., Murray, T., & Bromme, R. (2012). Is adaptation to task complexity really beneficial for performance? *Learning and Instruction*, 22, 281–289. doi:10.1016/j.learninstruc.2011.08.005

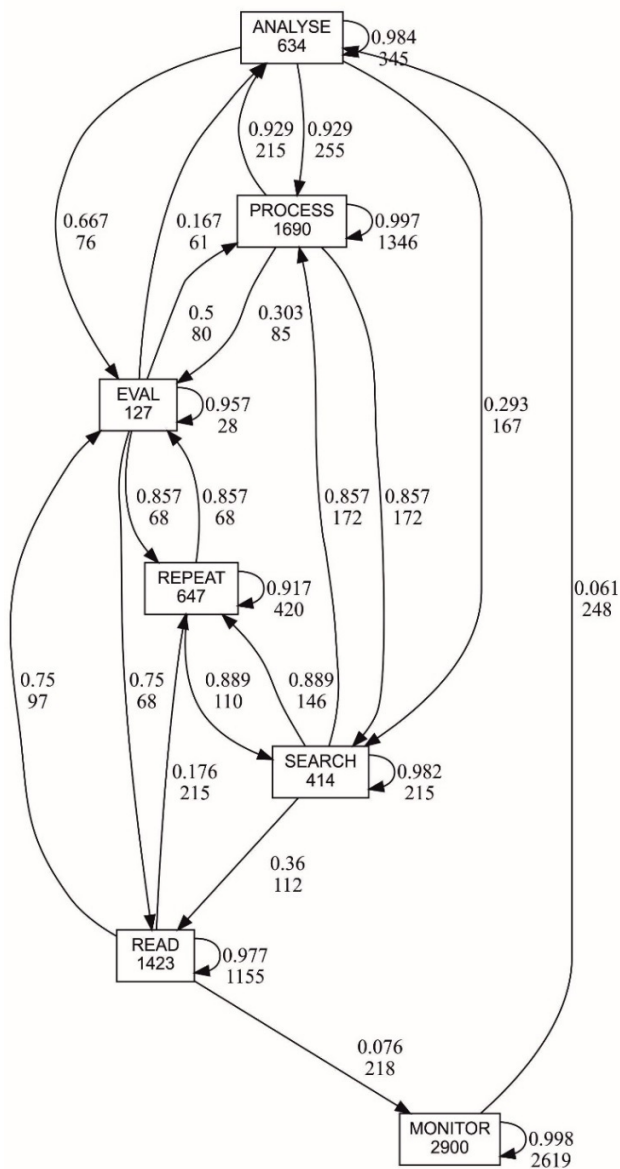
- Pressley, M., Woloshyn, V., Lysynchuk, L. M., Martin, V., Wood, E., & Willoughby, T. (1990). A primer of research on cognitive strategy instruction: The important issues and how to address them. *Educational Psychology Review*, 2(1), 1–58. doi:10.1007/BF01323528
- ProM (Version 5.2) [Computer Software]. (2008). Retrieved from <http://www.processmining.org>
- Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, 40, 1–12. doi:10.1207/s15326985ep4001_1
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 239–257. doi:10.1007/s11412-009-9070-z
- Reimann, P., Frerejean, J., & Thompson, K. (2009). Using process mining to identify models of group decision making processes in chat data. In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Computer-supported collaborative learning practices: CSCL2009 conference proceedings* (pp. 98–107). Rhodes, Greece: International Society for the Learning Sciences.
- Reimann, P., Markauskaite, L., & Bannert, M. (2014). E-Research and learning theory: What do sequence and process mining methods contribute? *British Journal of Educational Technology*, 45(3), 528–540. doi:10.1111/bjet.12146
- Reimann, P., & Yacef, K. (2013). Using process mining for understanding learning. In R. Luckin, P. Puntambekar, P. Goodyear, B. Grabowski, J. Underwood, & N. Winters (Eds.), *Handbook of Design in Educational Technology* (pp. 472–481). New York: Routledge.
- Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction*, 21, 267–280. doi:10.1016/j.learninstruc.2010.07.004
- Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, 2, 7–12. Retrieved from <http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/4491>
- Roll, I., Yee, N., & Briseno, A. (2014). Students' adaptation and transfer of strategies across levels of scaffolding in an exploratory environment. *Lecture Notes in Computer Science*, 8474, 348–353. doi:10.1007/978-3-319-07221-0_43
- Rozinat, A., & van der Aalst, W. M. P. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems*, 33(1), 64–95. doi:10.1016/j.is.2007.07.001

- Schoor, C., & Bannert, M. (2012). Exploring regulatory processes during a computer-supported collaborative learning task using process mining. *Computers in Human Behavior*, *28*, 1321–1331. doi:10.1016/j.chb.2012.02.016
- Schunk, D. H., & Ertmer, P. A. (2005). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 631–649). San Diego, California: Academic Press. doi:10.1016/B978-012109890-2/50048-2
- Severiens, S., Ten Dam, G., & Van Hout-Wolters, B. (2001). Stability of processing and regulation strategies: Two longitudinal studies on student learning. *Higher Education*, *42*(4), 437–453. doi:10.1023/a:1012227619770
- Siadaty, M., Gašević, D., & Hatala, M. (2016a). Associations between technological scaffolding and micro-level processes of self-regulated learning: A workplace study. *Computers in Human Behavior*, *55*, 1007–1019. doi:10.1016/j.chb.2015.10.035
- Siadaty, M., Gašević, D., & Hatala, M. (2016b). Measuring the impact of technological scaffolding interventions on micro-level processes of self-regulated workplace learning. *Computers in Human Behavior*, *59*, 469–482. doi:10.1016/j.chb.2016.02.025
- Sonnenberg, C., & Bannert, M. (2015). Discovering the effects of metacognitive prompts on the sequential structure of SRL-processes using process mining techniques. *Journal of Learning Analytics*, *2*(1), 72–100. Retrieved from <http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/4090>
- Sonnenberg, C., & Bannert, M. (2016). Evaluating the impact of instructional support using data mining and process mining: A micro-level analysis of the effectiveness of metacognitive prompts. *Journal of Educational Data Mining*, *8*(2), 51-83.
- Souvignier, E., & Trenk-Hinterberger, I. (2010). Implementation eines Programms zur Förderung selbstregulierten Lesens. Verbesserung der Nachhaltigkeit durch Auffrischungssitzungen [Implementation of a program to foster reading competence: Improving long-term effects by using booster-sessions]. *Zeitschrift für Pädagogische Psychologie*, *24*, 207–220. doi:10.1024/1010-0652/a000017
- Stark, R., & Krause, U. M. (2009). Effects of reflection prompts on learning outcomes and learning behaviour in statistics education. *Learning Environments Research*, *12*, 209–223. doi:10.1007/s10984-009-9063-x
- Thillmann, H., Künsting, J., Wirth, J., & Leutner, D. (2009). Is it merely a question of “what” to prompt or also “when” to prompt? *Zeitschrift für Pädagogische Psychologie*, *23*(2), 105–115. doi:10.1024/1010-0652.23.2.105
- Trčka, N., Pechenizkiy, M., & van der Aalst, W. M. P. (2010). Process mining from educational data. In C. Romero, S. Ventura, M. Pechenizkiy, & R. Baker (Eds.),

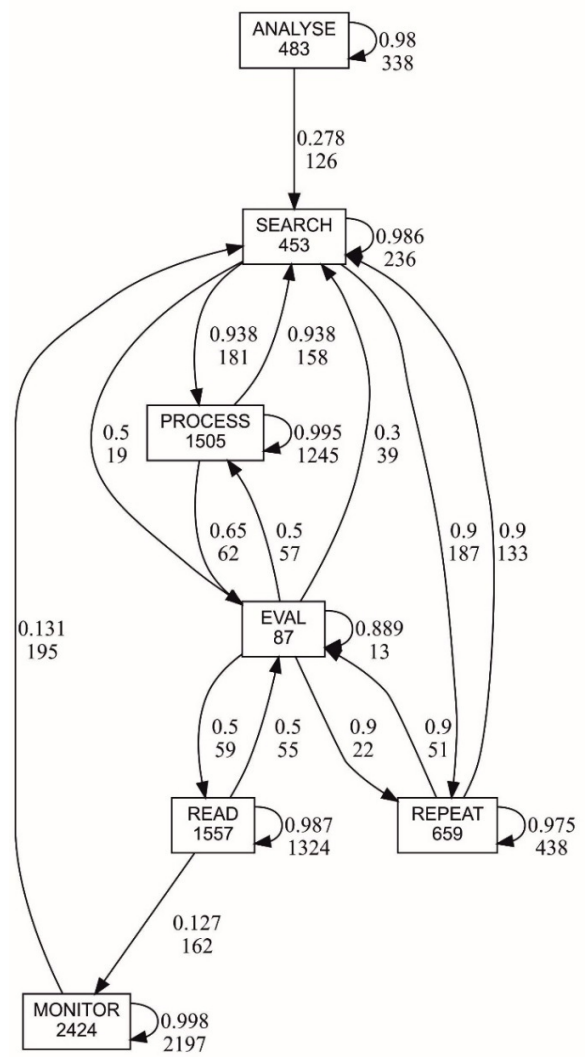
- Handbook of Educational Data Mining* (pp. 123–142). Boca Raton, FL: Chapman&Hall/CRC.
- Trevors, G., Feyzi-Behnagh, R., Azevedo, R., & Bouchet, F. (2016). Self-regulated learning processes vary as a function of epistemic beliefs and contexts: Mixed method evidence from eye tracking and concurrent and retrospective reports. *Learning and Instruction, 42*, 31–46. doi:10.1016/j.learninstruc.2015.11.003
- Van der Aalst, W. M. P. (2011). *Process mining: Discovery, conformance and enhancement of business processes*. Berlin, Heidelberg: Springer. doi.org/10.1007/978-3-642-19345-3
- Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences, 8*(4), 327–353. doi:10.1016/S1041-6080(96)90022-9
- Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning, 9*, 229–237. doi:10.1007/s11409-014-9113-3
- Winne, P. H., & Baker, R. S. (2013). The potentials of educational data mining for researching metacognition, motivation and self-regulated learning. *Journal of Educational Data Mining, 5*(1), 1–8. Retrieved from <http://www.educationaldatamining.org/JEDM/index.php/JEDM/article/view/28>
- Winne, P. H., & Hadwin, A. F. (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). NY: Taylor & Francis.
- Winne, P. H., & Perry, N. E. (2005). Measuring self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 531–566). San Diego, California: Academic Press.
- Wirth, J. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie, 23*(2), 91–94. doi:10.1024/1010-0652.23.2.91
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 17*(2), 89–100. doi:10.1111/j.1469-7610.1976.tb00381.x
- Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: A meta-analysis. *Asia Pacific Education Review, 17*(2), 187–202. doi:10.1007/s12564-016-9426-9
- Zhou, M., Xu, Y. X., Nesbit, J. C., & Winne, P. H. (2010). Sequential pattern analysis of learning logs: Methodology and applications. In C. Robero, S. Ventura, M. Pechenizkiy, & R. Baker (Eds.), *Handbook of Educational Data Mining* (pp. 107–121). Boca Raton, FL: Chapman&Hall/CRC.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183. doi:10.3102/0002831207312909

Appendix



Experimental Group



Control Group

Figure A.1. Process models represented as a heuristic net based on the sequence of coded activities of the first learning session, from Sonnenberg and Bannert (2015, p. 91). For a description of the codes please see the coding scheme presented in Table 1.