The order matters: Sequencing strategies in example-based learning

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie
an der Ludwig-Maximilians-Universität München

vorgelegt von

Julia Murböck

aus Schweinfurt

2018

Erstgutachter: PD Dr. Karsten Stegmann

Zweitgutachter: Prof. Dr. Jan-Willem Strijbos

Datum der mündlichen Prüfung: 13.07.2018

Danksagung

„Leider lässt sich eine wahrhafte Dankbarkeit nicht mit Worten ausdrücken.“ (Goethe)

Ich will es trotzdem versuchen. Zuerst möchte ich mich bei meinen Doktorvätern PD Dr. Karsten Stegmann und Prof. Dr. Jan-Willem Strijbos bedanken. Danke, Karsten, für deine Offenheit und die Freiheit, die du mir gegeben hast, meinen eigenen Weg zu gehen. Danke für dein offenes Ohr und deine wertvollen Ratschläge, die mir immer wieder geholfen haben, den Blick aufs Wesentliche (zurück) zu richten. Danke, Jan-Willem, für die Struktur, mit der du mir einen hilfreichen Rahmen zum Schreiben gegeben hast und die verhindert hat, dass ich dem „auf den letzten Drücker“-Phänomen verfalle. Danke für dein konstruktives Feedback, das mich Schritt für Schritt weiter voran gebracht hat.

Darüber hinaus möchte ich mich auch bei meinem Lehrstuhl-Chef Prof. Dr. Frank Fischer bedanken. Danke, Frank, dass du mich immer in meinem Vorhaben unterstützt hast und mir den Freiraum gegeben hast, meine Dissertation gut zu Ende zu bringen.

Ein großer Dank gilt auch meinen Kolleginnen und Kollegen. Danke, Jana, Uli, Michi, Sarah, Barbara und viele mehr, dass ihr mich kollegial wie auch mental und emotional in den Höhen und Tiefen während der Diss unterstützt habt und mir mit wertvollen Tipps und Feedback zur Seite standet. Vielen Dank, Georg, für das Drehen und Schneiden der Videos für meine Studien.

Danke an die Studentinnen und Studenten, die im Rahmen der EFM-Ausbildung an meinen Studien teilgenommen haben. Auch vielen Dank an die Studierenden, die während ihrer Projektarbeiten oder Hiwijobs einen wesentlichen Beitrag zur Auswertung der Daten geleistet haben.

Auch meinen Freunden und meiner Familie gebührt großer Dank. Danke, dass ihr in angenehmer Unaufgeregtheit meinen Doktorarbeitsprozess begleitet habt und mich in der heißen Phase zum Schluss auch mal für kurze Zeit „untertauchen“ habt lassen.

Last, but not least: Vielen Dank an meinen Mann Michael. Danke, dass du mich in jeder Doktorarbeitslage unterstützt hast und mich immer zur rechten Zeit damit in Ruhe gelassen hast, aber auch gepusht hast. Danke, dass du mir daheim den Rücken frei gehalten hast und immer für mich da warst und bist.

Deutsche Zusammenfassung

Beispielbasiertes Lernen ist eine effektive Methode, den anfänglichen kognitiven Fertigkeitserwerb zu unterstützen (Renkl, 2014, Van Gog & Rummel, 2010). Beim beispielbasierten Lernen wird die Lösung eines Problems anhand von Lösungsbeispielen oder Modellierungsbeispielen demonstriert. Typischerweise werden Lösungs- und Modellierungsbeispiele mit Problemlöseaufgaben kombiniert (z. B. Cooper & Sweller, 1987; Crippen & Earl, 2007; Hilbert & Renkl, 2009; Paas & Van Merriënboer, 1994). Dabei stellt sich die Frage, in welcher Reihenfolge Modellierungsbeispiele und Problemlöseaufgaben dargeboten werden sollten, um den anfänglichen kognitiven Fertigkeitserwerb zu fördern. Wenn für ähnliche Problemtypen mehrere Modellierungsbeispiele und Problemlöseaufgaben kombiniert werden, sind mindestens vier verschiedene Sequenzierungsstrategien möglich: Sequenzierung der Modellierungsbeispiele (1) abwechselnd vor den Problemlöseaufgaben, (2) geblockt vor den Problemlöseaufgaben, (3) abwechselnd nach den Problemlöseaufgaben und (4) geblockt nach den Problemlöseaufgaben.

Die Sequenzierung von Modellierungsbeispielen (1) abwechselnd vor den Problemlöseaufgaben hat den Vorteil, dass die Lerner dadurch kognitive Schemata erwerben können, wodurch die extrinsische kognitive Belastung reduziert wird, weil die arbeitsgedächtnis-belastende Mittel-Ziel-Strategie durch die erworbenen Schemata unnötig wird (Renkl, 2014; Sweller, 2010). Ein zweiter Vorteil ist, dass durch die erworbenen Schemata die Problemlöseprinzipien, die durch die Beispiele verdeutlicht werden, besser angewendet werden können (Trafton & Reiser, 1993; Van Gog et al., 2011). Durch die geringere extrinsische kognitive Belastung und die höhere Anwendungsqualität der Prinzipien soll der kognitive Fertigkeitserwerb gefördert werden (Renkl, 2014). Für eine Sequenzierung der Modellierungsbeispiele (2) geblockt vor den Problemlöseaufgaben sprechen theoretisch die gleichen Argumente. Allerdings könnten die Lernenden Probleme haben, alle Beispiele zu erinnern und auch die richtigen Beispiele den entsprechenden Problemlöseaufgaben zuzuordnen (Trafton & Reiser, 1993). Das könnte dazu führen, dass die Lernenden bei der Bearbeitung der Problemlöseaufgaben doch auf die arbeitsgedächtnis-belastende Mittel-Ziel-Strategie angewiesen sind. Bei einer Sequenzierung der Modellierungsbeispiele (3) abwechselnd nach den Problemlöseaufgaben sollten die Lernenden aufgrund der fehlenden Schemata auf die Mittel-Ziel-Strategie zur Bearbeitung der Problemlöseaufgaben zurückgreifen müssen, was in der Regel zu einer höheren extrinsischen kognitiven Belastung führt (Renkl, 2014; Sweller, 2010) und sich darüber hinaus negativ auf die Anwendungsqualität auswirken kann und dementsprechende auch negativ auf den kognitiven Fertigkeitserwerb wirken sollte (Van Gog et al., 2011). Ein Vorteil dieser Sequenzierung könnte allerdings sein, dass die Lernenden dadurch merken, was sie nicht wissen und dadurch ihre Aufmerksamkeit in den jeweils folgenden Modellierungsbeispielen auf die Aspekte richten, die sie als kritisch identifiziert haben (Loibl, Roll, & Rummel, 2017; Van Gog et al., 2011). Dies wiederum sollte helfen, die nachfolgende Modellierung besser an das Vorwissen anknüpfen zu können und dadurch den kognitiven Fertigkeitserwerb fördern (Loibl & Rummel, 2014). Für eine Sequenzierung der Modellierungsbeispiele (4) geblockt nach den Problemlöseaufgaben zählen theoretisch die gleichen Argumente, sowohl in Bezug auf eine mögliche negative Auswirkung auf den kognitiven Fertigkeitserwerb als auch eine mögliche positive Wirkung. Ein Unterschied zur Sequenzierung abwechselnd nachher ist jedoch, dass die Lernenden die kognitiven Schemata, die sie durch die Modellierung erwerben, nicht mehr anwenden können, wenn die Modellierungsbeispiele geblockt nach den Problemlöseaufgaben dargeboten werden. Die Studienlage zum Vergleich der verschiedenen Sequenzierungsstrategien liefert gemischte Befunde. Zum einen sprechen Befunde dafür, Modellierungsbeispiele (abwechselnd) vor Problemlöseaufgaben anstatt (abwechselnd) danach zu sequenzieren, um den kognitiven Fertigkeitserwerb zu fördern (z. B. Hsu et al., 2015; Kant et al., 2017; Leppink et al., 2014; Van Gog et al., 2011). Zum anderen zeigen die Studien, dass Modellierungsbeispiele (geblockt) nach der Bearbeitung von Problemlöseaufgaben erfolgreicher für den kognitiven Fertigkeitserwerb sind (geblockt) davor (z. B. Baggett, 1987; Kapur, 2012; Schwartz et al., 2011). Mögliche Erklärungen für die unterschiedliche Befundlage könnten sein, dass ein positiver Effekt der Sequenzierung von Modellierungsbeispielen vor Problemlöseaufgaben durch eine reduzierte extrinsische kognitive Belastung sowie eine bessere Anwendungsqualität vermittelt wird (Van Gog, 2011; Van Gog et al., 2011). Ein positiver Effekt scheint dagegen von der Sequenzierung von Modellierungsbeispielen nach Problemlöseaufgaben von einer Förderung der metakognitiven Überwachung und damit einhergehender Vorwissensdifferenzierung oder der Art der Lernaufgabe abzuhängen (Alfieri et al., 2013; Loibl, Roll & Rummel, 2017).

Die bisherigen Studien haben, wenn überhaupt, nur einen Teil der möglichen Erklärungen betrachtet, weshalb hier noch deutlicher Forschungsbedarf besteht. Die vorliegende Dissertation geht deshalb der Frage nach, welchen Einfluss die Sequenzierung von Modellierungsbeispielen und Problemlöseaufgaben auf den kognitiven Fertigkeitserwerb hat und vor allem, wie ein möglicher Einfluss erklärt werden kann. Speziell wird untersucht, ob ein möglicher positiver Einfluss der Sequenzierung von Modellierungsbeispielen vor Problemlöseaufgaben durch extrinsische kognitive Belastung und Anwendungs-Qualität vermittelt wird und ob ein möglicher positiver Einfluss der Sequenzierung von Modellierungsbeispielen nach Problemlöseaufgaben davon abhängt, ob (meta-)kognitive Prozesse (d.h. Selbsterklärungen und Überwachung) angeregt werden oder ob es von der Art der Lernaufgabe abhängt, dass ein positiver Effekt von nachher gezeigt werden kann. Um diesen generellen Forschungsfragen nachzugehen, wurden zwei empirische Studien durchgeführt.

Studie I untersuchte, inwiefern die Sequenzierung von Modellierungsbeispielen und Problemlöseaufgaben (geblockt vorher, abwechselnd vorher oder geblockt nachher) einen Einfluss auf den kognitiven Fertigkeitserwerb hat. Außerdem wurde untersucht, inwieweit ein Effekt der Sequenzierung durch extrinsische kognitive Belastung sowie die Anwendungsqualität vermittelt wird und ob ein Effekt der Sequenzierung von der Anregung von (meta-)kognitiven Prozessen abhängt. Die Stichprobe der Studie I bestand aus 126 Pädagogik-Studenten im ersten Semester. Die Studierenden wurden zufällig auf eine von sechs Bedingungen zugeteilt, welche in einem 2×3 faktoriellen Design mit zwei Zwischensubjekt-Faktoren (Anregungen: ja vs. nein; Sequenzierung: geblockt vorher vs. abwechselnd vorher vs. geblockt nachher) und einem Messwiederholungs-Faktor (kognitiver Fertigkeitserwerb zu T1 und T2) variiert wurden. Extrinsische kognitive Belastung und Anwendung-Qualität wurden erhoben, während die Studierenden die Lernaufgaben bearbeitet haben. Die Ergebnisse zeigen einen positiven Effekt der Sequenzierung von Modellierungsbeispielen abwechselnd vor den Problemlöseaufgaben auf den kognitiven Fertigkeitserwerb im Vergleich zu geblockt vorher und geblockt nachher. Außerdem wird dieser Effekt durch eine höhere Anwendungs-Qualität vermittelt. Extrinsische kognitive Belastung ist dagegen kein signifikanter Mediator, auch wenn die

Sequenzierung von Modellierungsbeispielen abwechselnd vor den Problemlöseaufgaben die extrinsische kognitive Belastung im Vergleich zu den beiden anderen Sequenzen reduziert hat. Außerdem war der Sequenzierungseffekt nicht abhängig von der Anregung von Selbsterklärungen und Überwachung. Diese Ergebnisse werden im Anschluss an die Darstellung der Studie II und ihrer Befunde gemeinsam diskutiert.

Studie II ging der Frage nach, inwieweit die Sequenzierung von Modellierungsbeispielen abwechselnd vor oder abwechselnd nach Lernaufgaben einen Effekt auf den kognitiven Fertigkeitserwerb hat. Außerdem wurde untersucht, inwieweit ein möglicher positiver Effekt der Sequenzierung abwechselnd vorher durch extrinsische kognitive Belastung sowie die Anwendungsqualität vermittelt wird und ob ein möglicher positiver Effekt der Sequenzierung abwechselnd nachher von der Art der Lernaufgabe abhängt. Dazu wurden drei Lernaufgaben verglichen: (1) Problemlöseaufgaben, (2) Vergleichsaufgaben (Vergleichen von Beispielen) und (3) Problemlöseaufgaben mit Modell-Vergleich. Die Stichprobe bestand aus 145 Pädagogik-Studenten im ersten Semester. Die Studierenden wurden zufällig auf eine von sechs Bedingungen zugeteilt, welche in einem 2×3 faktoriellen Design mit zwei Zwischensubjekt-Faktoren (Sequenzierung: abwechselnd vorher vs. Abwechselnd nachher; Art der Lernaufgabe: Problemlöseaufgaben vs. Vergleichsaufgaben vs. Problemlöseaufgaben mit Modell-Vergleich) variiert wurden. Extrinsische kognitive Belastung und Anwendung-Qualität wurden erhoben, während die Studierenden die Lernaufgaben bearbeitet haben. Die Ergebnisse zeigen einen positiven Effekt der Sequenzierung von Modellierungsbeispielen abwechselnd vor den Lernaufgaben auf den kognitiven Fertigkeitserwerb, zumindest für Lernende mit geringerem Vorwissen. Außerdem wird dieser Effekt durch eine höhere Anwendungs-Qualität vermittelt. Extrinsische kognitive Belastung ist dagegen kein signifikanter Mediator, auch wenn die Sequenzierung von Modellierungsbeispielen abwechselnd vor den Problemlöseaufgaben die extrinsische kognitive Belastung im Vergleich zu abwechselnd nachher reduziert hat. Außerdem war der Sequenzierungseffekt nicht abhängig von der Art der Lernaufgabe.

Zusammengefasst bestätigen die Befunde dieser Dissertation die Modelle zum kognitiven Fertigkeitserwerb (z. B. Anderson, 1982; Van Lehn, 1996; Renkl, 2014). Im Einklang mit diesen Modellen bieten beide empirischen Studien dieser Dissertation

Unterstützung für eine Erklärung des positiven Effekts der Sequenzierung von Modellierungsbeispielen abwechselnd vor Lernaufgaben basierend auf dem Erwerb von kognitiven Schemata, die sich in einer besseren Anwendungsqualität zeigen: Die Sequenzierung abwechselnd vorher führt zu höherem Fertigkeitserwerb als geblockt-vorher (Trafton & Reiser, 1993) und als geblockt und abwechselnd nachher (Van Gog et al., 2011), zumindest für Lernende mit geringem Vorwissen (Reisslein et al., 2006). Aus einer kognitiven Fertigkeitserwerbs-Perspektive bedeuten diese Befunde, dass Lernende in frühen Phasen des Fertigkeitserwerbs Beispiele benötigen, die ihnen helfen, kognitive Schemata zu erwerben, die sie beim Problemlösen anwenden und modifizieren können (Renkl, 2014). Eine Erklärung des positiven Effekts Sequenzierung von Modellierungsbeispielen abwechselnd vor Lernaufgaben basierend auf der Theorie der kognitiven Belastung (z.B. Sweller, 2010) scheint auf den ersten Blick ebenfalls passend. Allerdings zeigen die Befunde aus beiden Studien, dass extrinsische Belastung kein signifikanter Mediator ist. Diese Befunde sind im Einklang mit Kritik an der Theorie der kognitiven Belastung wie zum Beispiel von De Jong (2010), der die Validität und Generalisierbarkeit der Theorie in Frage stellt. Die Befunde aus beiden empirischen Studien legen den Schluss nahe, dass Anwendungsqualität als inhaltlicher und qualitativ aussagekräftiger Indikator des Lernprozesses wichtiger ist als extrinsische kognitive Belastung.

Die Erklärungen für einen möglichen positiven Effekt der Sequenzierung von Modellierungsbeispielen nach Lernaufgaben wurden in den beiden Studien nicht unterstützt. Zum einen wurde erwartet, dass die Anregung von Selbsterklärungen und Monitoring (in Studie I) und die Problemlöseaufgaben mit Modell-Vergleich (in Studie II) die Wahrnehmung der eigenen Wissenslücken fördert. Dies wiederum sollte die Integration der nachfolgenden Modellierungsbeispiele mit dem Vorwissen unterstützen (Loibl & Rummel, 2014; Loibl, Roll, & Rummel, 2017). Beide Studien haben jedoch nicht gezeigt, dass ein Effekt der Sequenzierung davon abhängt, ob die Wahrnehmung von Wissenslücken unterstützt wird. Dies könnte unter anderem mit Befunden von Loibl und Rummel (2014) erklärt werden, die gezeigt haben, dass es auch wichtig ist, die Lösungen der Lerner mit der richtigen Lösung gegenüberzustellen und zu vergleichen. Dies wurde in Studie I nicht getan, in Studie II nur teilweise. Das zeigt auch, dass die

Anregung von Selbsterklärungen und Überwachung scheinbar nicht ausreichte, um die Wahrnehmung von Wissenslücken zu fördern. Eine Limitation in diesem Kontext ist jedoch, dass die Prozesse (d.h. Wahrnehmung von Wissenslücken), die durch die Anregung von Selbsterklärungen und Überwachung gefördert werden sollten, in beiden Studien dieser Dissertation nicht gemessen wurden. Das bedeutet, dass keine genauen Aussagen darüber möglich sind, ob die Intervention überhaupt die gewünschte Wirkung erzielt hat. Zukünftige Forschung sollte deshalb zuerst Experimente entwickeln und durchführen, die den vermuteten Mechanismus für einen möglichen positiven Effekt der Sequenzierung nachher zu erfassen und erst dann Experimente durchführen, die diese Prozesse anregen (Bannert, 2009). Als zweite Erklärung für einen möglichen positiven Effekt der nachherigen Sequenzierung wurde die Art der Lernaufgabe vermutet, speziell die Verwendung von Vergleichsaufgaben (Alfieri et al., 2013). Studie II konnte allerdings nicht zeigen, dass ein Effekt der Sequenzierung von der Art der Lernaufgabe abhängt. Dies könnte unter anderem mit Befunden von Nokes-Malach und Kollegen (2013) erklärt werden, die gezeigt haben, dass Vergleichsaufgaben zu besserem weiten Transfer führen als Modellierungsbeispiele abwechselnd vor Problemaufgaben. Weiter Transfer wurde in Studie II jedoch nicht erhoben.

Abschließend wird auf Basis der vorliegenden Befunde dieser Dissertation empfohlen, Modellierungsbeispiele abwechselnd vor Lernaufgaben wie dem Problemlösen zu sequenzieren, weil dadurch die Anwendungsqualität von Prinzipien erhöht wird, was wiederum zu besserem kognitivem Fertigkeitserwerb führt.

Content

1 Introduction

Imagine the following situation: Michael is a teacher and he plans the first session on calculating stochastic probabilities. He knows from research on example-based learning that a combination of modeling examples and problem solving is an effective approach to foster early skill acquisition (e.g., Renkl, 2014; Van Gog & Rummel, 2010). Therefore Michael searches for modeling examples that instantiate the probability-calculations to combine with the problem solving tasks he already had from the last year. As it comes to the more detailed planning, Michael is wondering, how he should sequence the modeling examples and problem solving tasks. He decides to ask a colleague and she suggested two possibilities: either to provide the modeling examples all together (i.e., blocked) before the students start with problem solving or to sequence the modeling examples alternated before problem solving. His colleague argues that both ways should help the learners to apply the rules for calculating stochastic probabilities, but she thinks that the alternated sequence should be even more helpful, because the learners can directly apply what they have learned from the modeling example. In addition to his colleague's suggestion, Michael can also imagine of presenting the modeling examples blocked after problem solving or alternated after problem solving. In Michael's opinion, this could have the advantage that his students experience what they do not know about calculating stochastic probabilities and then are curious to see how it really works.

In case that Michael would search for empirical findings to answer his questions, he would be faced with mixed and inconclusive results (Baggett, 1987; Hsu et al., 2015; Kant et al., 2017; Kapur, 2012; 2014; Leppink et al., 2014; Reisslein et al., 2006; Schwartz et al., 2011; Stegmann et al., 2012; Van der Meij et al., 2017; Van Gog, 2011; Van Gog et al., 2011). While some studies show that sequencing modeling examples before problem solving has a positive influence on cognitive skill acquisition compared to a reverse sequence (Hsu et al., 2015; Kant et al., 2017; Leppink et al., 2014; Van Gog et al., 2011), other studies show a superiority in terms of cognitive skill acquisition for sequencing modeling examples after problem solving instead of before (Baggett, 1987; Kapur, 2012;

Schwartz et al., 2011). Furthermore, there are also some studies that found no significant differences between sequencing modeling examples before or after problem solving (Stegmann et al., 2012; Van Gog, 2011; Van der Meij et al., 2017).

On the one hand, it seems that a superiority of providing modeling examples before problem solving is explained by having the benefit that they reduce detrimental working memory load and foster the application of the learned procedures (e.g., Leppink et al., 2014; Van Gog et al., 2011). On the other hand, when the modeling examples are presented after the leaners have tried to solve problems, they might be more aware what they do not know when they struggle while problem solving and then might better connect the following instruction to their (prior) knowledge. Thus, the benefits of sequencing modeling examples after problem solving might depend on specific conditions, such as supporting the awareness of knowledge gaps (e.g., Loibl, Roll, & Rummel, 2017; Van Gog et al., 2011). Furthermore, studies that show a superiority of sequences modeling examples afterward often used comparison tasks (e.g., Schwartz et al., 2011). Therefore, the effectiveness of sequencing modeling example afterwards might depend on the kind of learning tasks the learners engage in.

However, none of the studies on sequencing fully investigated both potential explanations for the effectiveness of different sequencing strategies: (1) sequences with modeling examples before learning tasks are effective because they reduce extraneous cognitive recourses and allow for application of the mental model (e.g., Kant et al., 2017, Van Gog et al., 2011), and (2) sequences with modeling examples after learning tasks are effective when the awareness of knowledge gaps is supported and when comparison tasks are implemented (e.g., Alfieri et al., 2013; Loibl et al., 2017). What is needed now to solve this problem is not another study that shows a superiority of this or that sequence, but a study that investigates the explanations why this or that sequence is more beneficial for cognitive skill acquisition. Therefore, this dissertation aims at closing these gaps by investigating different sequencing strategies for example-based learning tasks with respect to learning mechanisms (i.e., extraneous load and application quality) that could mediate potential effects of sequencing on cognitive skill acquisition as well as learning conditions that might moderate potential effects of sequencing on cognitive skill acquisition (i.e., supporting the awareness of knowledge gaps with prompts and the type of learning task).

In order to have a profound basis for these empirical investigations, I will first clarify what cognitive skills are and how they are acquired (Chapter 2.1). In Chapter 2.2 I will explain how example-based learning can be used to foster cognitive skill acquisition. Especially, extraneous cognitive load and germane processes will be discussed as mechanisms for the effectiveness of example-based learning. Furthermore, I will elaborate how effective different example-based learning tasks are with respect to triggering these processes as well as with regard to cognitive skill acquisition. Also, the effectiveness of direct prompting germane processes is delineated. After elaborating this theoretical foundation, in Chapter 2.3 I will review the research on different sequencing strategies in example-based learning with respect to cognitive skill acquisition, regarding the cognitive mechanisms (i.e., extraneous cognitive load and application-quality) and with regard to possible interacting effects under different learning conditions (i.e., prompting and type of learning task). In Chapter 3, I will derive two general research questions from the theoretical and empirical foundations presented in Chapter 2, that result in a model for the effectiveness of different sequencing strategies. Chapter 4 provides details for the first empirical study on effects of sequencing and prompting. The second empirical study on effects of sequencing and the type of learning task is described in Chapter 5. In Chapter 6, I will provide answers to the general research questions of both empirical studies. Furthermore, I will discuss the theoretical and practical implications of the two empirical studies and the investigated model. Moreover, general limitations are discussed with respect to the validity of the intervention, the process and outcome measures, and the developed model. Lastly, Chapter 7 provides an outlook where I summarize the main findings of this dissertation and where I draw conclusions for future research.

2 Fostering cognitive skill acquisition with example -based learning

In order to optimally foster cognitive skill acquisition with example-based learning, a precise understanding of how cognitive skills are determined and of how cognitive skills are acquired and develop is important (Chapter 2.1). Furthermore, it is important to identify why example-based learning is an effective approach to foster cognitive skill acquisition. Thus, I will clarify which mechanisms can explain the effectiveness of example-based learning and how those processes can be supported (Chapter 2.2). Lastly, I will discuss how and under which conditions different sequencing strategies in example-based learning can influence those mechanisms as well as cognitive skill acquisition (Chapter 2.3).

2.1 Cognitive skills and their acquisition

In this subsection, I will describe the basis for fostering cognitive skill acquisition that is how cognitive skill can be defined and how cognitive skills are assumed to be acquired.

2.1.1 Cognitive skills

An example which illustrates a cognitive skill can be the description of research designs according to the scheme of Campbell and Stanley (1963) in contrast to a more psycho-motor skill such as playing football. Thus, a cognitive skill is determined by its focus on mental activities compared to physiological activities (Ireson, 2008). Furthermore, I will describe cognitive skills with respect to the (1) structure of cognitive skills, (2) the complexity (3) the problem type, and (5) the problem solving processes.

The (1) *structure* of cognitive skills can be described by production systems (Anderson, 1982). A production system contains a goal-stack that is determined by several productions that can be considered as rules to solve problems. For example, when confronted with the task to describe a research design, the goal-stack first contains the goal to solve this problem. The first production will then lead to a definition of sub-goals to reach that goal. Thus, productions have a hierarchical goal-sub-goal structure in the form of a goal stack. A production consists of a condition that specifies a goal, for example under which circumstances the production is applied (the if-part), and an action that determines what is done when applying the production (the then-part) (Anderson, 1982). An example for a production that determines the goal-stack can be the following task: Please describe a research design according to the scheme of Campbell & Stanley (1963) for this research question: To what extent does the integration of illustrations in a text have a positive effect on knowledge acquisition? As illustration for the beginning of the goal-stack I will describe four productions.

1. If the goal is to solve the task (condition), then describe the design according to the scheme of Campbell and Stanley for this research question (condition)
2. If the goal is to describe the design according to the scheme of Campbell and Stanley for this research question (condition), then identify the independent as well as dependent variable and the numbers of measurements (action).
3. If the goal is to determine the independent variable (condition), then ask for the variable which is assumed to influence something (action).
4. If the goal is to determine the dependent variable (condition), ask for the variable which is assumed to be influenced (action)

Thus, every single production determines which the next production is and also which goals are in the goal-stack. In the presented example, the first production defines that the goal-stack contains one goal overall. The next production then differentiates three sub-goals that are added to the goal-stack. With the productions three and four two sub-goals are reached.

The (2) *complexity* of cognitive skills refers to the number of different constituent skills, which must be coordinated and integrated during performance. This also implies that the task can be more or less complex, depending on the number of subskills that need to be coordinately applied for solving the task (Lim, Reiser, & Olina, 2009; Van Merriënboer, Clark, & De Croock, 2002). Thus, according to this view a cognitive skill is complex, when it consists of many different productions that need to be applied such as for the skill of describing a research design; whereas a cognitive skill is less complex, when there is only a little number of productions such as when identifying a dependent variable. The complexity also depends on the prior knowledge of the learner, which is related to the number of how many subskills are entirely new to learn (Singley & Anderson, 1989). This means, that the learners can regard an objectively seen less complex cognitive skill still as complex, when they have little prior knowledge about the cognitive skill.

The (3) *problem type* can be differentiated with respect to well- and ill-structured problems (Jonassen, 1997; Schraw, Dunkle, & Bendixen, 1995). Well-structured problems have a clear solution that can be classified as (in-)correct as well as a prescribed solution procedure, that contains a limited number of productions. Furthermore, the solution procedure is comprehensible, meaning that the relations between the solution strategies and elements of the problem description are known (Wood, 1983). For example, describing a research design according to the scheme of Campbell and Stanley (1963) can be classified as well-structured problem type, whereas writing of a research paper might be regarded as ill-structured problem. Ill-structured problems can have multiple possible solutions solution procedures or no correct solution at all, but rather more or less correct solutions (Kitchner, 1983).

The (4) *problem solving processes* are assumed to consist of the construction of a problem space and searching in this space for problem solving operators (Gick, 1986). During the construction of a problem space, the learners are assumed to understand and define the problem with respect to five features: (a) beginning state, (b) goal state, (c) different problem states, (d) the constraints for solving the problem, and (e) problem solving operators to transform the beginning state to the goal state (Newell & Simon, 1972; Simon, 1999). To illustrate these features more concretely, I detail a problem representation for describing research designs according to the scheme of Campbell and Stanley (1963). The beginning state (a) is determined by research questions that need to be transformed to a design according to the Campbell and Stanley scheme, which represents the goal state (b).

Different problem states (c) are for instance to identify the type of design, to determine the independent variable(s), to determine the dependent variable(s) and so on. Constraints for solving the problem (d) are that the learner only has a certain amount of time (e.g., ten minutes) to solve the problem. After the problem is represented, learners search the problem space which means that they try to transform the different (problem) states with the help of problem solving operators (e) (Newell & Simon, 1972; Simon, 1999). Problem solving operators are typically used as synonym for productions (Anderson, 1982). Although the description of the problem solving processes and the structure of cognitive skill appear quite similar, for example problem states could be seen as sub-goals in the goal-stack, the former vied adds information with regard to the search for and the selection of problem solving operators (i.e., productions). The quality of the problem representation can influence the selection of productions (Gick, 1986). When the problem representation activates an already existing cognitive schema, a *schema-driven strategy* is implemented. Thereby, a schema is regarded as a cluster of knowledge relevant for a specific problem type, like goal and solution procedures for that type of problem (Gick & Holyoak, 1980). When a schema-driven strategy is applied, a solution procedure is directly implemented, thus no search in the problem space is needed (Gick, 1986). For example, a learner who has a cognitive schema for describing designs with the Campbell and Stanley scheme can infer will directly implement the solution without selecting further productions that define sub-goals and lead to additional productions. When the problem representation does not directly lead to schema activation, *search-based problem solving strategies* need to be applied. Typically, this is either done by means-end-analysis or through search by analogy (Gick, 1986). During means-end-analysis, first the difference between the actual problem state and the goal state is determined. Then, a sub-goal is build which reduces that difference. In a third step, an operator is applied to reach that sub-goal; this implies that the means for achieving a (sub)goal is the (sub)goal itself (Gick, 1986). This is regarded as a less useful strategy, because many sub-goals need to be active in working memory. This induces high working memory load, which interferes with schema-acquisition (Renkl, 2014). During search by analogy, an example is retrieved that can guide the selection of operators, which is considered a useful strategy for the acquisition of cognitive skills (e.g., Anderson et al., 1981). Thus, learners can apply search by analogy when they know examples for a specific type of problem, whereas learners without knowledge about analog examples have to rely on means-end-analysis (Gick, 1986). From this perspective, presenting examples before problem solving is favorable, because it can lead to a more effective problem solving strategy via search by analogy.

To this end, I have delineated how cognitive skills can be described with respect to their structure and complexity as well as the problem type and how cognitive skills are used during problem solving, but not how they are acquired and develop. Therefore, the next section will provide a detailed insight to the acquisition of cognitive skills.

2.1.2 Cognitive skill acquisition

There are three established models in cognitive skill acquisition research, which build upon each other: Anderson's (1982, 1983, 1993) ACT(-R) (i.e., adaptive control of thought - rationale) model, Van Lehn's (1996) model on phases of skill acquisition, and Renkl's (2014) model on phases of skill acquisition within his integrative theory of example-based learning.

Anderson's ACT model on cognitive skill acquisition (1982) is based on Fits (1964) stages for the development of perceptual-motor skills (cognitive, associative and autonomous stage), but applied on cognitive skills. Anderson (1982) differentiates three stages of cognitive skill acquisition: (1) declarative stage, (2) knowledge compilation stage, and (3) procedural stage. Within the *declarative stage*, instructions and information about the skill are encoded and represented in a declarative form. The encoded facts need to be rehearsed in working memory while executing the skill, which has the disadvantage of relative high costs with respect to time and working memory capacity. The facts "are used interpretively by general-purpose productions" (Anderson, 1982, p. 374), like searching by analogy or means-ends analysis. In this stage, examples play an important role because they can trigger 'search by analogy' referred to by Gick (1986) as *search-based problem solving strategies* and by Anderson (1982) as *general-purpose productions*.

Accordingly, instruction with examples should be provided to learners before problem solving. In the *knowledge compilation stage*, the declarative knowledge is converted into procedural knowledge in a gradual process as a result of practice. Characteristic for this stage is speedup and that strategies such as piece-by-piece matching and verbal rehearsal are no longer needed. This also means that the speed by which the declarative knowledge is accessed and used becomes faster. During knowledge compilation, two sub-processes occur: composition and proceduralization. Composition means that sequences of productions are integrated into one production, which leads to higher speed and unitary application. This means that for example the two productions for identifying the dependent and independent variable are compiled to only one production that needs to be applied. Proceduralization means that no declarative knowledge is involved in the production, but only the product itself; for example the production to identify the dependent and independent variable would contain the product (i.e., illustrations and knowledge acquisition) instead of the declarative knowledge about the rule how to identify the (in-)dependent variable. As a result the declarative knowledge does not have to be active in working memory, which frees up working memory capacity; for which the elimination of the verbal rehearsal is an indicator. The compilation depends on the goal structure that is generated during problem solving; this is indicated by the order of the productions that are chunked in the new production. The *procedural stage* (3) is marked by automatization and further speedup (Anderson, 1982).

ACT was further developed to *ACT-R* (Anderson, 1993). The major changes concern the relation between declarative and procedural knowledge and the source of declarative knowledge (Anderson & Fincham, 1994). In ACT, declarative knowledge is regarded as a prerequisite of procedural knowledge, whereas in ACT-R, declarative knowledge about the example must not be stored in long term memory, but it needs to be active in working memory during the process of analogy by the learner. This would also be the case, when the learners have access to learning resources with examples, because the declarative knowledge about the example will be active in working memory when they process the example. Second, it is assumed that the declarative origins of procedural knowledge come from examples instead of instruction, which has the advantage that the declarative knowledge is connected with the analogy instead of disconnected from its application. This reconceptualization also concerns the phases of skill acquisition. ACT-R distinguishes four phases of cognitive skill acquisition: (1) analogy, (2) forming declarative rules, (3) forming production rules within knowledge compilation, and (4) retrieval of specific examples. During the phase of analogy, example features are mapped and applied to problem features. Next, declarative rules are abstracted from the interpretive problem solving by analogy. With further practice, the declarative rules and analogies are compiled to productions, where the procedural rules are automated. In phase four, the solution to specific problems can be directly retrieved due to a large body of examples in memory (Anderson, 1993; Anderson, Fincham, & Douglass, 1997). Thus, examples are especially important in the first two phases and should be provided before problem solving.

In *Van Lehn's (1996) model*, which is also based on the model of Fitts (1964), an early, intermediate and late phase of cognitive skill acquisition are distinguished. The first phase consists of activities to acquire declarative information like reading a textbook, discussing the information, listening to a lecture and so on (Van Lehn, 1996), which is comparable to the first phase in ACT. In ACT-R, there is no such phase with pure declarative information independent from examples. During the intermediate phase, four processes are assumed to take place: (1) retrieval, (2) mapping, (3) application, and (4) generalization. This phase is similar to both ACT-R's phases of analogy and declarative rules. Retrieval of an example or principle can be spontaneous or deliberate, although deliberate retrieval is more effective for retrieval of an example or principle (e.g., Gick & Holyoak, 1980), because spontaneous retrieval is likely guided by surface features of the example and problem (e.g., Catrambone & Holyoak, 1989) and deliberate retrieval by structural similarities (e.g., Faries & Reiser, 1988). During mapping the learners relate features of the problem with features of the example or principle. Next, the principle is applied during problem solving. Reduction of errors during problem solving should improve cognitive skill acquisition, when the goal is to learn the application of certain rules to solve specific problems (e.g., Anderson, Corbett, Koedinger, & Pelletier, 1995; Renkl & Atkinson, 2003). Subsequently, a generalized schema is acquired for the principle, in the sense that surface features will no longer mislead retrieval, mapping and application. As in ACT-R, this last phase is characterized by automatization through practice effects by the power law of practice (Van Lehn, 1996).

Renkl's (2014) model on cognitive skill acquisition within his integrative theory of example-based learning differentiates four phases: (1) principle encoding, (2) relying on analogs, (3) forming declarative rules, and (4) fine tuning. In the first phase of principle encoding, declarative knowledge about domain principles is acquired, which is mainly encoded as fact and not yet as schemata. This phase is similar to the first phase in ACT and the first phase in Van Lehn's model, and to the second phase in ACT-R. In the second phase, the learner begins to solve problems. The problem solving processes highly rely on analogies, which ideally are encoded and connected with the underlying principles. This in turn usually fosters the formation of abstract schemata (Renkl, 2014). In phase three, declarative rules are formed. The learner verbalizes specific if-then-rules for solving a problem. Application conditions for those rules are encoded as schemata. In the fourth phase, procedural rules are formed by chunking the single subskills. This leads to automatization and speedup, and frees up working memory capacity. Using the skill in different contexts fosters flexibility in application of the skill. In this phase, example-based learning is not that relevant, especially compared to phase one and two, because the goal is a fast, automated and flexible use of the skill.

These three different models provide differing detailed insights into the assumed stages of skill acquisition. Table 1 illustrates which stages are similar and which ones are unique. For example, all models share that in phase I and/or II the encoded principles need to be related to analogies, and the analogies need to be connected to problem solving to acquire abstract schemata. Thus, examples play a major role in the early phases of cognitive skill acquisition, because they provide analogies that are mapped and applied to the problem features and then generalized as schemata. In the late phase(s), examples play a minor role, because the focus is on practice and not on building abstract schemata with support of examples (Anderson, 1982; Anderson et al., 1997; Renkl, 2014; Van Lehn, 1996). Interestingly, the right time for problem solving differs between the models. In

ACT(-R), problem solving starts in stage I (Anderson, 1982; 1983), whereas Van Lehn (1996) and Renkl (2014) separate an example-phase (stage I) from the second stage in which problem solving also plays a role. However, in all models there is no clear separation of the different stages of cognitive skill acquisition, but rather the assumption of a smooth transition (Anderson et al., 1997; Renkl, 2014; Van Lehn, 1996).

Taken together, examples and problem solving are equally important for early cognitive skill acquisition. It is unclear, however, in which sequences examples and problem solving should be provided, to optimally support cognitive skill acquisition. From the models it can be derived that examples should be provided before problem solving, although it is unclear whether they should be provided as a block of examples before problem solving or alternated with problem solving. Nevertheless, there are also arguments for providing examples after problem solving, which will be discussed in Chapter 2.3. To approach the matter of sequencing properly, it will first be elaborated in Chapter 2.2, when and how example-based learning is effective.

Table 1. *Comparison of different models on cognitive skill acquisition*

	Anderon's ACT model	**Anderon's ACT-R model**	**Van Lehn's model**	**Renkl's model**
Phase I	Declarative phase: encoding declarative information and interpretative problem solving	Analogy: encoding, mapping and application of example features	Early phase: encoding declarative information	Principle encoding: encoding of declarative information
Phase II	Knowledge compilation stage: composition and proceduralization	Forming declarative rules: abstraction of rules from interpretative problem solving by analogy	Intermediate phase: retrieval, mapping, application, generalization	Relying on analogs: interpretative problem solving by analogy
Phase III	Procedural stage: automatization	Forming production rules: composition, proceduralization and automatization	Late phase: automatization	Forming declarative rules: abstraction of rules from interpretative problem solving by analogy
Phase IV		Retrieval of specific examples: direct application of solution		Fine tuning: composition, proceduralization and automatization

2.2 Example-based learning of cognitive skills

Example-based learning (EBL) has been widely investigated over the past two decades and it is an instructional approach to support the early acquisition of cognitive skills (for reviews see Renkl, 2014; Van Gog & Rummel, 2010). In order to apply EBL, it is important to know the underlying mechanisms for the effectiveness of EBL. The following two subsections will first elaborate on explanations as to why EBL can foster cognitive skill acquisition. These explanations rely on the construct of cognitive load and germane processes (Renkl, 2014). The implementation of EBL requires diverse learning tasks to trigger different processes. Therefore, the third subsection will clarify which learning tasks are typically used in EBL and how effective they are in fostering the relevant processes and performance. The final subsection will discuss how germane processes in EBL can be supported with prompts.

2.2.1 Cognitive load as mechanism for the effectiveness of example-based learning

Sweller and Cooper (1985) argue that worked examples support the acquisition of schemata about the relation of problem states, operators and goal states in early phases of skill acquisition, whereas (too) early problem solving attempts often focus on search processes to reduce the differences between problem and goal states. Problem solving with means-ends analysis usually requires more working memory capacity than studying worked examples, because many sub goals need to be active in working memory, which leads to high cognitive load (Sweller & Cooper, 1985; Sweller, 1988). Due to the high working memory demands on the one hand and the limited working memory capacity on the other hand, schema construction – which also needs working memory resources – is hindered (Sweller, Van Merriënboer, & Paas, 1998). Based on the assumption of limited working memory capacity, cognitive load theory distinguishes three types of cognitive load, which all require a portion of the available working memory capacity: intrinsic cognitive load, extraneous cognitive load, and germane cognitive load (Sweller & Cooper, 1985; Sweller et al., 1998).

Intrinsic cognitive load is influenced by the complexity of the learning material in the sense of element interactivity and by the learner's prior knowledge. Element interactivity refers to "the number of elements that must be processed simultaneously in working memory" (Sweller et al., 1998, p. 259). Intrinsic cognitive load is also assumed to not directly contribute to learning, and it can also not directly be manipulated (Brünken et al., 2010; Sweller et al., 1998). For example, learning isolated facts like Latin vocabulary has low element interactivity and therefore a low intrinsic cognitive load, relative to the learner's prior knowledge in Latin. Learning complex procedures like applying Latin grammar and vocabulary to translate a Latin text into German has high element interactivity. *Extraneous cognitive load* refers to the amount of load imposed by the instructional design of the learning material. This means, everything that is not directly relevant for learning interferes with schema construction and should be minimized (Brünken, Moreno, & Plass, 2010; Sweller, 2010). Extraneous cognitive load can occur, for instance, when a table or graph on one page needs to be related to a connected text on another page. Several studies have shown that providing worked examples compared to problem solving leads to a reduction in extraneous cognitive load (e.g., Renkl, Gruber, Weber, Lerche, & Schweizer, 2003; Van Gog et al., 2006). *Germane cognitive load* is considered as the learning relevant load which is needed to understand the learning material and to construct schemata. This load should be maximized (Brünken et al., 2010; Sweller et al., 1998). However, this type of load has received critique over the past decade and its existence is debated, as well as its measurement (De Jong, 2010). For example, De Jong (2010) criticizes conceptual problems concerning the distinction between intrinsic and germane cognitive load, because encoding and understanding the learning material, which is a process dependent on element interactivity, and therefore intrinsic cognitive load, also involves processes that contribute to germane cognitive load.

Mental effort is often used as a synonym for (extraneous) cognitive load (Kirschner & Kirschner, 2012). Paas and colleagues (2003, p. 64) define mental effort as "the aspect of cognitive load that refers to the cognitive capacity that is actually allocated to accommodate the demands imposed by the task; thus, it can be considered to reflect the actual cognitive load”. Mental effort is typically measured with the self-report rating-item of Paas (1992). This mental effort rating scale is widely used in EBL-research due to its several advantages (Paas, Touvinen, Tabbers, & Van Gerven, 2003). Paas and colleagues (2003) argue that this measurement is sensitive, reliable, valid, inexpensive, easy to use, and does not interfere with the learning task performance. With respect to sensitivity, Van Gog, Kirschner, Kester and Paas (2012) showed that a frequent and repeated measurement of mental effort after each task should be preferred to an overall measurement of mental effort at the end of learning phase due to a more precise and sensitive measurement.

However, cognitive load is critically approached by authors such as De Jong (2010). One major criticism denotes the post-hoc explanations in terms of the three types of cognitive load, because these are impossible to falsify. In these cases, authors (e.g., Mayer, Hegarty, Mayer, & Campbell, 2005) derive conclusions about cognitive load from posttest scores. When learners performed well, germane load was maximized by the instructional design or intervention, but when learners performed poor, extraneous load was not reduced (or even increased) by the instructional design or intervention. Likewise, mental effort measures are also susceptible for post-hoc explanations due to their unidimensional structure: when mental effort is negatively related to performance, it is explained as extraneous cognitive load, but when it is positively related to performance it is explained as germane cognitive load (De Jong, 2010). Furthermore, many studies that measure mental effort calculate a new score to relate mental effort to test performance: mental efficiency, which is based on mental effort invested in the learning phase in relation to final test performance. Even though Van Gog and Paas (2008) are aware of the difficulties in interpreting mental effort during learning, they still recommend the mental efficiency score when studies aim to reduce extraneous cognitive load during learning. They claim that when studies investigate an instructional format that aims to reduce extraneous cognitive load and are successful in showing an improvement in performance, these studies can conclude that mental effort represents extraneous cognitive load (Van Gog & Paas, 2008). The question is, however, if it is more suitable to determine a reduction in extraneous cognitive load by investigating mental efforts as a (negative) predictor for test performance and/or as a mediating variable instead of calculating mental efficiency. One advantage of a regression analysis lies in comparing the predictive value of mental effort with other qualitative learning process measures. This is not possible for mental efficiency because performance and mental effort are combined in a new score and are no longer separated measures. This also results in questionable validity of the mental efficiency score.

To overcome unidimensional measurements of mental effort, multidimensional scales to measure the three different types of cognitive load were developed and tested (e.g., Leppink et al., 2013; 2014; Opfermann, 2008). For example, Leppink, Paas, Van der Vleuten, Van Gog and Van Merriënboer (2013) found a three factorial solution with the factors intrinsic, germane and extraneous cognitive load, but they found no significant correlations of either type of load with performance. Leppink, Paas, Van Gog, Van der Vleuten and Van Merriënboer (2014) replicated the three-dimensional solution with intrinsic, germane and extraneous cognitive load. They argue, however, that it is not clear whether the three factors actually represent the three types of load, because they only found significant negative correlations of intrinsic and extraneous cognitive load with performance, but no significant correlation of germane cognitive load and performance. In a follow-up study, they found no significant correlations between either type of cognitive load and performance. A problem of the multidimensional scale, however, can lie in the subjectivity of the scales, that is, the validity of the scales is compromised when the learner cannot differentiate between the three scales and/or when the learner's interpretation differs from the intended one. Leppink and colleagues (2014) concluded that their findings provide support for a recent reconceptualization by Sweller, Ayres and Kalyuga (2011). In this reconceptualization only intrinsic and extraneous cognitive load are distinguished, where intrinsic cognitive load should be maximized by applying germane processes such as elaboration and extraneous cognitive load should be minimized.

Taken together, extraneous cognitive load can be reduced by providing examples before problem solving, because no means-end-analysis – which imposes high demands on working memory – is required. This frees up working memory resources that can be allocated to germane cognitive load instead. However, due to the contested nature of germane cognitive load it is more accurate to focus instead on germane processes, such as elaboration and monitoring, as mechanism for the effectiveness of EBL.

2.2.2 Germane processes as mechanisms for the effectiveness of example-based learning

Germane processes are regarded as highly important for schema construction (Paas & Sweller, 2014). Depending on the learning goal, germane processes can be differentiated in cognitive processes such as organization and elaboration, and metacognitive processes like monitoring and regulation (Renkl & Atkinson, 2003). Research on EBL mainly derives from three perspectives: observational learning, learning with worked examples, and analogical reasoning (i.e., contrasting cases). Dependent on the perspective, different (meta-) cognitive processes are assumed to take place (Van Gog & Rummel, 2010; Renkl, 2014). In order to better understand the processes, the three EBL-perspectives will be briefly explained.

Observational learning derives from Bandura's (1986) social-cognitive learning theory. Several studies have shown positive effects on academic skills for observational learning (e.g., Schunk, 1981; Zimmerman & Kitsantas, 2002; Rummel, Spada, & Hauser, 2009). In observational learning, a model demonstrates a procedure on how to solve a specific (set of) problem(s). Bandura (1971, 1986) assumes two phases in observational learning, an acquisition and a performance phase, with two central processes in each phase. During the acquisition phase, the learner has to pay attention to the model and construct a cognitive schema of the modeled behavior. In the performance phase, the learner needs to be able to reproduce the behavior and has to be motivated to show the behavior. From this perspective, monitored enactment is an important mechanism for translating thoughts into action (Bandura, 1991; Carroll & Bandura, 1987). Monitored enactment can be regarded as a concept-matching process, where conceptions (i.e., schemata) serve as cognitive guidance for problem solving and as internal standards for comparison, detection and correction of discrepancies between conceptions and behavior. Therefore, a combination of problem-solving and modeling examples as pairs seems helpful, so that the learner can compare his/her performance with the model and then focus attention on critical aspects in the following modeling examples (Bandura, 1991). Also, from a broader perspective, monitoring is an important mechanism in learning. Hartwig and colleagues (2012) showed that general monitoring accuracy significantly predicted exam performance. Similarly, Dunlosky and Rawson (2012) showed the importance of monitoring accuracy for learning and retention. More specifically, their study revealed that inaccurate monitoring in the sense of overconfidence undermines learning achievement.

During *learning with worked examples*, learners study a worked example, which consists of a problem statement, the solution steps and the solution (e.g., Renkl, 2014, Van Gog & Rummel, 2010). In many studies, learning with worked examples has been shown to be more effective to foster cognitive skill acquisition in early stages of learning compared to learning with problem solving tasks (e.g., Sweller & Cooper, 1985; Paas, 1992; for an overview see Van Gog & Rummel, 2010; Renkl, 2014). Important elaboration processes when learning with worked examples are self-explanations (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Renkl, 2014). Chi and colleagues (1989) compared successful and poor leaners in a study on learning physics with worked examples. They revealed that good learners showed more self-explanations of single solution steps and underlying principles in the worked examples than poor learners. This phenomenon was called the self-explanation effect. However, learning time was not controlled by Chi and colleagues (1989) and could have confounded self-explanation quality. Renkl (1997) replicated Chi and colleagues' study with example-based learning of probability calculation, where all learners had the same amount of learning time, showing the self-explanation effect even when controlling for learning time. Moreover, Renkl (1997) identified two types of effective self-explanations that were both positively related to performance: anticipative and principle-based self-explanations. Anticipative self-explanations take place when the learner anticipates the next solution step(s). Principle-based self-explanations occur when learners self-explain the rationale behind the example by connecting operators to domain-principles or by explaining goal–operator combinations. Apart from self-explanations, metacognitive skills are also important for analogical problem solving in example-based learning (Muldner & Conati, 2010). Those meta-cognitive skills concern a min-analogy use of worked examples. Min-analogy means trying to solve a problem oneself without (or limited) copying from the example. Thereby, the learners only refer to an example when they monitor impasses or uncertainties (Van Lehn, 1998; Van Lehn & Jones, 1993).

During *analogical reasoning* the learners infer knowledge from examples and rely on this knowledge while problem solving (Gentner, 2003; Holyoak, 2005). Four stages are assumed to take place during analogical reasoning (Holyoak, 2005, 2012; Gentner, 2003): (1) provided examples (and principles) are encoded, sometimes already as a schema; (2) when solving a transfer problem, analogical examples are activated and selected; (3) features of the problem are mapped to the example; and (4) a schema is inferred or modified from this mapping process. Mapping is the core process of analogical reasoning, because concrete and abstract principles need to be interrelated for schema construction to occur (Renkl, 2014). Mapping processes usually consist of an effortful search for commonalities and an alignment of target features. More specifically, effortful search for commonalities means that either two examples or an example and problem are compared with respect to common features. Target feature alignment refers to the interrelation of concrete instances and abstract principles of two examples or an example and a problem (Alfieri et al., 2013).

Taken together, all three EBL-perspectives share the importance of EBL for cognitive skill acquisition in early stages of learning, because it can reduce extraneous cognitive load and thereby free up working memory resources for germane processes, which allow for the construction of schemata (Renkl, 2014). Whereas germane processes in learning with worked examples (self-explanation) and analogical reasoning (mapping) are mainly characterized by elaboration processes, germane processes in observational learning (monitored enactment) rely on both elaboration and metacognitive processes. The assumed processes in the three EBL-perspectives also fit well to the stages of cognitive skill acquisition, especially the first two stages in Anderson's (1993) ACT-R, Van Lehn's (1996) and Renkl's (2014) models, in which principles are encoded and the learner relies on analogies. The key processes in analogical reasoning (Gentner, 2003) are principally the same as in Van Lehn's (1996) intermediate phase, i.e. retrieval, mapping, application and generalization. Within the different EBL-perspectives, diverse learning tasks can be used or combined with examples that should trigger different germane processes. The next section will elaborate on the learning tasks that are typically used in the three EBL-perspectives and how effective these tasks are in fostering the relevant processes and performance.

2.2.3 Effectiveness of different example-based learning tasks

According to Doyle (1983), learning tasks can be characterized by at least three features. Learning tasks contain (1) the creation of products, (2) the application of operations to create the product, and (3) resources that influence working on the learning task. Some authors regard worked examples themselves as learning tasks (e.g., Germ, 2009; Schabram, 2007). Following Doyle (1983), this thesis concerns worked examples (only) not as learning tasks because no product is created by the learner and no operators are applied, but as resources that influence working on the learning task.

In line with their slightly different foci, the three EBL-perspectives use different learning tasks. In research on observational learning, modeling examples are usually combined with problem solving tasks (e.g., Rummel & Spada, 2005). Research on learning with worked examples typically uses worked examples combined with problem solving-tasks (e.g., Van Gog & Kester, 2012) and/or self-explanation of examples (e.g., Schworm & Renkl, 2007), and learning tasks shaped as incomplete worked examples (e.g., Stark, 2000), faded worked examples (e.g., Atkinson et al., 2003), or erroneous worked examples (e.g., Große & Renkl, 2007). In analogical reasoning research, contrasting cases and comparing examples are often used as learning tasks, either provided as invention task or not (e.g., Rittle-Johnson & Star, 2007).

Regardless of the EBL-perspective, all of these learning tasks can activate germane processes – albeit different germane processes. From an observational learning perspective, a combination of modeling examples and problem solving tasks should trigger monitored enactment, i.e. elaboration and monitoring processes (see Chapter 2.2.2). Monitored enactment should in turn affect attentional processes when receiving the next modeling example (Bandura, 1991). Within the perspective of learning with worked examples, a combination of worked examples and problem solving tasks should help learners to apply abstract principles on concrete problems (Renkl, 2014) and also motivate learners to process the worked example more deeply – especially when they have to solve a problem afterwards (Sweller & Cooper, 1985). Incomplete, faded or erroneous worked examples have gaps or errors that need to be identified, completed, corrected or inferred by the learners. This should activate anticipative self-explanation processes and also help learners to apply abstract principles on concrete problems (Renkl, 2014). From a cognitive load perspective, a combination of modeling examples with comparing of examples as learning tasks could lower extraneous cognitive load as opposed to a combination of modeling examples with problem solving tasks, because no search processes in the problem space are needed to solve this kind of task, which should reduce working memory demands (e.g., Sweller, 1988). From an analogical reasoning perspective, contrasting cases or comparing examples provide opportunities for mapping processes, which should foster the interrelation of abstract principles with concrete instances (Alfieri et al., 2013).

Research on the effectiveness of different types of learning tasks from different EBL-perspectives is quite rare. Most studies on different learning tasks have been conducted from the learning with worked examples perspective. For instance, Renkl and Atkinson (2003) compared example-problem pairs to fading of worked examples. In their study, fading means that a complete example was provided first, followed by a second example where the learners had to solve one step by themselves. The steps the learners had to solve increased from example to example so that the leaners had to solve a complete problem in the end. Renkl and Atkinson (2003) present evidence that fading of worked examples reduced errors during learning – i.e., higher application quality – and thus resulted in better near transfer. However, Reisslein and colleagues (2006) showed no superiority of fading of worked examples compared to example-problem pairs. Closer scrutiny suggests that a limitation of the practicability of fading might lie in the number of steps that need to be faded. As it turns out, the (faded) examples in both studies (Renkl & Atkinson, 2003; Reisslein et al., 2003) consisted of only three steps. When more steps are required to solve a certain type of problem, fading of worked examples could require much more time than alternating examples and problem solving; and thus does not appear very practical to integrate in lessons. An argument in favor for the combination of worked examples with problem solving tasks is the study by Baars and colleagues (2017). They showed that problem solving after worked examples led to better monitoring accuracy than learning with worked examples only. Stark and colleagues (2000) also showed the benefits of the combination—albeit for the reverse order, that is, the combination of problem solving before a worked example, significantly led to more monitoring and regulation compared to learning with worked examples only. Finally, Shebilske and colleagues (2006) studied whether the combination of practice sessions (i.e., problem solving) and a demonstration in the form of modeling examples fosters the skills to work with the Airborne Early Warning and Control System compared to practice sessions only. Their study showed that an alternated-after sequence of providing modeling examples after problem solving significantly led to a higher degree of cognitive skill acquisition compared to problem solving only.

Nokes-Malach and colleagues (2013) compared learning tasks from the worked examples perspective and analogical reasoning perspective with respect to transfer performance. They investigated whether (a) analogical comparison of worked examples, (b) self-explanation of worked examples, and (c) worked examples-problem pairs influenced near, intermediate and far transfer. The measurement of near transfer was interleaved with three problem solving tasks throughout the learning phase. Intermediate and far transfer was measured in a test phase directly after the learning phase. Nokes-Malach and colleagues (2013) showed that self-explanation of worked examples as wells as example-problem pairs lead to better near transfer than comparing examples. Moreover, although all groups performed similarly on intermediate transfer, the self-explanation and comparison groups outperformed the example-problem pairs on far transfer. With respect to a comparison of learning tasks from the observational learning perspective and worked examples and/or analogical reasoning perspectives, systematic research is missing. Similarly, thus far research has not investigated what type of learning task (from what EBL-perspective) should be combined with examples.

Taken together, modeling examples can be effectively combined with problem solving to reduce extraneous cognitive load and support application quality. Also, a combination with self-explanation can support elaboration of the example. However, there is little to no research as to which combination is better for cognitive skills acquisition, i.e. whether modeling examples with problem solving should be preferred or whether modeling examples should rather be combined with comparing of examples. Besides the encouragement of germane processes (i.e. application, elaboration and monitoring processes) by the type of learning task, these germane processes can also directly be supported by integrating prompts in the learning tasks. Thus, the next section will explain how the germane processes such as application, elaboration and monitoring processes can directly be supported by prompting.

2.2.4 Prompting germane processes in example-based learning

Prompts can be an effective way to reduce passivity and activate learners to engage in germane processes in example-based learning (Renkl, 2014). Prompts are scaffolds that direct the learner’s attention to relevant aspects of the learning task (Quintana et al., 2004) or worked example (Renkl, 2014, Van Gog & Rummel, 2010) and act as support for activities the learner is capable of but does not show automatically (Pressley et al., 1992). In the context of example-based learning, these activities mainly include the germane processes of elaboration (self-explanation and comparing) and monitoring. This section will delineate how self-explanation and comparing can be supported by cognitive prompts, how monitoring processes can be supported by metacognitive prompts, and whether and how a combination of cognitive and metacognitive prompts can be fruitful for early cognitive skill acquisition.

Self-explanation prompts aim to activate principle-based and anticipative self-explanations (Renkl, 2014). Two meta-analyses showed small to moderate effects of self-explanation prompts on transfer in structured domains like math and science compared to no self-explanation prompts (Durkin, 2011; Mugford, Corey, Bennell, & Martens, 2009). Self-explanation prompts have also shown to be effective for learning in business apprenticeship (Stark, 1999), biology (Chi, De Leeuw, Chiu, & LaVancher, 1994) as well as in less structured and more complex domains like argumentation (Schworm & Renkl, 2007). However, Berthold and colleagues (2011) showed double-edged effects of conceptually-oriented self-explanation prompts compared to no prompts: self-explanation prompts had a positive effect on conceptual knowledge of tax law, but a negative effect on procedural knowledge. The authors assumed that these double-edged effects of prompting may have been due to the complexity of their learning task on tax law in the sense of high element-interactivity (i.e., Sweller, 2006), because it might have induced high working memory load with respect to intrinsic cognitive load (Berthold, Röder, Knörzer, Kessler, & Renkl, 2011). Furthermore, several studies showed no effects of self-explanation prompts like in the studies by Hausmann and Chi (2002), Mwangi and Sweller (1998, Experiment 3), and Conati and Van Lehn (2000). Huang (2007) hypothesized that the mixed findings might be explained by the more general type of prompts used in these studies and that these prompts might not have been specific enough compared to studies that reported effects of prompting.

When comparing examples different objectives can be prompted to stimulate the learner to identify commonalities, differences or both (Alfieri et al., 2013). In a meta-analysis on case comparison, Alfieri and colleagues (2013) explored that the type of the comparison that was prompted (i.e. searching for commonalities, differences or both) moderates the effectiveness of case comparisons. More specifically, comparison prompts that stimulate learner to identify commonalities were superior with respect to learning outcomes over prompts that stimulated identifying both differences and commonalities and, this in turn was more effective than prompts to stimulate identifying differences only. The authors argue that searching for similarities could be more helpful to focus on the critical features of the examples as opposed to surface features that could be more likely induced by prompts that stimulate finding differences (Alfieri et al., 2013).

Prompting monitoring processes is best approached from the broader context of prompting metacognitive processes. Stark and colleagues (2008) investigated the effect of metacognitive prompting within example-based learning on knowledge acquisition in statistics education. They prompted students to provide reasons for their learning behavior to induce a mindful use of the learning environment. They showed that metacognitive prompting significantly and sustainably led to more knowledge acquisition than no prompts (Stark, Tyroller, Krause, & Mandl, 2008). In the narrower context of prompting monitoring processes within example-based learning research is quite limited. One exception is the study that by Carroll and Bandura (1987). They explored the role of concurrent matching (i.e. possibility of concurrent comparison) and of visual monitoring (i.e., prompting monitoring processes) in observational learning of motor skills. They showed that concurrent matching as well as visual monitoring led to higher reproduction accuracy. However, they also revealed an interaction of concurrent matching and visual monitoring: concurrent matching was more effective than separate matching for those learners who could not monitor their performance, but no superiority was shown for leaners that monitored their behavior.

With respect to a combination of cognitive (e.g., elaboration) and metacognitive (e.g., monitoring) prompts in the context of example-based learning, systematic research is missing. In the broader context of research on self-regulated learning several studies showed that a combination of cognitive and metacognitive prompts is more effective for learning than no prompting (e.g., Berthold, Nückles, & Renkl, 2007, Wichmann & Leutner, 2009). For example, Wichmann and Leutner (2009) showed that combining cognitive and metacognitive prompts within a simulation-based inquiry learning environment led to better learning compared to cognitive or metacognitive prompts only or no prompts at all. Similarly, Berthold, Nückles and Renkl (2007) revealed that cognitive as well as a combination of cognitive and metacognitive prompts for writing learning protocols led to better comprehension and more retention compared to metacognitive prompts only or no prompts at all. The cognitive prompts addressed the use of organization and elaboration strategies, whereas metacognitive prompts concerned monitoring strategies. A follow-up study by Nückles, Hübner and Renkl (2009) expanded the prompt-types by adding prompts for planning for remedial strategies. They investigated five conditions: (1) no prompts, (2) cognitive prompts only, (3) metacognitive prompts only with planning for remedial strategies, (4) a combination of cognitive and metacognitive prompts with planning for remedial strategies, and (5) a combination of cognitive and metacognitive prompts without planning for remedial strategies. Their study showed that all types and combinations of prompts led to better understanding and retention compared to no prompts. Moreover, the condition with a combination of cognitive and metacognitive prompts with planning for remedial strategies significantly outperformed all other conditions. Recently, Roelle, Nowitzki and Berthold (2017) further investigated the effects of the order in which cognitive and metacognitive processes are prompted when writing learning protocols. They showed in two experiments that the order affects the quality of the (meta-)cognitive processes and the learning outcomes: learners who first received metacognitive prompts showed a higher quality of metacognitive and organization processes as well as higher learning outcomes compared to learners who first received cognitive prompts.

To sum up, prompting principle-based self-explanations as well as comparisons processes share the same function, namely relating concrete examples to abstract principles, which is helpful for schema construction (Renkl, 2014). Compared to self-explanation or comparison-prompts, monitoring prompts are less often investigated in research on example-based learning. However, based on the findings of research on self-regulated learning, it appears a fruitful approach to combine self-explanation and monitoring prompts, because both cognitive and metacognitive processes will then be activated. How these processes can be influenced by different sequencing strategies for modeling examples and learning tasks and how different learning tasks and prompting might interact with different sequences will be discussed in the next section.

2.3 Sequencing of example-based learning activities

A combination of modeling or worked examples with problem solving – compared to problem solving only – is an effective mean to foster early cognitive skill acquisition, because it reduces extraneous cognitive load and enables the learners to apply the principles that are instantiated by the examples (see Chapter, 2.2). However, when combining modeling examples and problem solving, different sequencing strategies can be implemented. Thereby, an important question is how different sequencing strategies affect cognitive skill acquisition and whether this occurs by influencing extraneous cognitive load and application of principles. To approach this question, I first will shed light on what sequencing is about and which sequences could be implemented, before I explore the question on different sequencing effects based on theoretical arguments and empirical evidence.

So far, research on sequencing did not provide a clear description of how sequencing strategies in example-based learning are defined and how they are distinct from sequences of example-based learning activities. Therefore, I will provide an own definition of sequencing. In this dissertation, I understand sequencing as intentional activity of the teacher where (s)he planes the sequence of – in this case – modeling examples and problem solving or other learning tasks, as opposed to any orders or sequences as (incidental) feature of the learning situation. When teachers plan the use of modeling examples and problem solving in their lessons, they have several options how to sequence modeling examples and problem solving. An often used sequence in research on EBL is to provide modeling examples *alternated-before* problem solving compared to problem solving only (e.g., Cooper & Sweller, 1987; Paas & Van Merriënboer, 1994). An alternated-before sequence of modeling examples and problem solving was furthermore compared to a *blocked-before sequence* of modeling examples and problem solving (Trafton & Reiser, 1993) as well as compared to an *alternated-after sequence* (e.g., Leppink et al., 2014; Van Gog, 2011; Van Gog et al., 2011). Additionally, sequencing of modeling examples *blocked-after* problem solving was also compared to a blocked-before sequence (e.g., Schwartz et al., 2011). Before I will discuss the studies on those comparisons in detail, I will first explore how different sequences effect cognitive skill acquisition by influencing extraneous cognitive load and application-quality.

Sequencing modeling examples *alternated-before* problem solving has the advantage that learners can acquire cognitive schemata about the problems, which (1) fosters the reduction of extraneous cognitive load (e.g., Renkl, 2014) and (2) helps to better apply the principles during later problem solving (e.g., Trafton & Reiser, 1993; Van Gog et al., 2011). A reduction of extraneous cognitive load and better application-quality should in turn support cognitive skill acquisition (e.g., Renkl, 2014). For a *blocked-before* sequence of modeling examples and problem solving, – theoretically – the same argumentation should hold true. However, it might be difficult for the learners to remember all modeling examples and to select the appropriate examples to map to the problem solving tasks (e.g., Trafton & Reiser, 1993, Van Lehn, 1996). Thus, they might need to rely on week problem solving strategies such as means-end-analysis, which induces high cognitive load (e.g., Renkl, 2014). When providing an *alternated-after* sequence of modeling examples and problem solving, extraneous cognitive load should be higher during learning compared to an (alternated-) before sequence, because the learners mainly rely on means-end-analysis when they have no example to map. Accordingly, application-quality should also be affected (Van Gog et al., 2011). However, an advantage of sequencing modeling examples alternated after can be, that the learners experience desirable difficulties during problem solving (Bjork, 1999) and thus be aware of their knowledge gaps (e.g. Loibl et al., 2017). This awareness should lead to a different attentional focus when receiving the modeling examples (e.g., Van Gog et al., 2011). This in turn might help to better connect the information in the modeling example with the learners' prior knowledge (e.g., Loibl & Rummel, 2014). For sequencing modeling examples *blocked-after* problem solving the same pattern and argumentation as for the alternated-after sequence should hold true. This is, that extraneous cognitive load is higher and application-quality is lower than in before-sequences (e.g., Van Gog et al., 2011). The described advantage of the awareness of knowledge gaps should also be the case for this sequence (Loibl et al., 2017). A difference to the alternated-after sequence is that when the information of the modeling examples is connected to the prior knowledge of the learners, they cannot test and apply their acquired schemata.

Because the main interest in all studies on sequencing is whether it influences cognitive skill acquisition, I will first discuss studies whether sequencing actually effects cognitive skill acquisition, , before I will shed light on findings with respect to extraneous cognitive load, application-quality and awareness of knowledge gaps as potential explanations for sequencing effects on cognitive skill acquisition.

2.3.1 Effects of sequencing on cognitive skill acquisition

Whether worked examples should be sequenced alternated- or blocked-before problem solving was investigated by Trafton and Reiser (1993). They compared – besides other conditions– an alternated-before sequence with a blocked-before sequence of worked examples and problem solving with respect to cognitive skill acquisition. In their study, 20 students learned about LISP-programming skills in those two conditions. Results showed that the students who learned with the alternated-before sequence significantly outperformed students in the blocked-before condition in the posttest. Trafton and Reiser (1993) interpreted these results in light of cognitive skill acquisition models (see Chapter 2.1), that is that the examples that are sequenced alternated before could be easier mapped and applied during problem solving compared to the blocked-before sequence, which supported cognitive skill acquisition. However, due to the very small sample of 10 students per condition, a replication of the findings would be preferable to provide further evidence. So far, further systematic research on comparing an alternated-before and blocked-before sequence with respect to cognitive skill acquisition is missing.

The question whether modeling examples should be sequenced before or after learning tasks such as problem solving can be regarded in the broader context of sequencing instruction and practice (Kant et al., 2017). Research on different sequencing strategies provided mixed and inconclusive results. There are several studies that provide evidence in favor of instruction (i.e. modeling examples) before practice (i.e. problem solving or other learning tasks) (e.g. Van Gog et al., 2011; Leppink et al., 2014; Hsu et al., 2015, Kant et al., 2017). Van Gog and colleagues (2011), for example, investigated whether different sequencing strategies of worked examples and problem solving influence cognitive skill acquisition. Therefore 51 secondary educational science students learned 30 minutes with two pairs of worked examples and problem solving tasks how to apply Ohm's law for reasoning about two different electrical circuit faults, either in an alternated-before or alternated-after sequence. Besides other results their study showed that sequencing worked examples and problem solving tasks in an alternated-before sequence significantly led to better cognitive skill acquisition than an alternated-after sequence. The authors argue that worked examples before problem solving can help to acquire cognitive schemata that guide later problem solving, which can better be applied and refined during problem solving and therefore leads to better performance than an alternated-after sequence. Van Gog et al. (2011) provide three possible explanations, why problem-example pairs are not that effective. This is, that novice learners cannot assess their deficiencies during problem solving accurately (Dunning et al., 2003), that leaners may not be motivated due to the experience of failure before or, that the students may be inclined to hindsight bias (Bjork, 1999). Leppink and colleagues (2014) replicated Van Gog and colleagues' (2011) study in the domain of statistics – more specifically the application of Bayes' theorem. The learning time in their study was, however, very short (i.e., ten minutes only). Their study showed that sequencing worked examples and problem solving tasks in an alternated-before sequence significantly led to better cognitive skill acquisition than an alternated-after sequence. Kant and colleagues (2017) conceptually replicated the study of Leppink et al. (2014), but in the domain of learning physics in school. Seventh-graders learned in approximately one school session about the control-of-variable strategy in a simulation-based learning environment. They also showed that alternated-before sequence of modeling examples and learning tasks significantly led to better cognitive skill acquisition than alternated-after sequence. Lastly, Reisslein and colleagues (2006) examined the effects of sequencing worked examples either alternated-before or alternated-after problem solving. Second they investigtaed, whether a potential effect is dependent on prior knowledge. The pretest was one week before the intervention and the students with low and high prior knowledge were randomly assigned to the conditions. The learning goal was to calculate the resistance of two types of electrical circuits. The computer-based learning environment consisted of four examples and problems (i.e., two per calculation-type). The results showed a significant interaction effect of sequencing and prior knowledge on cognitive skill acquisition: for the alternated-before sequence learners with low and high prior knowledge had an equal level of cognitive skill acquisition, but for the alternated-after sequence learners with high prior knowledge significantly outperformed learners with low prior knowledge on cognitive skill acquisition (Reisslein et al., 2006).

Based on the reported studies sequencing examples before learning tasks should be preferred with respect to cognitive skill acquisition, because it should help to better apply the cognitive schema acquired from the example than in a reverse sequence, at least for learners with low prior knowledge. However, the research lines on productive failure (e.g., Kapur, 2012, 2014) and preparation for future learning (e.g. Schwartz et al., 2011) provide evidence in favor of sequencing instruction after practice with respect to performance. An explanation for the effectiveness of those problem solving first-approaches is that leaners might identify their knowledge gaps during problem solving and focus their attention on those critical aspects during the following instruction, which should lead to a better connection and integration of the provided knowledge (Van Gog et al., 2011; Loibl et al., 2017). Schwartz, Chase, Oppezzo, & Chin (2011) investigated whether sequencing worked examples blocked before the learning tasks (i.e., tell and practice-format) or sequencing worked examples blocked after learning tasks (i.e., inventing with contrasting cases) would lead to more transfer and structural understanding in physics education. The students with the blocked-before sequence had a lecture beforehand and were instructed with a worked example before they practiced with the contrasting cases. The students with the blocked-after sequence were instructed to invent a procedure for computing an index and had the lecture at the end. Students in both conditions learned in pairs over eight days. The students in the blocked-after condition had a significant better conceptual understanding as well as transfer than blocked-before students. Kapur (2012) compared sequencing modeling examples before the problem solving (i.e., direct instruction before problem solving) with sequencing modeling examples blocked after learning tasks (i.e., productive failure during problem solving before direct instruction) with regard to learning the concept of variance. The study had a pre-post quasi experimental between-subjects design with 133 9th grade students. The instructional phase lasted about four sessions with 50 minutes each. The students with the before-sequence got two modeling example-problem pairs in the first session. In the second session the students with the before-sequence individually worked on three data analysis problems and the solutions were discussed in class. During the third session, the before-students worked in triads on generation problems where they had to examine quantitative indices for data analysis problems and the teacher discussed the solution. In the fourth session, the before-students individually solved three problems. The students with the blocked-after sequence worked in triads on the same generation problems as they were used in the before-condition, but during the first two sessions. In the third session the teacher compared and contrasted the student's solution and modeled and explained the canonical solution as in the before-condition. During the fourth session the blocked-after students practiced with three problems and the teacher discussed the solution. The posttest took place after the four learning phase sessions. Results of Kapur's (2012) study showed that learners with blocked-after sequence significantly outperformed students with the before-sequence with respect to conceptual understanding and transfer. However, it is not clear whether this effect can actually be attributed to the different sequencing of the modeling examples and learning tasks or for example to the fact, that the after-students had two sessions with generation problems compared to only one session in the before-condition. The number and the duration of working on specific learning tasks should be equal in both conditions to eliminate this explanation. A further difference between the conditions is that in the before-conditions, the modeling examples are sequenced alternated before problem solving in the first session, whereas in the after-conditions, the modeling examples were provided blocked after problem solving.

Studies in a more observational learning tradition also provide (some) evidence for a problem-example sequence (e.g. Baggett, 1987; Gräsel & Mandl, 1993). For example, Baggett (1987) investigated the effects of different sequences of instructions (i.e. modeling examples) and practice (i.e. problem solving) on the skills to build a model helicopter from 54 pieces. Besides others, Baggett (1987) compared two sequences: (1) modeling example before problem solving and (2) modeling example after problem solving. 60 psychology students were randomly assigned to one of the two conditions. Results showed that students with the after-sequence performed significantly better than all other conditions on a seven days delayed posttest. Gräsel and Mandl (1993) investigated the effect of modeling examples on the acquisition of diagnostic reasoning strategies and whether the modeling example should be provided before or after solving a problem. A major limitation is, however, that they only report descriptive analyses. Whether the differences are also statistically important was not investigated. Furthermore, the sample only consisted of 18 medicine students. They showed that learners without the modeling example and only solving a problem had the lowest level of cognitive skill acquisition, whereas learners with the modeling example before problem solving a higher level, and leaners with the modeling example after problem solving had the highest level of cognitive skill acquisition.

There are also some studies which did not find significant differences between example-problem pair and reverse sequences (Stegmann et al., 2012; Van Gog, 2011; Van der Meij et al., 2017). All studies with non-significant results raise the question of statistical power problems. Studies with little statistical power do not help in providing answers on the matter of sequencing when showing non-significant results. For example, Van Gog (2011) showed no significant differences in test performance alternated-before and an alternated-after sequence of modeling examples and problem solving tasks. However, the sample consisted of only 32 students. According to G*Power analysis (Faul et al., 2007), Van Gog's (2011) study would have needed more than 100 participants to show a medium effect with a probability of 80 %. Similarly, Van der Meij and colleagues (2017) also did not show significant differences in test performance between an alternated-before and an alternated-after sequence of modeling examples and problem solving tasks. Their sample size for this two conditions was 55 in total, and thus most probably not high enough to show a medium effect.

Taken together, the presented studies on sequencing provide mixed and inconclusive results. For the positive effect of examples before learning tasks, literature provides two main explanations: (1) providing modeling examples before learning tasks can lead to an abstract schema that guides later problem solving (e.g. Van Gog et al., 2011). This implies, that extraneous cognitive load should be reduced, because the working-memory-intensive means-end-analysis is not necessary (Renkl, 2014; Sweller & Cooper, 1985); and (2) that the acquired schema can be tested, redefined or generalized by applying it in problem solving after the modeling example was provided (e.g., Van Lehn, 1996). Although some of the studies investigated sequencing effects at least for one cognitive process (i.e. extraneous cognitive load or application-quality), none of the studies addressed them as mediators (see Chapter 2.3.2). The explanations for a positive effect of sequencing modeling examples after learning tasks are twofold: (1) one reason is seen in supporting the awareness of knowledge gaps (e.g., Loibl & Rummel, 2014; Loibl et al., 2017). (2) Another explanation is seen in the implementation of specific design features, especially the use of comparing examples as learning tasks (e.g., Alfieri, et al., 2013; Loibl et al., 2017).

2.3.2 Effects of sequencing on extraneous cognitive load

A reduction in extraneous cognitive load could be an explanation for a positive effect of an (alternated-) before sequence on cognitive skill acquisition. However, only parts of the studies presented in the previous Chapter do also empirically address this explanation. Van Gog and colleagues (2011) also criticized that previous studies on sequencing did not address extraneous cognitive load although they are arguing with it. Their study showed that extraneous cognitive load indicated with mental effort during the learning phase was significantly lower in conditions for sequencing worked example *alternated-before* problem solving compared to an *alternated-after* sequence. Van Gog and colleagues (2011) argue that the pattern of less extraneous cognitive load effort and better test performance is an indicator for a higher efficiency of the learning process, even though mental efficiency is not reported. Although extraneous cognitive load is seen as a mechanism by the authors, why the example-problem sequence should lead to better performance, their study did not investigate the possible mediating role of extraneous cognitive load. Hsu and colleagues (2015) showed that for learners with low prior knowledge, an alternated-before sequence significantly reduced extraneous cognitive load indicated with mental effort than an alternated-after sequence of worked examples and problem solving. For learners with higher prior knowledge, extraneous cognitive load did not differ significantly between the sequencing conditions. Kant and colleagues (2017) revealed that extraneous cognitive load indicated with mental effort was significantly lower in the first learning phase, when a modeling example was provided compared to a learning task. However, in the second learning phase, no significant differences were shown for an alternated-before and alternated-after sequence of modeling examples and learning tasks. Similarly, Leppink and colleagues (2014) also showed no significant differences in extraneous cognitive load between an alternated-before and alternated-after sequence of worked examples and problem solving. Furthermore, their study did not find a relation of extraneous cognitive load and test performance. However, a sample of only 18 students per condition (i.e., 36 for both conditions) could point to a not enough statistical power to show a medium effect of sequencing on extraneous cognitive load.

Taken together, the majority of the studies on sequencing-effects showed a positive effect of an alternated-before sequence for extraneous cognitive load. However, none of the studies did investigate extraneous cognitive load as mediator for an effect of sequencing on cognitive skill acquisition, even though the authors regard it a mechanism. Therefore, the actual mediating role of extraneous cognitive load is still an open question so far and should be clarified in detail.

2.3.3 Effects of sequencing on application-quality

Different sequencing strategies in example-based learning should influence schema construction indicated by application-quality, that means sequencing modeling examples (alternated) before learning tasks should better allow for retrieval, mapping and application of the specific example features to the learning task (Van Lehn, 1996) compared to sequencing modeling examples (alternated) after learning tasks (e.g., Van Gog et al., 2011). However, only mixed evidence is provided by the few studies that investigate sequencing effects on application-quality. Trafton & Reiser (1993), for example, showed no significant differences in application quality – indicated as accuracy of first solution attempts - between an alternated-before or blocked-before sequence. However, their sample size with only ten participants per condition points to statistical power problems. With regard to an alternated-before or -after sequence of worked example and problem solving, Reisslein and colleagues (2006) also showed no significant difference in application-quality. In their study, 61 students participated per condition, thus there would have been enough power to show a medium effect. In contrast to this, Van Gog (2011) showed that students with an alternated-before sequence significantly outperformed students with an alternated-after sequence with respect to application-quality.

Taken together, the potential role of application-quality as mediator for sequencing effects is not yet clear, because only little research with mixed results took place on sequencing and application-quality and none of the reported studied did analyze application-quality as mediator. However, this should be clarified in detail, because it is assumed to be a mechanism for the effectiveness of different sequencing strategies in example-based learning. As long as application-quality is not investigated as mediator, this cannot be explored.

In conclusion, research on effects of sequencing with respect to extraneous cognitive load and application-quality provides mixed results. So far, none of the studies investigated extraneous cognitive load or application-quality or even both as mediator for sequencing effects, although it is argued for a superiority of (alternated) before-sequences with these processes. My dissertation contributes in closing this gap.

2.3.4 Sequencing and type of learning tasks

As presented in Chapter 2.3.1, research on sequencing effects revealed mixed results. One reason for this might lie in the different learning tasks used in previous research on sequencing. Studies that provide evidence for sequencing modeling examples before learning tasks used problem solving as learning task (e.g. Van Gog et al., 2011; Leppink et al., 2014), whereas studies that provide support for a reverse sequence used invention or comparing tasks (e.g. Kapur 2012, 2014 Schwartz et al., 2011). Alfieri and colleagues' (2013) meta-analysis on case comparison could be seen as indicator for a possible interacting relationship of sequencing and type of learning task. Their meta-analysis explored findings of 57 experiments that include 336 comparisons. They showed that providing principles (such as in modeling examples) after example comparison is more effective with respect to conceptual and procedural knowledge compared to providing no principles or before comparing example, because it might better support the modification of an abstracted schema (Holyoak, 2012). The authors also highlight, that example comparison as learning task may be especially helpful to provide "preparation for future learning" (Schwartz & Bransford, 1998; Schwartz et al., 2011) in order to profit from following instruction (Alfieri et al., 2013). This finding might be regarded as hint that there could be an interaction between sequencing and type of learning task in the sense that for that for contrasting cases as learning, task modeling examples afterwards might be more effective than examples before. Moreover, Schwartz and colleagues (2011) showed that a blocked-after sequence had a significant higher recall of deep structure as well as transfer then the students in the blocked-before sequence. They also analyzed the cognitive processes and showed that blocked-after students made more transitions between the examples by searching for commonalities than the students in the blocked-before condition. Also Loibl and colleagues (2017) highlight in their review on problem solving-first approaches, that specific design features need to be implemented in order for problem solving-first approaches to be effective. Those design features are (1) the implementation of comparing example or contrasting cases and (2) supporting the awareness of knowledge gaps (e.g. by prompting) (Loibl et al., 2017). Whereas this section focused on the first design feature that could interact with sequencing, the next section will clarify, how the effectiveness of different sequencing strategies could depend on prompting.

2.3.5 Sequencing and prompting

One advantage of a (alternated) after-sequence of modeling examples and learning tasks might be that learners experience struggle and uncertainties during solving the task, which could affect their attentional processes when receiving an example after problem solving (Van Gog et al., 2011). However, learners often need to be prompted to perform (meta-) cognitive processes like self-explanation and monitoring activities (see Chapter 2.2.3). So far, systematic research on possible interacting effects of sequencing of modeling examples and learning tasks and prompting is missing, therefore we mainly rely on theoretical arguments.

From a theoretical perspective different sequencing strategies and prompting self-explanations and monitoring might interact, because learners who are prompted to self-explain an example and diagnose critical problem solving aspects should be more aware of their uncertainties and then focus their attention in the modeling examples on those critical aspects (Bandura, 1986). Indirect evidence for this might be seen in Stark and colleagues' (2000) study. They investigated, whether an alternated-after sequence of worked examples and problem solving fosters (meta-) cognitive processes like example elaboration and monitoring as well as transfer compared to learning with worked examples only. 30 students learned with four problems and/or worked examples in the respective condition approximately 45 minutes. Their study showed that the alternated-after sequence significantly led to more elaboration activity of the examples, monitoring and regulation activities as well as higher (near, medium and far) transfer than learning with worked examples only. Although the study by Stark and colleagues (2000) did not vary different sequences or prompts, it might provide hints with respect to the role of (meta-) cognitive activities when providing examples after problem solving. This is, that prompting (meta-) cognitive processes might influence the effectiveness of specific sequences and leaners that are prompted might profit more from a sequence that provides modeling examples after problem solving.

Taken together, there are different possible explanations for a positive effect of sequencing modeling examples before learning tasks as well as for a positive effect of sequencing modeling examples after learning tasks. A positive effect of providing modeling examples before learning tasks is typically explained by having the benefit that this reduces extraneous cognitive load and fosters the application of the procedures, which in turn leads to better cognitive skill acquisition (e.g., Leppink et al., 2014; Van Gog et al., 2011), even though this relation is not empirically tested yet. A positive effect of providing modeling examples after learning tasks is typically explained by the awareness of knowledge gaps and a better connection of the instruction with prior knowledge (e.g., Loibl & Rummel, 2014; Loibl, Roll, & Rummel, 2017). Another possible explanation for a positive effect of sequencing modeling examples afterwards is the type of learning task the studies used, that is they often used comparison tasks (e.g., Schwartz et al., 2011, Alfieri et al., 2013). Based on the theoretical and empirical arguments and explanations for differential sequencing effect that I discussed in this Chapter, I will next derive two general research questions in Chapter 3 and will present a model where those explanations are integrated.

3 General research questions and the present studies

The first general research question refers to an overall effect of sequencing modeling examples and learning tasks on cognitive skill acquisition.

General research question I: To what extent do different sequencing strategies for modeling examples and learning tasks have an effect on cognitive skill acquisition?

> ***GRQ Ia:*** To what extent does an alternated vs. blocked sequence of modeling examples and learning tasks have an effect on cognitive skill acquisition?

An alternated sequence of modeling examples and learning tasks may lead to higher cognitive skill acquisition compared to a blocked sequence (Trafton & Reiser, 1993), because it should be easier to map and apply the specific features of the example to the learning task such as problem solving (Van Lehn, 1996).

> ***GRQ Ib:*** To what extent does sequencing of modeling examples before vs. after learning tasks have an effect on cognitive skill acquisition?

Thus far, research has provided mixed and inconclusive results for the influence of sequencing modeling examples before or after learning tasks on cognitive skill acquisition (Reisslein et al., 2006; Van Gog, 2011; Van Gog et al., 2011; Stegmann et al., 2012; Leppink et al., 2014; Baggett, 1987; Kapur, 2012; Schwartz et al., 2011, Hsu et al., 2015; Van der Meij et al., 2017; Kant et al., 2017). On the one hand, some studies have found positive effects of sequencing modeling examples before learning tasks compared to a reverse sequence for cognitive skill acquisition (Hsu et al., 2015; Kant et al., 2017; Leppink et al., 2014; Van Gog et al., 2011). On the other hand, studies have also shown a superiority of sequencing modeling examples after learning tasks instead of before in terms of cognitive skill acquisition (Baggett, 1987; Kapur, 2012; Schwartz et al., 2011). Lastly, some studies found no significant differences in cognitive skill acquisition between sequencing modeling examples before or after learning tasks (Van Gog, 2011; Stegmann et al., 2012; Van der Meij et al., 2017).

However, potential explanations for the effects of sequencing are rarely investigated by the previous studies. Therefore, in this dissertation I want to contribute to explaining (potential) effects of sequencing on cognitive skill acquisition.

General research question II: How can the different effects of sequencing of modeling examples and learning tasks on cognitive skill acquisition be explained?

There are some arguments for either sequence, i.e. providing modeling examples before or after learning tasks. A prototypical argument for examples before learning tasks is based on cognitive load theory (e.g., Van Gog et al., 2011; Leppink et al., 2014), namely that examples reduce extraneous cognitive load and this reduction is assumed to lead to better cognitive skill acquisition (Sweller, 2010). Also, providing modeling examples before learning tasks can lead to the construction of schemata that can be applied during the learning tasks (Van Gog et al., 2011), which should foster cognitive skill acquisition (Renkl, 2014).

A prototypical argument for providing modeling examples before learning tasks lies in the preparation for future learning (e.g., Schwartz and Bransford, 1998), that means that a differentiation of prior knowledge and the awareness of knowledge gaps should lead to better learning from the following modeling examples (e.g., Kapur, 2012; Loibl, Roll, & Rummel, 2017). Furthermore, 'comparing examples' as learning task is seen as especially helpful to provide preparation for future learning and that with this learning task modeling examples after the learning tasks can be more effective than before (Alfieri et al., 2013; Schwartz & Bransford, 1998; Schwartz et al., 2011).

The next question aims at investigating the main explanations for an assumed superiority in terms of cognitive skill acquisition of modeling examples before learning tasks.

> ***GRQ II a:*** To what extent is a potential positive effect of sequencing of modeling examples before learning tasks on cognitive skill acquisition mediated by extraneous cognitive load and application-quality during learning?

With respect to extraneous cognitive load there is – albeit indirect – evidence from research on example-based learning in general that it could play a mediating role (e.g., Renkl, 2014): when modeling examples are provided before learning tasks, working memory capacity is relieved because no means-end-analysis for problem solving is needed (Sweller & Cooper, 1985). Although some studies on sequencing investigated effects on extraneous cognitive load (e.g., Leppink et al., 2014) and application-quality (e.g., Van der Meij et al., 2017) as potential explanations for an effect of sequencing modeling examples before learning tasks instead of after learning tasks, none of them explored whether extraneous cognitive load and application-quality actually are mediators for effects of different sequences.

A typical argument with regard to application-quality is that providing modeling examples before learning tasks should allow for mapping and application of the specific example features to the learning task, when compared to sequencing modeling examples after the learning tasks (e.g., Van Gog, 2011), because no features or principles are provided to map and apply (Van Lehn, 1996). Thus, application quality should play a mediating role when the sequence of providing modeling examples before learning tasks appears effective for cognitive skill acquisition.

The next question aims at investigating the main explanations for an assumed superiority in terms of cognitive skill acquisition of providing modeling examples after learning tasks instead of before.

> ***GRQ II b:*** To what extent does a potential positive effect of sequencing modeling examples after learning tasks on cognitive skill acquisition depend on prompting self-explanations and monitoring or on different types of learning tasks?

Prompting self-explanations and monitoring is expected to interact with sequencing modeling examples before or after learning tasks such as problem solving, because learners who are prompted to self-explain an example and diagnose critical problem solving aspects should be more aware of their knowledge gaps (e.g., Loibl, Roll, & Rummel, 2017) and subsequently focus their attention in the modeling examples on those critical aspects (Bandura, 1986). Therefore, learners might profit more from modeling examples after learning tasks, when they are prompted to self-explain and monitor.

Also, an interaction between sequencing and type of learning task for cognitive skill acquisition is assumed. This means for 'comparing examples' as a learning task, a modeling example after a learning task might be more effective than before a learning task, because comparing examples may be especially helpful to provide "preparation for future learning" (Schwartz & Bransford, 1998; Schwartz et al., 2011). This preparation should help to profit from subsequent learning resources such as modeling examples (Alfieri et al., 2013). Alfieri and colleagues' (2013) meta-analysis on case comparison showed that providing principles after case comparison is more effective with respect to the acquisition of conceptual and procedural knowledge compared to providing before comparing cases or providing no principles at all, because it might better support the modification of an abstract schema (Holyoak, 2012). Thus, an effect of sequencing modeling examples after learning tasks might depend on the type of learning task. Hence, the effect of sequences that provide modeling examples before learning tasks might be reduced or inversed when 'comparing examples' is the learning task as opposed to 'problem solving' as learning task (Alfieri et al., 2013). This kind of interaction might also occur when the learning tasks consists of problem solving accompanied by a comparison of one's own solution with the example solution, because the learners might become more aware of their uncertainties and knowledge gaps (e.g., Loibl et al. ,2017) and can then focus their attention in the modeling examples on those critical aspects (Bandura, 1986).

As basis for the investigation of these general research questions, a model for the effectiveness of different sequencing strategies (see Figure 1) was developed, that is based on the theoretical assumptions and empirical findings described in Chapter 2.

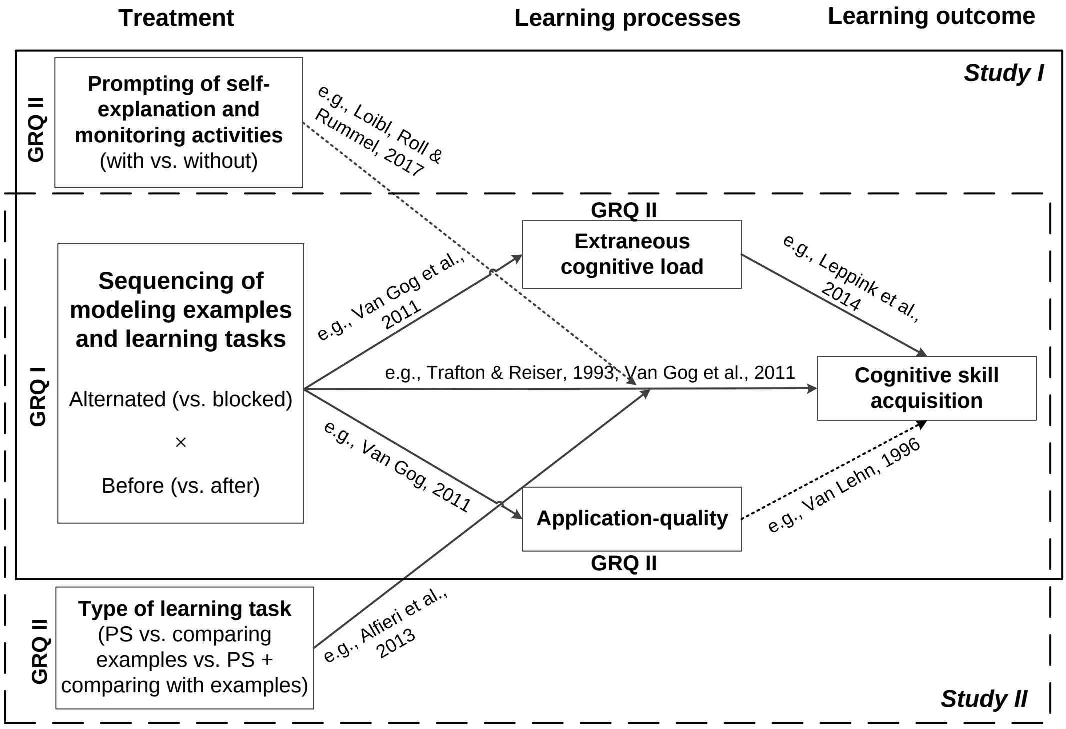


Figure 1. Model for the effectiveness of different sequencing strategies in example-based learning

Note. 'GRQ' refers to general research question, 'PS' stands for problem solving; dotted lines indicate that the assumed relation is based on theoretical arguments "only", whereas continuous lines indicate expected relations based empirical evidence.

The present studies

To examine these general research questions and assumptions, I conducted two empirical studies. Study I (Chapter 4) investigated, whether an alternated or blocked sequencing of modeling examples and learning tasks (i.e. problem solving) should be preferred with respect to cognitive skill acquisition (GRQ I a). To this end, Study I compared the effects of an alternated sequence of modeling example before problem solving with a blocked sequence of modeling example before problem solving. Furthermore, I investigated whether modelling examples should be provided before or after learning tasks (i.e. problem solving) for cognitive skill acquisition (GRQ I b). For a blocked-before and blocked-after sequence I expected larger differences in cognitive skill acquisition than for an alternated-before and alternated-after sequence, because when modeling examples are alternated with learning tasks, at least the second example can be regarded as both, after the learning task as well as before the next learning task. In a first step, I did not investigate all four combinations of alternated vs. blocked and before vs. after sequences in Study I due to potential power issues, but concentrated on comparing of an alternated-before, blocked-before and a blocked-after sequence. I also examined potential explanations for effects of sequencing (if any) on cognitive skill acquisition (GRQ II) in Study I. That is (1) whether extraneous cognitive load and application-quality mediate a potential positive effect of sequencing of modelling examples before learning tasks and (2) whether a potential positive effect of sequencing modelling examples after learning tasks is dependent on prompting self-explanations and monitoring. To investigate the two general research questions, I derived four specific research questions in Study I (see Chapter 4.1).

The sample of Study I consisted of 126 educational sciences students in their first semester. The students were randomly assigned to six groups in a 2×3 factorial design with two between-subject factors (i.e., provision of prompts; sequencing of modeling examples and learning tasks) and repeated measurement of cognitive skill acquisition at T1 directly after and at T2 four days after the intervention as within-subjects factor. Extraneous cognitive load and application-quality were assessed while students were working on the learning tasks.

In Study II (Chapter 5) I examined, whether modeling examples should be provided before or after learning tasks for cognitive skill acquisition (GRQ I b). Hence, an alternated-before sequence (i.e. modeling examples alternated before learning tasks) was compared with an alternated-after sequence (i.e. modeling examples alternated after learning tasks). Second, I examined potential explanations for an effect of sequencing (if any) on cognitive skill acquisition

(GRQ II). That is (1) whether a potential positive effect of sequencing modelling examples before learning tasks on cognitive skill acquisition is mediated by extraneous cognitive load and application-quality and (2) whether a potential positive effect of sequencing modelling examples after learning tasks is dependent on the type of learning task. Therefore, I compared three learning tasks: (1) problem solving only, (2) comparing examples, and (3) problem solving and comparing with examples. Problem solving only was implemented, because it is the standard learning task in research on sequencing strategies (e.g., Van Gog et al., 2011; Leppink et al., 2014, Reisslein et al., 2006). Comparing examples was implemented in line with the meta-analytical finding of Alfieri and colleagues (2013) that comparing examples is more effective when the principles are provided after observing modelling examples. Hence, it might also be possible that an effect of sequencing (if any) is dependent on the type of learning task. Problem solving and comparing with examples was included because it should support the observational learning process of monitored enactment (Bandura, 1991; see Chapter 2.2.2).

The sample of Study II consisted of 145 educational sciences students in their first semester. The students were randomly assigned to one of six conditions in a 2×3 factorial design with two between-subject factors (i.e., sequencing of modeling examples and learning tasks; the type of learning task: problem solving, comparing examples, and problem solving and comparing with examples). Cognitive skill acquisition was measured directly after the intervention. Extraneous cognitive load and application-quality were assessed three times while students worked on the learning tasks.

4 Effects of sequencing and prompting on cognitive skill acquisition (Study I)

4.1 Research Questions

The present study investigates (1) the effect of sequencing of modeling examples and learning tasks on cognitive skill acquisition (GRQ I). Second, my study explores (2) which explanations might clarify potential effects of sequencing on cognitive skill acquisition (GRQ II). More specifically, whether extraneous cognitive load and application-quality might meditate a potential positive effect of sequencing modeling examples before learning tasks on cognitive skill acquisition and whether a potential positive effect of sequencing modeling examples after learning tasks is dependent on prompting self-explanations and monitoring. To investigate these two general research questions (see Chapter 3), four specific research questions and corresponding hypotheses are formulated. The General Research Question I (GRQ I, see Chapter 3) is addresses with research question one. The General Research Question II (GRQ II, see Chapter 3) is specified by all four research questions of the present study.

RQ 1: To what extent does the sequencing of modeling examples and learning tasks, prompting and the interaction thereof have an effect on cognitive skill acquisition?

Different sequences of modeling examples and learning tasks should lead to different degrees of cognitive skill acquisition (Leppink et al., 2014; Van Gog et al., 2011) because they might affect cognitive processes like mapping and application in example-based learning. Mapping and application processes are important for schema construction (see Chapter 2.2.2), which is the main goal in early phases of cognitive skill acquisition (see Chapter 2.1). An alternated-before sequence should allow for mapping and application of the specific example features to the learning task compared to a blocked-before sequence, where it could be difficult to remember and select the right example features to map (Trafton & Reiser, 1993). When modeling examples are provided blocked after learning tasks, no features or principles are provided to map and apply (Van Lehn, 1996). Thus, an alternated- and blocked-before sequence should lead to higher cognitive skill acquisition than a blocked-after sequence.

Banduras theory on observational learning (1986) provides a further theoretical argument for an alternated-before sequence compared to a blocked-before sequence, since observational learning should have the advantage that the learners can directly compare their solution method with the modeling example and focus their attention on those critical aspects in the next modeling example they were not sure about during problem solving. However, these different attentional processes might also play a role when the modeling examples are provided blocked after problem solving. Prompting self-explanations and monitoring is assumed to interact with sequencing modeling examples before or after learning tasks such as problem solving, because learners who are prompted to self-explain an example and diagnose critical problem solving aspects should be more aware of their knowledge gaps (e.g., Loibl, Roll, & Rummel, 2017). Subsequently, the learners should focus their attention in the modeling examples on those critical aspects (Bandura, 1987). Therefore, modeling examples after learning tasks might be more beneficial, when the learners are prompted to self-explain and monitor.

None of the prior studies on sequencing of example-based learning activities (e.g., Leppink et al, 2014; Van Gog et al., 2011) investigated how stable short-term sequencing effects are and whether there is an increase in skill acquisition over time, dependent on sequencing. The present study wants to contribute to answer this open question.

- ***H1.1***: An alternated-before sequence of modeling example and learning tasks leads to higher cognitive skill acquisition than a blocked-before sequence (Trafton & Reiser, 1993).

- ***H1.2***: An alternated-before sequence of modeling example and learning tasks leads to higher cognitive skill acquisition than a blocked-after sequence (Leppink et al., 2014; Van Gog et al., 2011).
- ***H1.3***: A blocked-before sequence of modeling example and learning tasks leads to higher cognitive skill acquisition than a blocked-after sequence (Leppink et al., 2014; Van Gog et al., 2011).
- ***H1.4:*** A potential effect of sequencing is dependent on prompting self-explanations and monitoring. This means, a potential effect of an alternated-before sequence might be reduced or inversed, when self-explanations and monitoring are prompted as opposed to no prompting (Loibl & Rummel, 2014; Loibl, Roll, & Rummel, 2017).

RQ 2: To what extent does the sequencing of modeling examples and learning tasks, prompting and the interaction thereof, have an effect on extraneous cognitive load?

Learners who first receive a modeling example and then solve a learning task (such as problem solving) may be able to apply the cognitive schema they acquired from the modeling example during problem-solving. Thus, they do not have to rely on means-ends-analysis, a weak problem-solving strategy that requires a lot of working memory capacity (Renkl, 2014; Van Gog & Rummel, 2010). Therefore, an alternated- and blocked-before sequence may lead to lower extraneous cognitive load than a blocked-after sequence (Van Gog et al., 2011). However, with a blocked-before sequence learners may not be able to remember all modeling examples and need to rely on weak methods again (Trafton & Reiser, 1993).

- ***H2.1***: An alternated-before sequence of modeling example and learning tasks leads to lower extraneous cognitive load than a blocked-before sequence (Trafton & Reiser, 1993).

- ***H2.2***: An alternated-before sequence of modeling example and learning tasks leads to lower extraneous cognitive load than a blocked-after sequence (Van Gog et al., 2011).
- ***H2.3***: A blocked-before sequence of modeling example and learning tasks leads to lower extraneous cognitive load than a blocked-after sequence (Van Gog et al., 2011).

RQ 3: To what extent do the sequencing of modeling examples and learning tasks, prompting and the interaction thereof, have an effect on application-quality?

Providing an alternated- or blocked-before sequence of modeling examples and learning tasks can provide opportunities for mapping and application of the specific example features to the learning task (Van Lehn, 1996). However, when modeling examples and learning tasks are provided as blocked-before sequence, the learners might have problems to remember and retrieve the appropriate example features and therefore may be less able to map and apply the features to the learning task (Trafton & Reiser, 1993). In a blocked-after sequence no features or principles are provided to map and apply. Thus, an alternated- and blocked-before sequence should lead to higher application-quality than a blocked-before sequence (Van Gog, 2011).

- ***H3.1***: An alternated-before sequence of modeling example and learning tasks leads to higher application-quality than a blocked-before sequence (Van Lehn, 1996).
- ***H3.2***: An alternated-before sequence of modeling example and learning tasks leads to higher application-quality than a blocked-after sequence (Van Gog, 2011; Van Lehn, 1996).
- ***H3.3***: A blocked-before sequence of modeling example and learning tasks leads to higher application-quality than a blocked-after sequence (Van Gog, 2011; Van Lehn, 1996).

RQ 4: To what extent is a potential positive effect of sequencing modeling examples before learning tasks on cognitive skill acquisition mediated by extraneous cognitive load and application-quality?

So far, none of the prior studies on sequencing (see Chapter 2.3) investigated extraneous cognitive load and application-quality as mediating variables. Cognitive load theory (e.g., Sweller, 2010) assumes, that extraneous cognitive load is detrimental for learning, thus it should play a mediating role. However, research on the role of extraneous cognitive load in example-based learning in general provides mixed results regarding the mediating value of extraneous cognitive load (see Chapter 2.2.1).

From a theoretical perspective on cognitive skill acquisition (e.g., Van Lehn, 1996), application-quality should play a mediating role because retrieval, mapping and application of specific features of the example to the learning task are considered important learning processes in early phases of cognitive skill acquisition (see Chapter 2.1).

- ***H4.1***: A potential effect of sequencing is mediated by extraneous cognitive load (Sweller, 2010).
- ***H4.2***: A potential effect of sequencing is mediated by application-quality (Van Lehn, 1996).

4.2 Method

4.2.1 Sample

One hundred and twenty-six German educational science students in their first semester participated On average they were 22 years old (*M* = 21.75; *SD* = 4.85) and 88.7% was female.

A G*Power analysis (Faul et al., 2007) for a medium sized interaction effect (*partial* $\eta^2 = .06$) of sequencing and repeated measurement of cognitive skill acquisition on T1 and T2 in an ANOVA with repeated measurement with a power of 80% indicates an optimal sample size of 60 subjects. For a medium sized interaction effect (*partial* $\eta^2 = .06$) of sequencing and prompting on cognitive skill acquisition (at T1 or T2) in an ANCOVA with prior knowledge as covariate with a power of 80%, G*Power analysis (Faul et al., 2017) indicates an optimal sample size of 155 subjects. Thus, the actual sample size of 126 subjects is optimal for showing an interaction of sequencing and repeated measurement, but not optimal for showing an interaction of sequencing and prompting.

4.2.2 Design

We conducted a 3×2 factorial design with two between-subject factors (i.e., sequencing of modeling examples and learning tasks and provision of prompts) and repeated measurement of cognitive skill acquisition at T1 (directly after the intervention) and at T2 (four days after the intervention) as within-subjects factor. The first between-subjects factor was the sequencing of modeling examples and learning tasks (alternated-before vs. blocked-before vs. blocked-after sequence). The second between-subjects factor was the prompting self-explanations and monitoring (with vs. without). Prior knowledge was measured before the intervention, whereas extraneous cognitive load and application-quality were assessed four times, when the students worked on the learning tasks. The students were randomly assigned to one of the six conditions. Table 2 shows the distribution of participants across the six conditions.

Table 2. *Overview of participants per condition*

		Sequencing of ME and LT		
		Alternated-before	Blocked-before	Blocked-after
Prompting	With	18	27	13
	Without	26	18	24

Note. ME = Modeling examples, LT = learning tasks.

4.2.3 Learning Environment

The objective of the learning environment was to teach different between- and within-subject designs according to the scheme of Campbell and Stanley (1963). The declarative knowledge was covered in a lecture one month before the study. The procedural knowledge was to be acquired during the intervention study, which was implemented in a seminar session one month after the lecture. During this session, the students worked individually on four modeling examples and four learning tasks that each consisted of a solved example and a problem to solve, For standardization purposes, the modeling examples were provided as digital videos. Each modeling example was 5 to 7 minutes long. In total the four modeling videos lasted 24 minutes. All modeling examples consisted of problem-solving tasks, where the model described a design according to specific research questions. The problem-solving task was printed on an overhead-transparency and the model solved the task by demonstrating the procedure step by step on the overhead-transparency. The model also externalized the rationale behind the procedure. The first modeling example demonstrated how to describe a one-factorial within-subjects design (6 minutes). The second modeling example showed how to describe a one-factorial between-subjects design (5 minutes). The third example demonstrated how to describe a balanced within-subjects design (6 minutes). The last example showed how to describe a 2×2 factorial between-subjects-design (7 minutes). The complexity of the modeling examples increased from the first example to the final example. The order of the modeling examples was the same in all conditions.

The learning tasks were printed as a learning booklet and consisted of a solved example on the left side and a problem on the right side. The structure of the solved examples and problems was the same as in the modeling example, only the surface features (i.e., the cover story) differed. The solved example consisted of the task description and the solution. The problem consisted of a task description only. The learners had 9 minutes to work on each learning task. In total they had 36 minutes for all four learning tasks.

4.2.4 Independent Variables and manipulations

Sequencing of modeling examples and learning tasks

The first manipulation consisted of an alternated-before sequence, where the modeling examples were provided alternated before the learning tasks. In the blocked-before sequence the students first received the all four modeling examples and then worked on the four learning tasks. The third manipulation consisted of a blocked-after sequence. In these conditions, the leaners first had to work on all four learning tasks and then received the four modeling examples. Figure 2 illustrates the manipulation of sequencing of modeling examples and learning tasks.

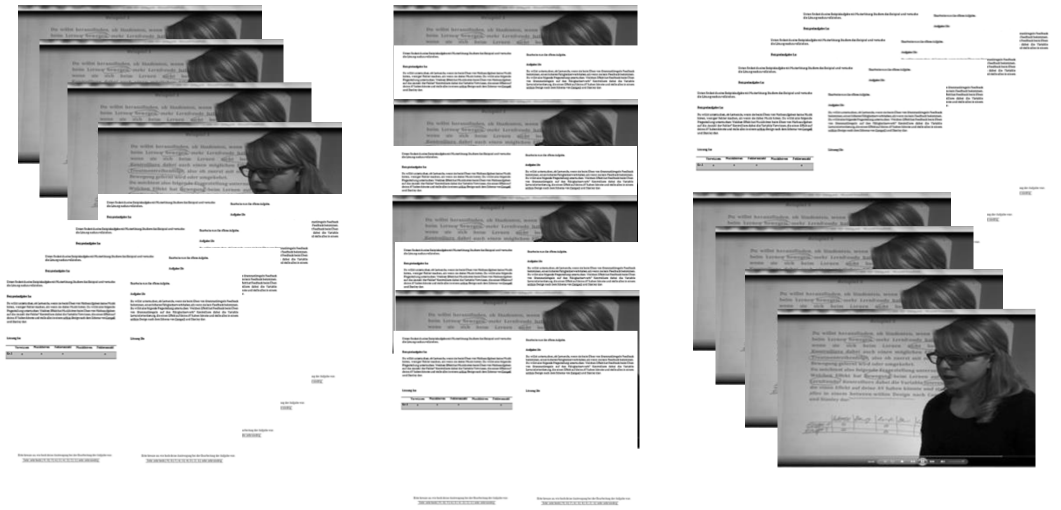
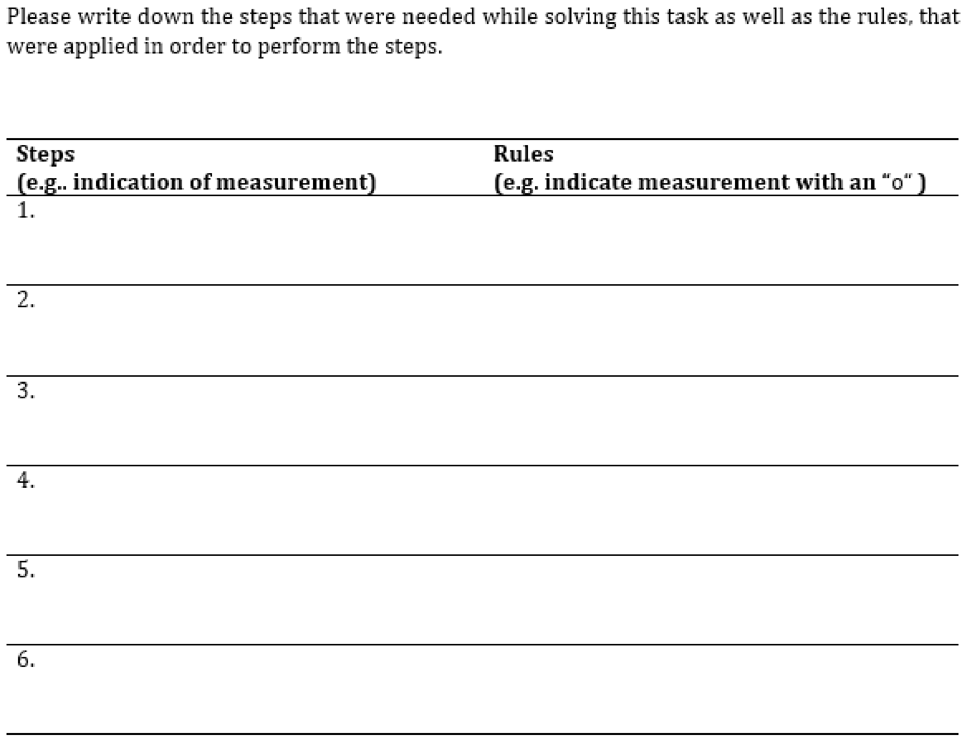


Figure 2. Overview of the manipulation

Note. The left column consists of blocked modeling examples before problem solving, the middle column consists of alternating modeling examples with problem solving and the right side consists of blocked modeling examples after problem solving.

Prompting self-explanations and monitoring

Prompting was integrated in the learning booklet. In the conditions with prompting, the students were prompted to self-explain the examples and to monitor themselves while solving the problems. The self-explanation prompts were placed underneath the solved example (see Figure 3).



Please write down the steps that were needed while solving this task as well as the rules, that were applied in order to perform the steps.

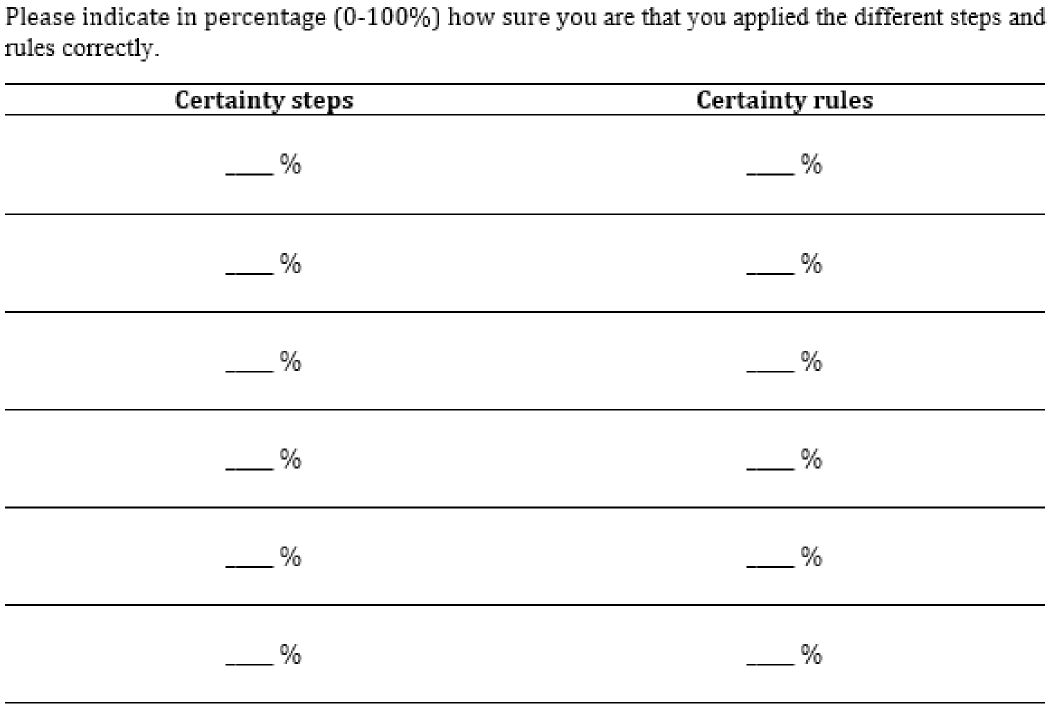
Steps **(e.g.. indication of measurement)**	**Rules** **(e.g. indicate measurement with an "o")**
1.	
2.	
3.	
4.	
5.	
6.	

Figure 3. Example of a self-explanation prompt

The prompts to monitor their own problem-solving process were placed underneath the problem-solving task and asked the students to indicate in percentage, how certain they were that they applied the steps and rules correctly (see Figure 4).

Conditions without prompting self-explanations and monitoring were only encouraged to study the examples carefully, but were not prompted to further self-explain the example or to monitor their problem solving,

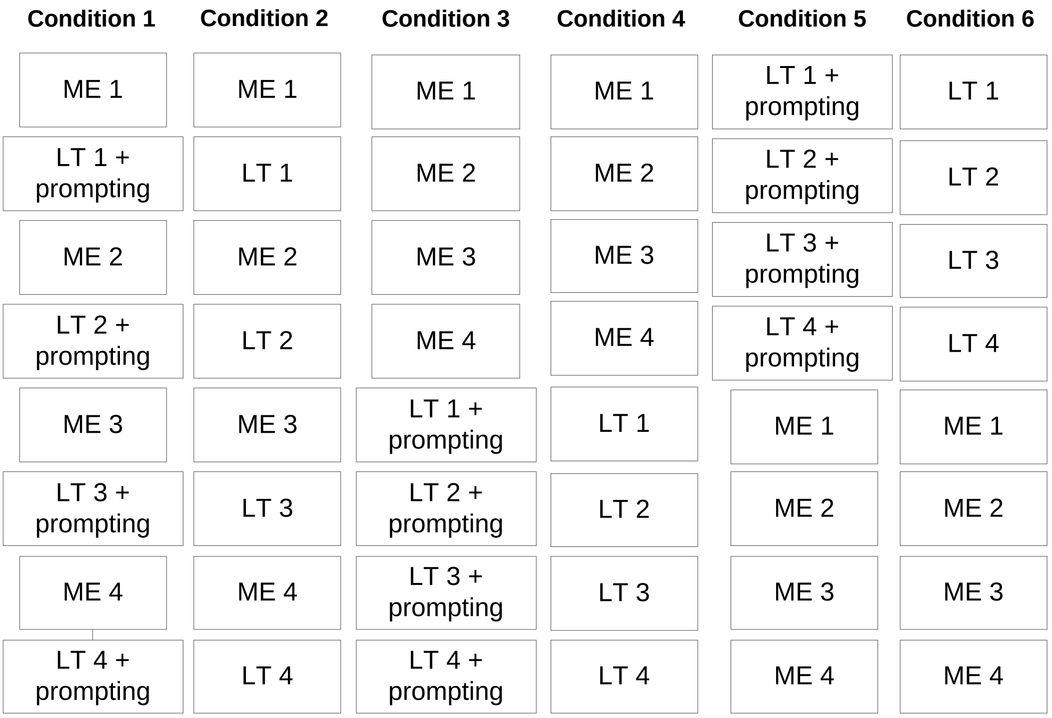
Combinations. The combinations of the both factors (i.e. the six conditions) are represented in Figure 5.



Please indicate in percentage (0-100%) how sure you are that you applied the different steps and rules correctly.

Certainty steps	Certainty rules
____ %	____ %
____ %	____ %
____ %	____ %
____ %	____ %
____ %	____ %
____ %	____ %

Figure 4. Example of a monitoring prompt



Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
ME 1	ME 1	ME 1	ME 1	LT 1 + prompting	LT 1
LT 1 + prompting	LT 1	ME 2	ME 2	LT 2 + prompting	LT 2
ME 2	ME 2	ME 3	ME 3	LT 3 + prompting	LT 3
LT 2 + prompting	LT 2	ME 4	ME 4	LT 4 + prompting	LT 4
ME 3	ME 3	LT 1 + prompting	LT 1	ME 1	ME 1
LT 3 + prompting	LT 3	LT 2 + prompting	LT 2	ME 2	ME 2
ME 4	ME 4	LT 3 + prompting	LT 3	ME 3	ME 3
LT 4 + prompting	LT 4	LT 4 + prompting	LT 4	ME 4	ME 4

Figure 5. Visualization of the six different conditions

Note. ME stands for modeling example, LT for learning task.

4.2.5 Measures

4.2.5.1 *Prior knowledge*

Before the intervention, prior knowledge of basic terms in empirical research methods was assessed directly before the intervention with a matching-task consisting of four items. Figure 6 shows an example item.

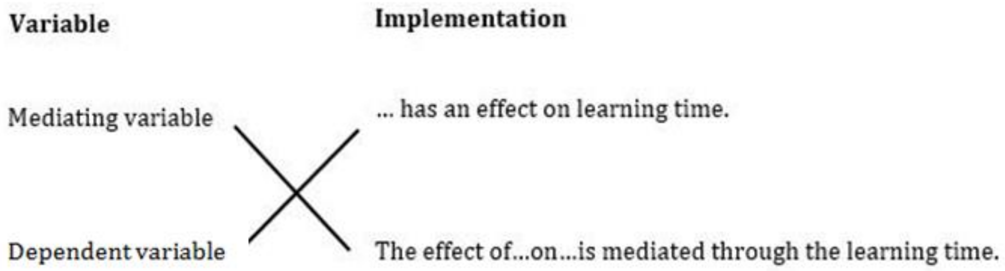


Figure 6. Example of an item for measuring prior knowledge

The four items were each scored with 1 point for a correct match and 0 points for an incorrect match (max. 4 points). The pretest scale had a Cronbach's alpha of .75.

4.2.5.2 *Extraneous cognitive load*

Extraneous cognitive load was assessed with the self-report mental effort rating-scale item translated from Paas (1992). The exact item was: "Bitte kreuze an, wie hoch deine Anstrengung bei der Bearbeitung der Aufgabe war.“ (Please report the amount of mental effort that you invested in working on the learning task). The scale ranged from *very, very high* (9) to *very, very low* (1). The students were asked to rate their extraneous cognitive load (indicated by mental effort) after each learning task, but not after studying the modeling example. Overall, students rated their mental effort four times. Cronbach's alpha across the four measurement-points was .83. Table 3 shows the descriptive values for the scale and measurement points.

4.2.5.3 Application-quality

Application-quality was assessed with the problem-solving parts within the learning tasks the students had to solve during the intervention. The first problem required the students to describe a one-factorial within-subjects design. The second problem was to describe a one-factorial between-subjects design. In the third problem the learners had to describe a balanced within-subjects design. The last problem was to describe a 2×2 factorial between-subjects-design. We created a separate coding scheme for each of the four tasks. All coding variables were coded as *present* (1) or *not present* (0). The coding scheme for the first and the second task consisted of 15 binary coding variables each. The coding scheme for the third task consisted of 27 binary coding variables and the coding scheme for the fourth task of 29 binary coding variables. The reason for the higher number of coding variables for the third and fourth task lies in the more complex nature of the task. For example one coding variable is:

Coding variable: indication of manipulation

- Code 1 = if the manipulation is indicated with an “x”
- Code 0 = if the manipulation is not indicated with an “x”

Two coders assessed the application-quality for each task. Cohen’s Kappa for the 86 coding variables ranged from .65 to 1 with a mean of $k = .95$. The scale for application-quality was built by computing the sum over all coding variables with a maximum of 86 points. Cronbach’s alpha for the entire scale is .93. Table 3 shows the descriptive values for the scale.

4.2.5.4 Cognitive skill acquisition

Cognitive skill acquisition was measured with a paper-and-pencil test with two part-task items and one whole-task item, where the students had to draw an empirical research design according to the scheme of Campbell and Stanley (1963). Part-task means that the students only had to fill in a part of the task, e.g. one part-task asked the students to describe the manipulation in a 2×2 design according to the scheme of Campbell and Stanley, the rest of the design was already given. The measurement took place directly after the intervention (T1) as well as in a delayed posttest four days later (T2). The coding scheme consisted of 25 variables coded as *present* (1) or *not present* (0). For example one coding variable is:

Coding variable: Constant indication of measurement in all conditions

- Code 1 = If the measurement is constantly indicated in all conditions
- Code 0 = If the measurement is not constantly indicated in all conditions

Cohen's Kappa for the 25 coding variables ranged from 0.6 to 1 with a mean of $k = 0.82$. Cronbach's alpha for the cognitive skill acquisition scale was .90 at T1 and .93 at T2. Table 3 shows the descriptive values for both scales.

Table 3. *Descriptive values for all metric variables*

	N	Minimum	Maximum	Mean	Standard Deviation
Prior knowledge	124	0	4	1.19	1.36
Extraneous cognitive load	123	1.25	9	5.21	1.71
Application-quality	125	0	78	59.97	12.71
Cognitive skill acquisition T1	126	0	25	18.85	5.25
Cognitive skill acquisition T2	147	0	25	17.42	7.28

4.2.6 Procedure

The entire session lasted 90 minutes. First, the students received a general introduction about the procedure of this session (3 minutes). Then they had max. seven minutes to answer the pretest. The intervention took 60 minutes. During this time, the students watched four modeling examples with a total length of 24 minutes (5-7 minutes each). They had 36 minutes in total to work on the four example-problem-pairs: nine minutes each for the first three pairs and ten minutes for the final pair. The students were distributed across six groups with different sequencing of tasks and with or without prompting (see section on design for more detail). After the intervention the students answered the posttest for which they had max. 20 minutes. Four days after the intervention the students answered the same posttest again at the start of a lecture in empirical research methods.

4.2.7 Statistical analysis

Alpha was set at 5% for all statistical analyses. To answer the research questions, ANOVAS and ANCOVAS (with repeated measurements) and mediation analyses were applied. A priori comparisons were conducted to analyze the difference between the three sequencing conditions. Conventions for the effect size measure partial η^2 for AN(C)OVAS are as following according to Cohen (1988): *partial* $\eta^2 = .01$ as a small effect size, *partial* $\eta^2 = .06$ as a medium effect size, and *partial* $\eta^2 = .14$ as large effect size. Conventions for correlations are $r = .10$ is a small effect size, $r = .30$ is a medium effect size and $r = .50$ is a large effect size (Cohen, 1988). The effect size d is considered as a small effect size with values of about d = .20, as a medium effect size with values of about d = .50, and as a large effect size with values above d = .80 (Cohen, 1988). The mediation analyses were conducted with PROCESS (Hayes, 2012); more specifically model 4 was used with all variables z-standardized and the bootstrap set on 50000. The metric variables were z-standardized to obtain standardized beta-values.

Prior to the analysis the sequencing factor was dummy-coded with the condition alternated-before coded as 1 and the conditions blocked-before and blocked-after coded as 0. This dummy-coding was chosen, because it was hypothesized that the alternated-before conditions would outperform the blocked-before and blocked-after sequence with respect to cognitive skill acquisition.

4.3 Results

4.3.1 Preliminary analyses: a priori differences in prior knowledge

To test any differences in prior knowledge, a 2×3 factorial ANOVA was conducted on prior knowledge with prompting and sequencing as between-subject factors. The analysis neither showed significant differences in prior knowledge between sequencing conditions, $F(2, 118) = .10$, $p = .902$, *partial* $\eta^2 < .01$, nor between prompting conditions, $F(1, 118) = .14$, $p = .905$, *partial* $\eta^2 < .01$. Also, there is no significant interaction of sequencing and prompting for prior knowledge, $F(2, 118) = 1.48$, $p = .231$, *partial* $\eta^2 = .03$. Table 4 provides means and standard deviations of prior knowledge in the six conditions.

Table 4. *Means and Standard Deviation of prior knowledge (max. 4 points)*

		Sequencing of ME and LT		
		Alternated-before	Blocked-before	Blocked-after
Prompting	With	1.44 (1.46)	1.04 (1.43)	1.23 (1.48)
	Without	0.85 (1.29)	1.44 (1.25)	1.33 (1.34)

Prior knowledge significantly correlated with cognitive skill acquisition at T1 ($r = .22$, $p = .012$), but neither with cognitive skill acquisition at T2 ($r = .13$, $p = .175$) nor with mental effort ($r = -.07$, $p = .429$) or process quality ($r = .12$, $p = .182$). Due to the significant correlation of prior knowledge and cognitive skill acquisition at T1, prior knowledge was included as covariate for all analyses that refer to cognitive skill acquisition at T1.

4.3.2 RQ1: Effects of sequencing, prompting and their interaction on cognitive skill acquisition

We conducted a 2×3 factorial repeated measurement ANCOVA with prompting and sequencing as the between-subject factors, repeated measurement of cognitive skill acquisition as the within-subject factor, and prior knowledge as a covariate.

The analysis showed a significant medium sized main effect of sequencing on cognitive skill acquisition over time, $F(2, 101) = 4.67$, $p = .012$, *partial* $\eta^2 = .09$: Students in the conditions with the alternated-before sequence had the highest mean-score in cognitive skill acquisition, whereas students with the blocked-before sequence had lower mean-score in cognitive skill acquisition and the students in the blocked-after sequence had the lowest mean-score. This pattern can be observed directly after the intervention at T1 as well as after four days at T2 (see Table 5 and Table 6). The interaction of sequencing and repeated measurement was not significant, $F(2, 101) = .43$, $p = .653$, *partial* $\eta^2 = .01$. Further, there was no significant interaction of sequencing and prompting on cognitive skill acquisition over time, $F(2, 101) = .94$, $p = .393$, *partial* $\eta^2 = .02$. The main effect of prompting on cognitive skill acquisition over time was also not significant, $F(1, 101) = 1.04$, $p = .309$, *partial* $\eta^2 = .01$. Independent from sequencing and prompting, the students significantly improved their skills significantly from T1 to T2, $F(1, 101) = 9.64$, $p = .002$, *partial* $\eta^2 = .09$. Further, there was a significant effect of prior knowledge on cognitive skill acquisition over time, $F(1, 101) = 5.90$, $p = .017$, *partial* $\eta^2 = .06$.

To reveal, whether there is an effect of sequencing on both measurement points of cognitive skill acquisition (T1 and T2), two further 2×3 ANCOVAs were conducted with prompting and sequencing as the between-subject factors, cognitive skill acquisition at T1 (respectively T2) as dependent variable, and prior knowledge as a covariate. The analysis for cognitive skill acquisition at T1 showed a significant medium sized effect of sequencing, $F(2, 117) = 4.24$, $p = .017$, *partial* $\eta^2 = .07$. Planned comparisons between the three sequencing conditions showed a significant difference between the alternated-before sequence and the blocked-before sequence in cognitive skill acquisition at T1 with a medium effect size, $t(123) = 2.34$; $p = .021$; d = 0.53 (hypothesis 1a was supported). The difference in cognitive skill acquisition at T1 between the alternated-before sequence and the blocked-after sequence was also significant with a medium effect size, $t(123) = 2.55$, $p = .012$, d = 0.60 (hypothesis 1b was supported). The small descriptive difference between the blocked-before and the blocked-after sequence in cognitive skill acquisition at T1 was not significant, $t(123) = .32$, $p = .750$, d = .07 (hypothesis 1c was not supported). The covariate prior knowledge also significantly influenced cognitive skill acquisition at T1, $F(1, 117) = 6.79$, $p = .010$, *partial* $\eta^2 = .06$. The interaction between sequencing and prompting was slightly not significant, $F(2, 117) = 2.54$, $p = .083$, *partial* $\eta^2 = .04$. Prompting did not significantly influence cognitive skill acquisition at T1, $F(1, 117) = .82$, $p = .367$, *partial* $\eta^2 < .01$.

Table 5. *Means, Estimated Means and Standard Deviation of cognitive skill acquisition at T1 (max. 25 points)*

		Sequencing of ME and LT		
		Alternated-before	Blocked-before	Blocked-after
Prompting	With	21.00 / 20.72	16.71 / 16.86	18.08 / 18.00
		(3.70)	(5.18)	(5.40)
	Without	20.52 / 20.86	19.80 / 19,59	16.68 / 16.47
		(4.19)	(4.18)	(6.28)

The analysis for cognitive skill acquisition at T2 showed slightly no significant difference between the sequencing conditions, $F(2, 101) = 2.84$, $p = .063$, *partial* $\eta^2 = .05$. The covariate prior knowledge had no significant influence on cognitive skill acquisition at T2, $F(1, 101) = 2.59$, $p = .111$, *partial* $\eta^2 = .03$. Furthermore, there was neither a significant interaction of sequencing and prompting ($F(2, 101) = .21$, $p = .814$, *partial* $\eta^2 < .01$) nor a significant effect of prompting on cognitive skill acquisition at T2, $F(1, 101) = 1.93$, $p = .168$, *partial*

$\eta^2 = .02$. The descriptive values are reported in Table 6.

Table 6. *Means, Estimated Means and Standard Deviation of cognitive skill acquisition at T2 (max. 25 points)*

		Sequencing of ME and LT		
		Alternated-before	Blocked-before	Blocked-after
Prompting	With	20.92 / 20.77 (2.10)	18.54 / 18.63 (6.05)	18.67 / 18.62 (4.08)
	Without	22.04 / 22.23 (3.14)	20.67 / 20.55 (4.24)	19.21 /19.09 (6.00)

4.3.3 RQ2: Effects of sequencing, prompting and their interaction on extraneous cognitive load

We conducted a 2×3 factorial ANOVA for extraneous cognitive load (indicated by mental effort) with prompting and sequencing as between-subject factors.

The analysis showed a significant medium sized main effect of sequencing on extraneous cognitive load, $F(2, 117) = 4.64$; $p = .012$; *partial* $\eta^2 = .07$. Extraneous cognitive load is perceived lowest in the conditions with the alternated-before sequence; higher in the conditions with the blocked-before sequence and highest in the conditions with the blocked-after sequence (see Table 7 for means and standard deviations). The difference between the alternated-before and blocked-before sequence was significant with a medium effect size, $t(120) = 2.14$. $p = .034$, d = 0.50 (hypothesis 2a was supported). The difference in extraneous cognitive load between the alternated-before and blocked-after sequence was significant with a medium to high effect size, $t(120) = 3.25$, $p = .001$, d = 0.74 (hypothesis 2b was supported). The descriptive difference between the blocked-before and blocked-after sequence was not significant, $t(120) = 1.24$, $p = .218$, d = .27 (hypothesis 2c was not supported). There is neither a significant interaction between sequencing and prompting on extraneous cognitive load, *F*(2, 117*)* = 0.12, *p* = .887, *partial* $\eta^2 < .01$, nor a significant effect of prompting, *F*(1, 117*)* = 0.18, *p* = .670, *partial* $\eta^2 < .01$.

Table 7. *Means and Standard deviation for extraneous cognitive load for each condition (max. 9 points)*

		Sequencing of ME and LT		
		Alternated-before	Blocked-before	Blocked-after
Prompting	With	4.77 (1.85)	5.36 (1.75)	5.76 (2.17)
	Without	4.43 (1.20)	5.28 (1.44)	5.79 (1.72)

4.3.4 RQ3: Effects of sequencing, prompting and their interaction on application-quality

We conducted a 2×3 factorial ANOVA with prompting and sequencing as between-subject factors and application-quality as dependent variable.

The ANOVA showed a significant medium sized main effect of sequencing on application-quality, *F*(2, 119) = 7.02, *p* = .001, *partial* $\eta^2 = .11$. The mean of application-quality is highest in the conditions with the alternated-before sequence, somewhat lower in the conditions with the blocked-before sequence and lowest in the conditions with the blocked-after sequence (see Table 8). Planned comparisons showed, that the descriptive difference between the alternated-before sequence and blocked-before sequence was not significant, *t*(122) = .58; *p* = .565; d = 0.13 (hypothesis 3a was not supported). The difference between the alternated-before and blocked-after sequence was significant with a large effect size, *t*(122) = 3.67, *p* < .001, d = 0.81 (hypothesis 3b was supported). The difference between the blocked-before and blocked-after sequence was significant with a medium effect size, *t*(122) = 3.15, *p* = .002, d = .68 (hypothesis 3c was supported). Moreover, there was neither a significant effect of prompting ($F(1, 116) = 3.27$, $p = .073$, *partial* $\eta^2 = .03$) nor a significant interaction of prompting and sequencing, $F(2, 116) = .95$, $p = .391$, *partial* $\eta^2 = .02$.

Table 8. *Means and Standard Deviations of application-quality for each condition (max. 86 points)*

		Sequencing of ME and LT		
		Alternated-before	Blocked-before	Blocked-after
Prompting	With	59.06 (14.56)	60.56 (10.60)	52.77 (16.21)
	Without	66.31 (6.61)	64.06 (12.84)	53.92 (12.78)

4.3.5 RQ4: Mediation effects of extraneous cognitive load and application- quality

I used PROCESS model 4 (Hayes, 2012) for two mediation analyses with cognitive skill acquisition – at T1 and T2 – as the criterion variable, extraneous cognitive load and application-quality as the mediators, sequencing as the predictor, and prior knowledge as a covariate.

Cognitive skill acquisition at T1. The mediation analysis showed a significant indirect effect of application-quality for the effect of sequencing on cognitive skill acquisition at T1, $\beta = 0.11$, CI_{95} [0.01, 0.31]). There was no significant indirect effect of extraneous cognitive load for the effect of sequencing on cognitive skill acquisition at T1, $\beta = 0.06$, CI_{95} [-0.07, 0.25]. Table 9 shows the beta- and p-values for the four steps in the mediation analysis.

Table 9. *Mediation analysis with cognitive skill acquisition at T1 as criterion variable*

Criterion	**Predictor**	***β***	***p***
Step I			
Cognitive skill acquisition T1	Sequencing dummy	0.58	.002
	Prior knowledge	0.23	.008
Step II			
Extraneous cognitive load	Sequencing dummy	- 0.57	.003
Step III			
Application-quality	Sequencing dummy	0.44	.023
Step IV			
Cognitive skill acquisition T1	Extraneous cognitive load	- 0.10	.259
	Application-quality	0.24	<.001
	Sequencing dummy	0.42	.025
	Prior knowledge	0.19	.027

Note. β = standardized regression-coefficient, N = 121.

Cognitive skill acquisition at T2. The mediation analysis for cognitive skill acquisition at T2 also showed a significant indirect effect of application-quality for the effect of sequencing on cognitive skill acquisition at T2 (β = 0.12, CI_{95} [0.02, 0.27]). There was no significant indirect effect of extraneous cognitive load for the effect of sequencing on cognitive skill acquisition at T2 (β = -0.03, CI_{95} [-0.08, 0.17]). Table 10 shows the beta- and p-values for the four steps in the mediation analysis.

Table 10. *Mediation analysis with cognitive skill acquisition at T2 as criterion variable*

Criterion	**Predictor**	***β***	***p***
Step I			
Cognitive skill acquisition T2	Sequencing dummy	0.31	.018
Step II			
Extraneous cognitive load	Sequencing dummy	- 0.70	<.001
Step III			
Application-quality	Sequencing dummy	0.56	.005
Step IV			
Cognitive skill acquisition T2	Extraneous cognitive load	- 0.04	.530
	Application-quality	0.21	.001
	Sequencing dummy	0.17	.217

Note. β = standardized regression-coefficient, *N* = 108.

4.4 Discussion

In the present study I investigated the effect of sequencing of modeling examples and learning tasks on cognitive skill acquisition as well as explanations for potential effects (i.e. the learning mechanisms that might mediate potential effects and the conditions under which specific sequences have specific effects). More specifically, whether a potential positive effect of sequencing modeling examples before learning tasks is mediated by extraneous cognitive load and application-quality and whether a potential positive effect of sequencing modeling examples after learning tasks is dependent on prompting self-explanations and monitoring.

4.4.1 Summary of findings

Preliminary analyses showed no significant differences in prior knowledge between the different sequencing as wells as prompting conditions. Thus, the randomization can be seen as successful, at least with respect to the distribution of prior knowledge between the conditions. With regard to RQ 1, the present study showed a medium sized effect of sequencing of modelling examples and learning tasks on cognitive skill acquisition: the alternated-before sequence leads to significant higher cognitive skill acquisition than the blocked-before and blocked-after sequence (hypotheses 1.1 and 1.2 were supported). However, there was no significant difference between the blocked-before and blocked-after sequence (hypothesis 1.3 was not supported). Furthermore, there was no significant interaction of sequencing and prompting (hypothesis 1.4 was not supported). Also there was no significant difference for sequencing in the delayed measurement of cognitive skill acquisition at T2. The effect size in this case was medium. However, as the G*Power-analysis revealed, the sample size was not optimal for showing a medium sized main-effect. Therefore, this could be a problem of statistical power. Regarding RQ 2, the present study showed a medium sized effect of sequencing of modelling examples and learning tasks on extraneous cognitive load: the alternated-before sequence leads to significant lower extraneous cognitive load than the blocked-before and blocked-after sequence (hypotheses 2.1 and 2.2 were supported). However, there was no significant difference in extraneous cognitive load between the blocked-before and blocked-after sequence (hypothesis 2.3 was not supported). Furthermore, the analyses showed a medium sized effect of sequencing on application-quality (RQ 3): the alternated-before as well as blocked-before sequence leads to significant higher application-quality than the blocked-after sequence (hypotheses 3.2 and 3.3 were supported). Between the alternated-before and blocked-before sequence, no significant difference in application-quality revealed (hypothesis 3.1 was not supported). The mediation analysis (RQ 4) showed that extraneous cognitive load did not significantly mediate the effect of sequencing (hypothesis 4.1 was not supported). However, the analyses showed a significant indirect effect of application-quality for the effect of sequencing on cognitive skill acquisition at T1 and T2 (hypothesis 4.2 was supported).

4.4.2 Theoretical implications

In the light of research on sequencing example-based learning activities, the results of this study are in line with the studies of Reisslein and colleagues (2006), Van Gog and colleagues (2011) and as Leppink and colleagues (2014), which provided evidence in favor of an alternated-before sequence compared to an alternated-after sequence with respect to cognitive skill acquisition. In addition, this study confirms the findings of Trafton and Reiser (1993), which showed that students who learned with an alternated-before sequence had a better performance than students who learned with the blocked-before sequence. From a theoretical perspective on cognitive skill acquisition (e.g., Van Lehn, 1996) an alternated-before sequence of modelling examples and learning tasks should provide easier retrieval, mapping and application of the specific features of the example to the learning task compared to a blocked-before or -after sequence. The findings of this study provide support for Van Lehn’s (1996) model, because the alternated-before sequence led to higher application-quality in the sense of better retrieval, mapping and application during the problem solving process, which partially mediated the effect of sequencing on cognitive skill acquisition. However, since application-quality only partially mediated the sequencing effect there seem to be further mechanisms which play a mediating role. According to Banduras theory on observational learning (1986), the alternated-before sequence should further have the advantage of monitored enactment, which is comparison, detection and correction of discrepancies between the model's and the learners' behavior (Bandura, 1991, Carroll & Bandura, 1990). In an alternated-before sequence the learners can directly compare their solution method with the modeling example and focus their attention on those critical aspects in the next modeling example the students were not sure about during problem solving. These attentional processes within monitored enactment might be further learning mechanisms that play a mediating role for sequencing effects on cognitive skill acquisition. Future research could investigate them by asking the participants, who learn with different sequences, after each modeling example and learning task, how often they compared their behavior to the model, and how often they detected and corrected discrepancies between the model's solution and their own. An open question prior to the present study was the stability of short-term sequencing effects and whether there would be an increase in skill acquisition over time dependent on sequencing. This study could not show an increase in cognitive skill acquisition over time dependent on sequencing, although the statistical power was high enough to show a medium sized interaction effect.

With respect to the effect of sequencing on application-quality, Van Gog (2011) also showed that an alternated-before sequence led to higher application-quality than an alternated-after sequence, but in their study application-quality was not analyzed as a potential mediator. The present study did include application-quality as mediator, which is apparently more precise to do so, because it allows for analyzing the assumed underlying mechanism of sequencing effects. Extraneous cognitive load was not a significant mediator for the effect of sequencing on cognitive skill acquisition in the present study, even though the alternated-before sequence reduced extraneous cognitive load compared to the blocked-before and -after sequence. The latter is in line with findings of Van Gog and colleagues (2011), which also showed that an alternated-before sequence led to lower extraneous cognitive load compared to an alternated-after sequence. Nevertheless, they did not analyze mediating effects of extraneous cognitive load. From a theoretical perspective (e.g., Sweller, 2010), extraneous cognitive load should be detrimental for learning and thus play a mediating role. However, the study of Leppink and colleagues (2014) did not show a significant correlation of extraneous cognitive load and test performance. It seems that application-quality as a content-wise measure is more important for predicting cognitive skill acquisition than extraneous cognitive load.

Furthermore, in the present study I assumed that prompting self-explanation and monitoring processes could interact with the sequencing of modeling examples and learning tasks. From a theoretical perspective, prompting self-explanations and monitoring was supposed to support a greater awareness of knowledge gaps which should help to better integrate the subsequent instruction with modeling examples (Loibl & Rummel, 2014; Loibl, Roll, & Rummel, 2017). When learners are aware of their knowledge gaps, the focus of their attention should be on the critical aspects in the modeling examples that are provided after the learning tasks, especially compared with learners, who did not monitor their knowledge gaps. However, the present study did not show that an effect of sequencing is dependent on prompting self-explanations and monitoring. One explanation might be too little statistical power to show an effect. Another explanation could be that an interaction between sequencing and prompting only takes place when the learners also have the possibility to regulate their learning behavior. In the conditions with the blocked-after sequence, the leaners could have focused their attention on the critical aspects, but then had no opportunity to regulate their problem solving behavior either by solving a new problem or by correcting the previous solved problem. Some support for the feasibility of this explanation is provided by a study of Hübner and colleagues (2007) on prompting (meta-)cognitive processes when writing learning protocols, which showed that the possibility to self-regulate is important for test performance. Future research could address the role of regulation opportunities for a potential interaction of sequencing and prompting with regard to cognitive skill acquisition in a between-within-subjects design, where sequencing and prompting are varied as between-subjects factors and the possibility for regulation as within-subjects factor.

4.4.3 Limitations

One limitation of the present study is that I only compared the alternated-before sequence to blocked-after, but not with an alternated-after sequence. Thus, we do not know whether there still is a superiority of the alternated-before sequence when compared to alternated-after. Study II investigates this by analyzing cognitive skill acquisition in conditions with an alternated-before sequence as opposed to conditions with an alternated-after sequence.

A further limitation might be that the learning task consisted of solved example-problem pairs that were combined with modeling examples in the three different sequencing conditions. A reason for this was to provide two examples for the same problem category because this can support abstracting the structural features of the examples instead of surface-features (Renkl, 2014). However, it is not clear, if the findings of the present study are valid for treatments without solved examples but only with modeling examples, because in the first case different processes might occur. Having a solved example beneath the problem solving task might seduce the learners to use max-analogy strategies, that is copying as much as possible as opposed to min-analogy which is trying to solve the problem on one's own before referring to the example (Muldner & Conati, 2010). For this reason, Study II did not provide solved examples beneath the problem solving task as learning tasks.

Another limitation concerns the unidimensional measurement of extraneous cognitive load with Paas' (1992) mental effort rating scale and that mental effort is not interpreted as reflecting the same type of cognitive load over different studies (De Jong, 2010). This is problematic because it concerns the validity of the instrument, as well as the validity of the measure and of the interpretation. Future research should use an instrument that takes the multidimensional structure into account, such as the scales of Leppink and colleagues (2013, 2014) or the scales of Opfermann (2008).

5 Effects of sequencing and learning tasks on cognitive skill acquisition (Study II)

5.1 Research Questions

In the present study I investigate (1) the effect of sequencing of modeling examples and learning tasks on cognitive skill acquisition (GRQ I), and (2) which explanations might clarify potential effects of sequencing on cognitive skill acquisition (GRQ II). More specifically, whether extraneous cognitive load and application-quality might meditate a potential positive effect of sequencing modeling examples before learning tasks on cognitive skill acquisition and whether a potential positive effect of sequencing modeling examples after learning tasks is dependent on the type of learning task. To investigate these two General Research Questions (see Chapter 3), four specific research questions and corresponding hypotheses are formulated. The General Research Question I is addressed by RQ 1. The General Research Question II is specified by all four research questions of the present study.

RQ1: To what extent does the sequencing of modeling example and learning tasks, the type of learning task and the interaction thereof, have an effect on cognitive skill acquisition?

An alternated-before sequence of modeling examples and learning tasks should lead to higher degrees of cognitive skill acquisition (Van Gog et al., 2011; Leppink et al., 2014), because it should allow for mapping and application of the specific example features to the learning task as opposed to an alternated-after sequence, when no features or principles are provided to map and apply (Van Lehn, 1996).

- ***H1.1:*** An alternated-before sequence of modeling example and learning tasks leads to higher cognitive skill acquisition than an alternated-after sequence (Van Gog et al., 2011; Leppink et al., 2014).

However, Alfieri and colleagues' (2013) meta-analysis on case comparison showed that providing principles after case comparison is more effective with respect to conceptual and procedural knowledge compared to providing no principles or before comparing cases, because it might better support the modification of an abstract schema (Holyoak, 2012). They also highlight, that case comparison as a learning task may be especially helpful to provide 'reparation for future learning' (Schwartz & Bransford, 1998; Schwartz et al., 2011) in order to benefit from instruction such as modeling examples (Alfieri et al., 2013). Thus, an effect of sequencing modeling examples afterwards might depend on the type of learning task.

- ***H1.2:*** The effect of sequencing is dependent on the type of learning task. This means, the effect might be reduced or inversed, when the learning task is that of comparing examples as opposed to problem solving (Alfieri et al., 2013).

Observational learning should provide the advantage that the learners can directly compare their solution method with the modeling example and focus their attention in the next modeling example on the critical aspects they were not sure about during problem solving (e.g., Bandura, 1986). Thus, providing problem solving and comparing with the example as learning task should lead to higher cognitive skill acquisition than problem solving or comparing examples as learning tasks.

- ***H1.3:*** Problem solving and comparing its result with examples as a learning task leads to higher cognitive skill acquisition than problem solving alone or only comparing examples as learning tasks (Bandura, 1986).

RQ2: To what extent does the sequencing of modeling example and learning tasks, the type of learning task and the interaction thereof, have an effect on extraneous cognitive load during learning?

Learners who first receive a modeling example and then solve a learning task can apply search-by-analogy strategy during problem-solving and thus do not have to rely on means-ends-analysis, a weak problem-solving strategy that requires a lot of working memory capacity (e.g., Sweller & Cooper, 1985; Renkl, 2014; see Chapter 2.1.1). Therefore, an alternated-before sequence should lead to lower extraneous cognitive load than alternated-before sequence (Van Gog et al., 2011).

- ***H2:*** An alternated-before sequence of modeling example and learning task leads to lower extraneous cognitive load than an alternated-after sequence (Van Gog et al., 2011).

RQ3: To what extent does the sequencing of modeling example and learning tasks, the type of learning task and the interaction thereof, have an effect on application-quality during learning?

Providing modeling examples and learning tasks as alternated-before sequence can allow for mapping and application of the specific example features to the learning task (Van Lehn, 1996). When modeling examples are provided alternated after the learning tasks no features or principles are provided to map and apply. Thus, an alternated-before sequence should lead to higher application quality than an alternated-after sequence (Van Gog, 2011).

- ***H3:*** An alternated-before sequence of modeling example and learning tasks leads to higher application-quality than an alternated-after sequence (Van Gog, 2011, Van Lehn, 1996).

RQ4: To what extent do extraneous cognitive load and application-quality during learning mediate an effect of sequencing on cognitive skill acquisition?

From a theoretical perspective on cognitive load (e.g. Sweller, 2010), extraneous cognitive load should be detrimental for learning and thus should play a mediating role. Research on the role of extraneous cognitive load in example-based learning in general provides mixed results regarding the mediating value of extraneous cognitive load (see Chapter 2.2.1). Even though studies on different sequencing strategies in example-based learning (e.g., Leppink et al., 2014; Van Gog et al., 2011) assume extraneous cognitive load as a mechanism, none of the studies on sequencing (see Chapter 2.3) investigated extraneous cognitive load as mediator.

Similarly, none of the studies on sequencing (see Chapter 2.3) analyzed application-quality as mediator. However, from a theoretical perspective on cognitive skill acquisition (e.g., Van Lehn, 1996), application-quality should play a mediating role because retrieval, mapping and application of specific features of the example to the problem solving task are considered important learning processes in early phases of cognitive skill acquisition (see Chapter 2.1).

- ***H4.1***: A potential effect of sequencing is mediated by extraneous cognitive load (Sweller, 2010).
- ***H4.2***: A potential effect of sequencing is mediated by application-quality (Van Lehn, 1996).

5.2 Methods

5.2.1 Sample

One hundred and forty-five German educational science students in their first semester participated. On average they were 21 years old (*M* = 21.08; *SD* = 3.90) and 91.7% of the participants were female. A G*Power analysis (Faul et al., 2007) for a medium sized interaction effect (*partial* η^2 = .06) of sequencing and type of learning task on cognitive skill acquisition in an ANCOVA with prior knowledge as covariate with a power of 80% indicates an optimal sample size of 155 subjects. Thus, the actual sample size of 145 subjects is nearly optimal.

5.2.2 Design

We conducted a 2×3 factorial design with two between-subject factors. The first between-subjects factor is the sequencing of modeling examples and learning tasks (alternated-before vs. alternated-after). The second between-subjects factor is the type of learning task (problem-solving vs. comparing examples vs. problem-solving and comparing with examples). Prior knowledge was measured before the intervention as a control variable. Extraneous cognitive load and application-quality were assessed three times after completing each learning task. Cognitive skill acquisition was measured directly after the intervention. The students were randomly assigned to one of six conditions. Table 11 shows the distribution of participants across the six conditions.

Table 11. *Distribution of participants across conditions*

		Type of Learning task		
		PS	Comparing examples	PS + comparing with examples
Sequencing of ME and LT	Alternated-before	22	27	25
	Alternated-after	21	23	24

Note. PS stands for problem-solving, ME for modeling examples and LT for learning tasks.

5.2.3 Learning Environment

The objective of the learning environment was to teach students how to describe different between- and within-subject designs according to the scheme of Campbell and Stanley (1963). The declarative knowledge was covered in a lecture one month before the study. The application of this knowledge was the goal of the intervention study, which was implemented in a seminar session one month after the lecture. During this session, the students worked individually on three modeling examples and three learning tasks. For standardization purposes, the modeling examples were provided as digital videos. Each modeling example was 4.5 to 6 minutes long. In total the three modeling videos lasted 16 minutes. All modeling examples consisted of problem-solving tasks, where the model described a design according to the scheme of Campbell and Stanley (1963) with respect to specific research questions. The task was shown in a power point slide and the model solved the task by demonstrating the procedure step by step on the power point presentation. The model also externalized the rationale behind the procedure. The first modeling example demonstrated how to describe a one-factorial between-subjects design and took 6 minutes. The second modeling example showed how to describe a 2×3 factorial between-subjects design and took 5.5 minutes. The third example demonstrated how to describe a balanced within-subjects design (4.5 minutes). The complexity of the modeling examples increased from example to example and their order was the same in all conditions, thus intrinsic cognitive load with respect to element interactivity was kept equal between the conditions. The learning tasks were printed as a learning booklet. The structure of each learning tasks was the same as in the corresponding modeling example; only the surface features, namely the cover story, differed. The learners had 9 minutes to work on each learning task; in total they had 27 minutes. After the first modeling example and learning task, the learners of all conditions received one practice problem.

5.2.4 Independent variables and manipulations

Sequencing of modelling examples and learning tasks

The factor sequencing of modeling examples and learning tasks contains two manipulations: the alternated-before and alternated-after sequence. Students in the conditions with the alternated-before sequence received the first modeling example and then worked on the first learning task, and then they received the second modeling example and after this the second learning task and so on. The learners in the conditions with the alternated-after sequence received each modeling example directly after they worked on the associated learning task. Figure 7 illustrates the manipulation of the sequence.

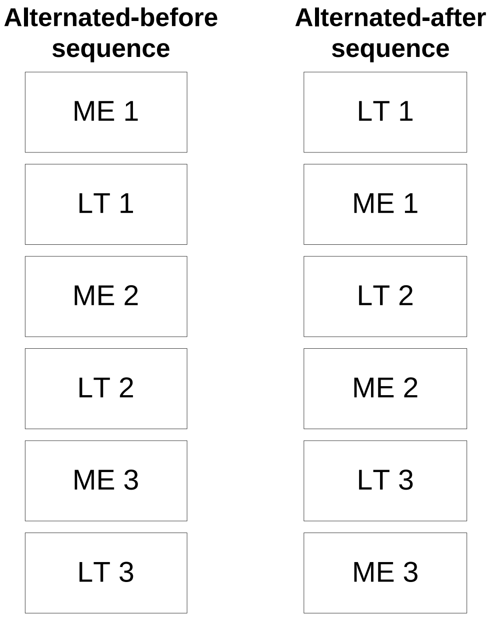
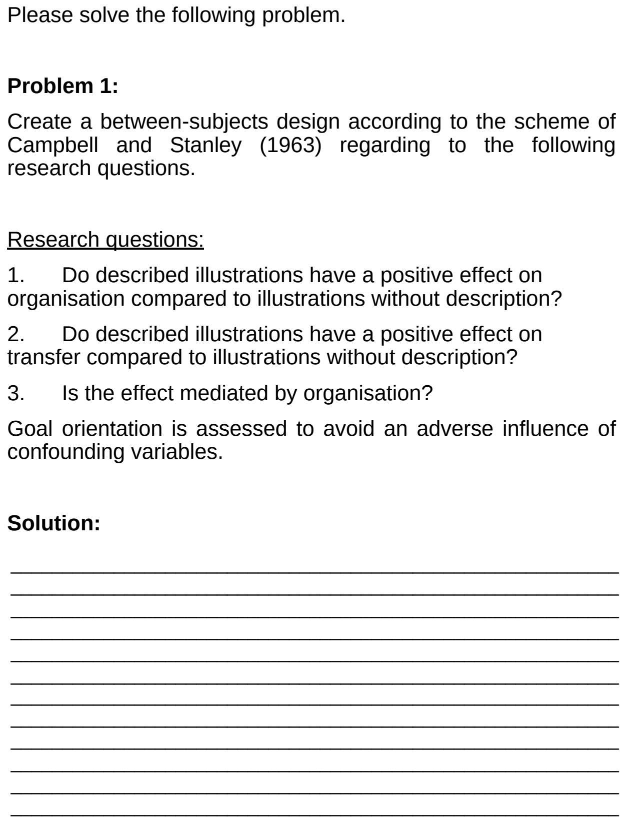


Figure 7. Overview of the sequencing manipulation

Note. ME stands for modeling example, LT for learning task.

Type of learning task

The factor type of learning task contains thee manipulations: problem-solving, comparing examples, and problem-solving and comparing with examples. The manipulation was integrated in the learning booklet. In the conditions with the learning task of problem-solving, the students had to solve three problem solving tasks (see Figure 8).



Please solve the following problem.

Problem 1:

Create a between-subjects design according to the scheme of Campbell and Stanley (1963) regarding to the following research questions.

Research questions:

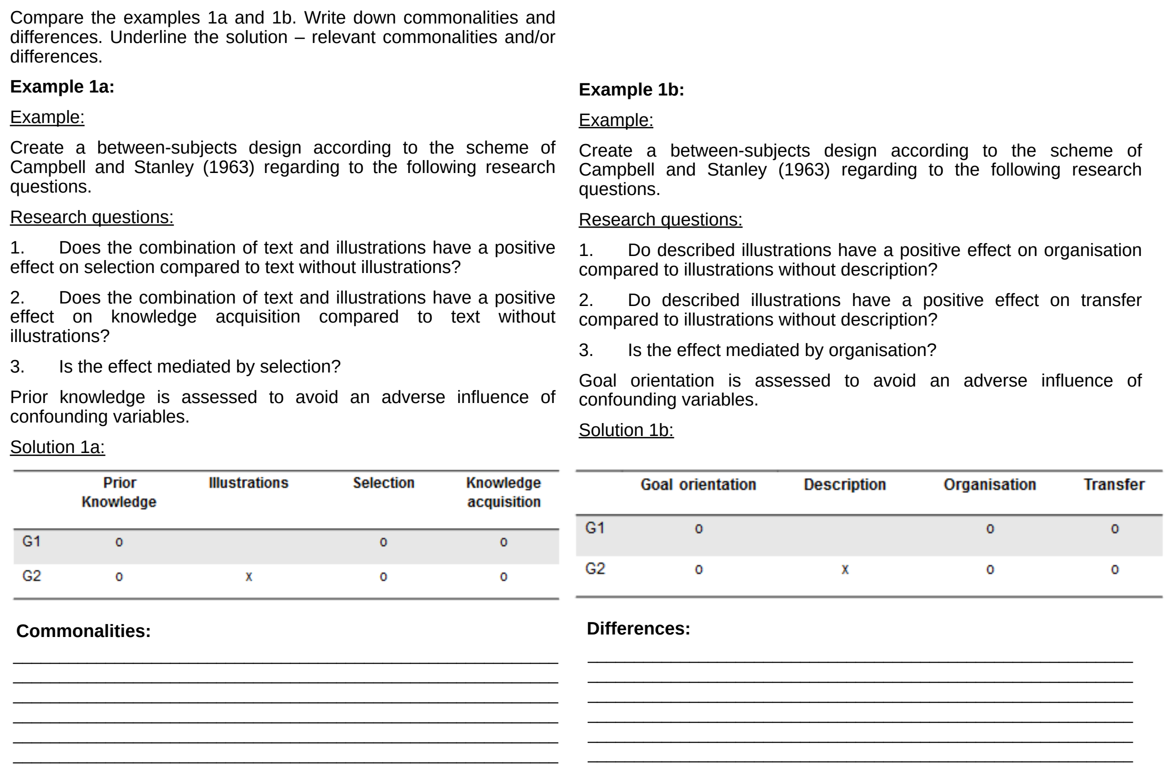
1. Do described illustrations have a positive effect on organisation compared to illustrations without description?
2. Do described illustrations have a positive effect on transfer compared to illustrations without description?
3. Is the effect mediated by organisation?

Goal orientation is assessed to avoid an adverse influence of confounding variables.

Solution:

Figure 8. Example for the type of learning task in the conditions with problem solving

Students in the conditions with the learning task of comparing examples had to compare two solved examples by writing down commonalities and differences. They should also underline solution-relevant features (see Figure 9). Whereas one solved example was the same as the modeling example, the other one consisted of the same problem, the problem-solving students had so solve.



Compare the examples 1a and 1b. Write down commonalities and differences. Underline the solution – relevant commonalities and/or differences.

Example 1a:

Example:

Create a between-subjects design according to the scheme of Campbell and Stanley (1963) regarding to the following research questions.

Research questions:

1. Does the combination of text and illustrations have a positive effect on selection compared to text without illustrations?
2. Does the combination of text and illustrations have a positive effect on knowledge acquisition compared to text without illustrations?
3. Is the effect mediated by selection?

Prior knowledge is assessed to avoid an adverse influence of confounding variables.

Solution 1a:

	Prior Knowledge	Illustrations	Selection	Knowledge acquisition
G1	o		o	o
G2	o	x	o	o

Commonalities:

Example 1b:

Example:

Create a between-subjects design according to the scheme of Campbell and Stanley (1963) regarding to the following research questions.

Research questions:

1. Do described illustrations have a positive effect on organisation compared to illustrations without description?
2. Do described illustrations have a positive effect on transfer compared to illustrations without description?
3. Is the effect mediated by organisation?

Goal orientation is assessed to avoid an adverse influence of confounding variables.

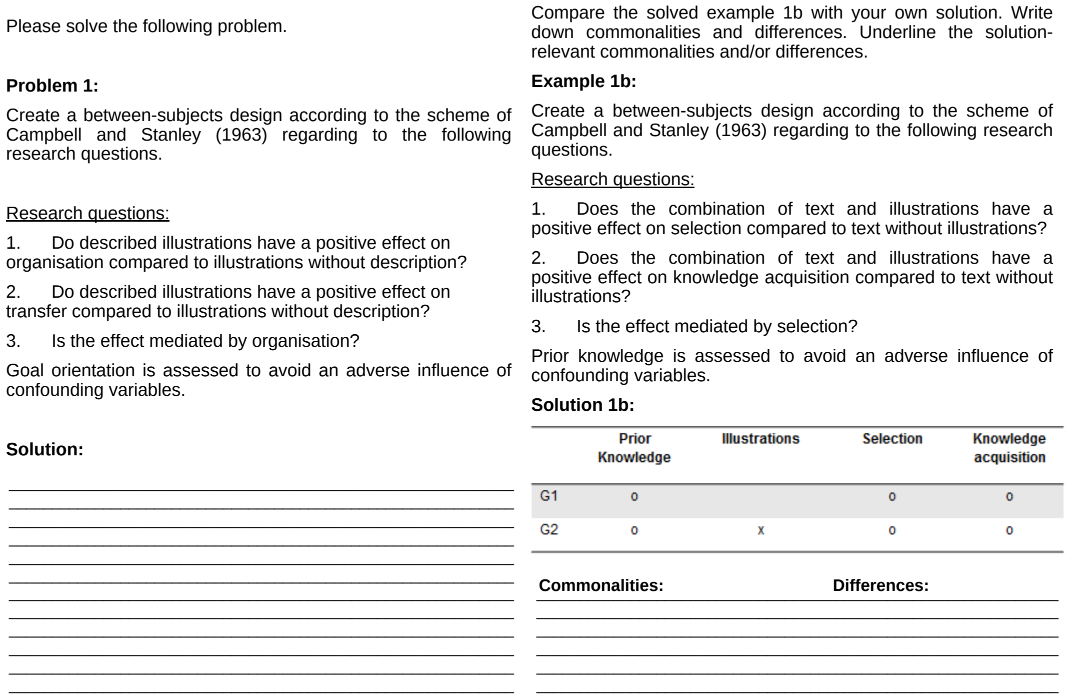
Solution 1b:

	Goal orientation	Description	Organisation	Transfer
G1	o		o	o
G2	o	x	o	o

Differences:

Figure 9. Example for the type of learning task in the conditions with comparing examples

The students in the third condition first had to solve the same problem as the problem-solving students and then had to compare their solution with the solved example (see Figure 10), which is the same example as presented to the students in the comparing examples condition and as it was presented in the modeling example.



Please solve the following problem.

Problem 1:

Create a between-subjects design according to the scheme of Campbell and Stanley (1963) regarding to the following research questions.

Research questions:

1. Do described illustrations have a positive effect on organisation compared to illustrations without description?
2. Do described illustrations have a positive effect on transfer compared to illustrations without description?
3. Is the effect mediated by organisation?

Goal orientation is assessed to avoid an adverse influence of confounding variables.

Solution:

Compare the solved example 1b with your own solution. Write down commonalities and differences. Underline the solution-relevant commonalities and/or differences.

Example 1b:

Create a between-subjects design according to the scheme of Campbell and Stanley (1963) regarding to the following research questions.

Research questions:

1. Does the combination of text and illustrations have a positive effect on selection compared to text without illustrations?
2. Does the combination of text and illustrations have a positive effect on knowledge acquisition compared to text without illustrations?
3. Is the effect mediated by selection?

Prior knowledge is assessed to avoid an adverse influence of confounding variables.

Solution 1b:

	Prior Knowledge	Illustrations	Selection	Knowledge acquisition
G1	o		o	o
G2	o	x	o	o

Commonalities:

Differences:

Figure 10. Example for the type of learning task in the conditions with problem solving and comparing with examples

The combinations of the two factors (i.e. sequencing and type of learning tasks) resulted in six conditions (see Figure 11). In condition 1, the students received the alternated-before sequence with problem-solving as learning task: they received the first modeling example and then worked on the first problem-solving task. This procedure repeated two times. The students in condition 2 received the alternated-after sequence with problem-solving as learning task: they worked on the first problem-solving task and then received the first modeling example. This procedure repeated two times. In condition 3, the students received the alternated-before sequence with comparing examples as learning task: they received the first modeling example and then compared two solved examples by searching for commonalities and differences, whereas one of the examples was the example from the modeling and the other example had the same story as the problem-solving task in the condition 1 and 2. This was repeated two times. Students in condition 4 received the alternated-after sequence with comparing examples as learning task: the learners first compared the two solved examples and then received the first modeling example. This repeated two more times. In condition 5, the students learned in the alternated-before sequence with problem-solving and comparing with examples as learning task: they were provided with the first modeling examples. After this, they worked on the first problem-solving task and then compared their solution with the solved example (i.e. the modeling example's solution) by searching for commonalities and differences. This procedure repeated two times. The solved example is the same as one example from the conditions 3 and 4 with comparing examples as learning task. Students in condition 6 learned in the alternated-after sequence with problem-solving and comparing with examples as learning task: they solved the first problem and compared their solution with the modeling example's solution and then received the modeling example. Then they solved the second problem and so on.

Condition 1	**Condition 2**	**Condition 3**	**Condition 4**	**Condition 5**	**Condition 6**
ME 1	PS 1	ME 1	Comparing Examples 1	ME 1	PS + comparing with example 1
PS 1	ME 1	Comparing Examples 1	ME 1	PS + comparing with example 1	ME 1
ME 2	PS 2	ME 2	Comparing Examples 2	ME 2	PS + comparing with example 2
PS 2	ME 2	Comparing Examples 2	ME 2	PS + comparing with example 2	ME 2
ME 3	PS 3	ME 3	Comparing Examples 3	ME 3	PS + comparing with example 3
PS 3	ME 3	Comparing Examples 3	ME 3	PS + comparing with example 3	ME 3

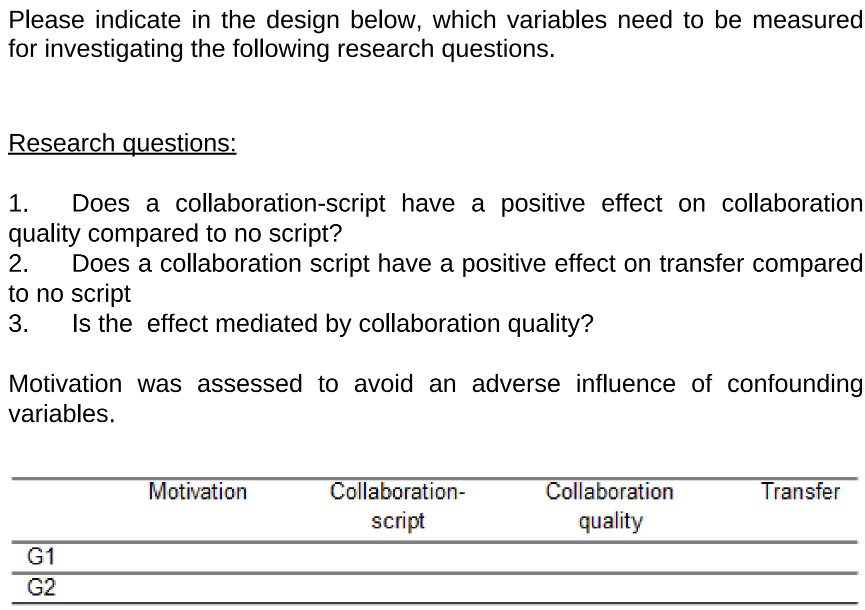
Figure 11. Visualization of the six different conditions

Note. ME stands for modeling example, PS for problem-solving.

5.2.5 Measures

5.2.5.1 *Prior knowledge*

Prior knowledge was assessed before the learning phase with four part-task problems, where the students had to apply only some features of the to be learned skills. In the first task for example, they had to examine variables in a text and determine the independent and dependent variable. The second task was to determine how many conditions one would need for testing a specific research question. The third task, which is shown in Figure 12, asked the students to indicate the measurements, whereas the fourth task asked the students to indicate the manipulation.



Please indicate in the design below, which variables need to be measured for investigating the following research questions.

Research questions:

1. Does a collaboration-script have a positive effect on collaboration quality compared to no script?
2. Does a collaboration script have a positive effect on transfer compared to no script
3. Is the effect mediated by collaboration quality?

Motivation was assessed to avoid an adverse influence of confounding variables.

	Motivation	Collaboration-script	Collaboration quality	Transfer
G1				
G2				

Figure 12. Example for a part-task problem in the pretest

The coding scheme to analyze the four part-task problems consisted of 19 binary variables coded 1 or 0. The scheme for the first part-task consists of six coding variables, two coding variables for the second part-task, eight coding variables for the third part-task and four coding variables for the fourth part-task.

One coding variable is for example:
Measurement of dependent variable is indicated with “o”

- Code 1 = If the dependent variable “transfer” is indicated with “o”
- Code 0 = If the dependent variable “transfer” is not indicated with “o”

Cohens Kappa for the 19 coding variables ranged from .64 to 1 with a mean score of .90. The scale has a Chronbach’s alpha of .77. The scale was built by computing the sum with a maximum of 19 points.

5.2.5.2 *Extraneous cognitive load*

Extraneous cognitive load was assessed with three self-report 5-point Likert-scale items adapted from Opfermann (2008). Figure 13 provides an example.

“How easy or difficult do you find working with the learning material.”

Very easy	Rather easy	Neither nor	Rather difficult	Very difficult
☐	☐	☐	☐	☐

Figure 13. Example-item of the scale for measuring extraneous cognitive load

Overall, students rated their extraneous cognitive load three times. The students did not rate extraneous cognitive load after studying the modeling examples, but only after working on each of the three learning tasks. The reliability for the three items over the three measurement-points is good (Cronbach’s Alpha = .89). The extraneous cognitive load scale was computed by calculating the mean score for the three items over the three measurement-points. Table 12 shows the descriptive values for the scale.

5.2.5.3 *Application-quality*

Application-quality was assessed for those conditions that learned with problem solving tasks. Those students worked on three problem solving tasks during the intervention phase. We created a coding scheme for each of the three tasks. All variables were coded binary as *present* (1) or *not present* (0).

The coding scheme for the first task consists of 26 binary coding variables, for the second task of 33 binary coding variables, and for the third task of 24 binary coding variables. For example one coding variable is:

> "Measurement of control variable is indicated with 'o'
> - Code 1 = If the control variable "goal orientation" is indicated with 'o'
> - Code 0 = If the control variable "goal orientation" is not indicated with 'o'"

Cohen's Kappa for the 83 coding variables ranged from .73 to 1 with a mean of $k = .99$. The application quality scale was built by computing the sum over all coding variables with a maximum of 83 points. Cronbach's alpha for the entire scale is .97. Table 12 shows the descriptive values for the scale.

5.2.5.4 *Cognitive skill acquisition*

Cognitive skill acquisition was measured with a paper-and-pencil test with two part-tasks and one whole-task, where the students had to describe a design according to the scheme of Campbell and Stanley (1963). Part-task means that the students only had to fill in a part of the task, e.g. students only had to indicate the manipulation in a 2×2 design according to the scheme of Campbell and Stanley, the rest of the design was already given. The second part-task asked the students to determine variables from a text, e.g. identify the independent and dependent variable. The whole task asked the students to describe a balanced within-subjects design according to the scheme of Campbell and Stanley (1963). The measurement took place directly after the intervention. The coding scheme consisted of 29 variables with the binary codes *present* (1) and *not present* (0). For analyzing the first part-task ten coding variables were used, for the second part-task two coding variables were applied. Finally, 17 coding variables were used for analyzing the whole-task. For example one coding variable is:

“Column for independent variable is labeled correctly (0/1)

- Code 1 = If the column for the independent variable is labeled with the name of the independent variable “type of instructional support”
- Code 0 = If the column for the independent variable is not labeled with the name of the independent variable “type of instructional support”

Cohen’s Kappa for the 29 coding variables ranged from .63 to 1 with a mean of k = .95. The scale for cognitive skill acquisition was built by computing the sum over all coding variables with a maximum of 29 points. Cronbach’s alpha for the entire scale is .81. Table 12 shows the descriptive values.

Table 12. *Descriptive values for all metric variables*

	N	Minimum	Maximum	Mean	Standard Deviation
Prior knowledge	145	1.00	19.00	8.67	3.49
Extraneous cognitive load	141	1.44	5.00	2.83	0.69
Application-quality	90	2.00	82.00	47.03	19.60
Cognitive skill acquisition	143	8.00	27.00	21.01	4.27

5.2.6 Procedure

The entire session lasted about 80 minutes. First, the students received a general introduction about the procedure of this session (5 minutes). Then they had maximum eight minutes to answer the pretest. The intervention phase took 43 minutes. During this time, the students first received an explanation about the tasks (2 minutes). Next they received three modeling examples with a total length of 16 minutes; 4.5 to 6 minutes each. The students had 27 minutes in total to work on the three learning tasks, nine minutes for each task. After the intervention the students answered the posttest for which they had maximum 12 minutes, whereas the students had a specific time for each test-item ranging from three to six minutes.

5.2.7 Statistical analysis

An alpha-error-level of 5 % was used. Conventions for interpreting effect sizes (*partial* η^2, d and *r*) are the same as in study one (Cohen, 1988, see Chapter 4.2.7). When performing the ANOVA for RQ1, RQ2 and RQ3, a defined model with an interaction term with sequencing and prior knowledge was computed. For RQ4, a mediated moderation model was performed with PROCESS model 5 and bootstrap-analysis with 50000 samples (Hayes, 2012). Prior to this analysis all metric variables were z-standardized to obtain standardized beta-values.

5.3 Results

5.3.1 Preliminary analyses

To test for any differences in prior knowledge, a 2×3 factorial ANOVA was conducted on prior knowledge with sequencing and type of learning task as between-subject factors.

The analysis neither showed significant differences in prior knowledge between sequencing conditions, $F(1, 139) = .12$, $p = .728$, *partial* $\eta^2 < .01$, nor between the type of learning task-conditions, $F(2, 139) = .48$, $p = .622$, *partial* $\eta^2 < .01$. Also, there was no significant interaction of sequencing and type of learning task for prior knowledge, $F(2, 139) = .25$, $p = .778$, *partial* $\eta^2 < .01$. Table 13 shows the descriptive values.

Table 13. *Means and Standard Deviation of prior knowledge (max. 19 points)*

	Alternated-before sequence			Alternated-after sequence		
	Problem solving ($n = 22$)	Comparing examples ($n = 27$)	Problem solving + comparing with example ($n = 25$)	Problem solving ($n = 23$)	Comparing examples ($n = 24$)	Problem solving + comparing with example ($n = 24$)
Prior knowledge	8.41 (2.22)	8.59 (3.43)	8.68 (3.91)	8.13 (3.27)	9.33 (4.45)	8.83 (3.44)

Prior knowledge is significantly and positively correlated with cognitive skill acquisition ($r = .31$, $p < .001$), negatively correlated with extraneous cognitive load ($r = -.17$, $p = .041$), but not significantly correlated with application-quality ($r = .09$, $p = .426$). However, testing the regression slopes for the correlation of prior knowledge and cognitive skill acquisition revealed different slopes for the alternated-before and alternated-after sequence (see Figure 14). Thus, it was decided to include prior knowledge as an interaction term in subsequent analyses.

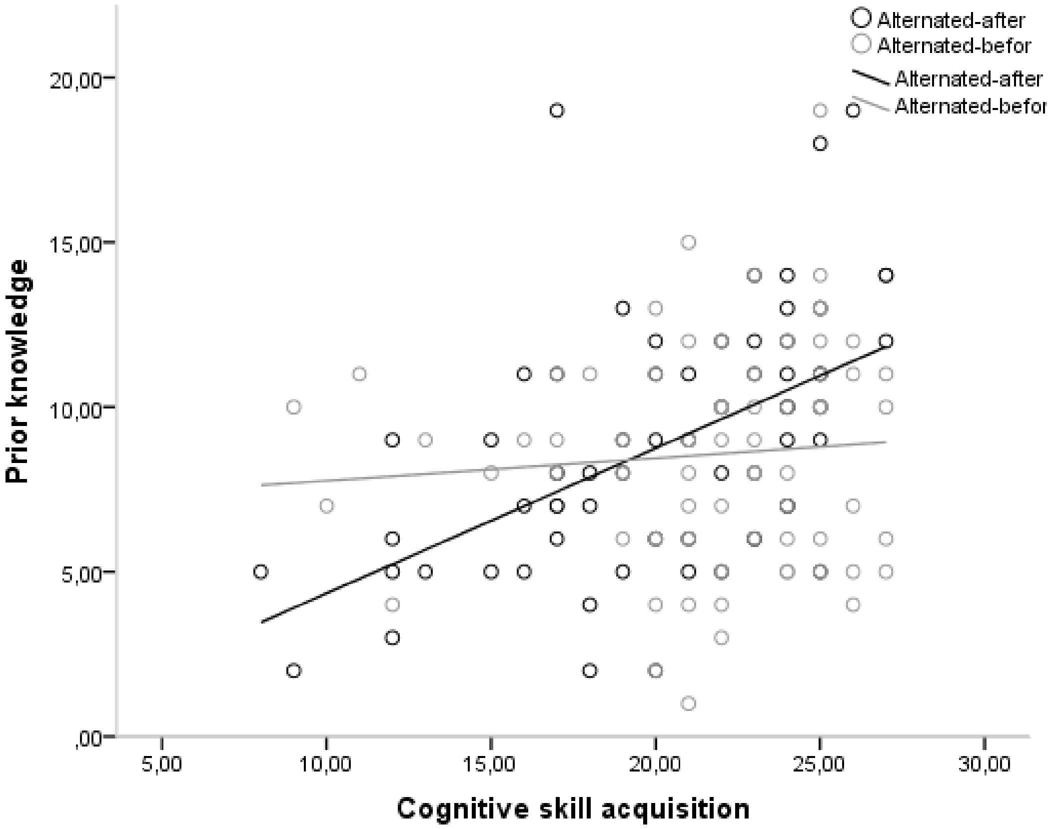


Figure 14. Visualization of different regression slopes for correlations of prior knowledge and cognitive skill acquisition in the two sequencing conditions.

5.3.2 RQ1: Effects of sequencing, type of learning task and the interaction thereof on cognitive skill acquisition

We conducted a 2×3 factorial ANOVA with an adapted interaction model with prior knowledge, sequencing and type of learning task as between-subject factors and cognitive skill acquisition as the dependent variable.

The analysis showed a significant interaction effect of sequencing and prior knowledge with a small to medium effect size on cognitive skill acquisition, $F(1, 131) = 6.14$, $p = .015$, *partial* $\eta^2 = .05$. A further simple main effect analysis showed a significant difference in sequencing in favor for the alternated-before sequence for learners with low prior knowledge, $F(1, 131) = 11.39$, $p = .001$, *partial* $\eta^2 = .08$, but no significant difference in sequencing for learners with high prior knowledge, $F(1, 131) = .31$, $p = .576$, *partial* $\eta^2 < .01$. For visualization purposes, prior knowledge was divided into two categories via median split. Figure 15 shows the interaction of prior knowledge and sequencing.

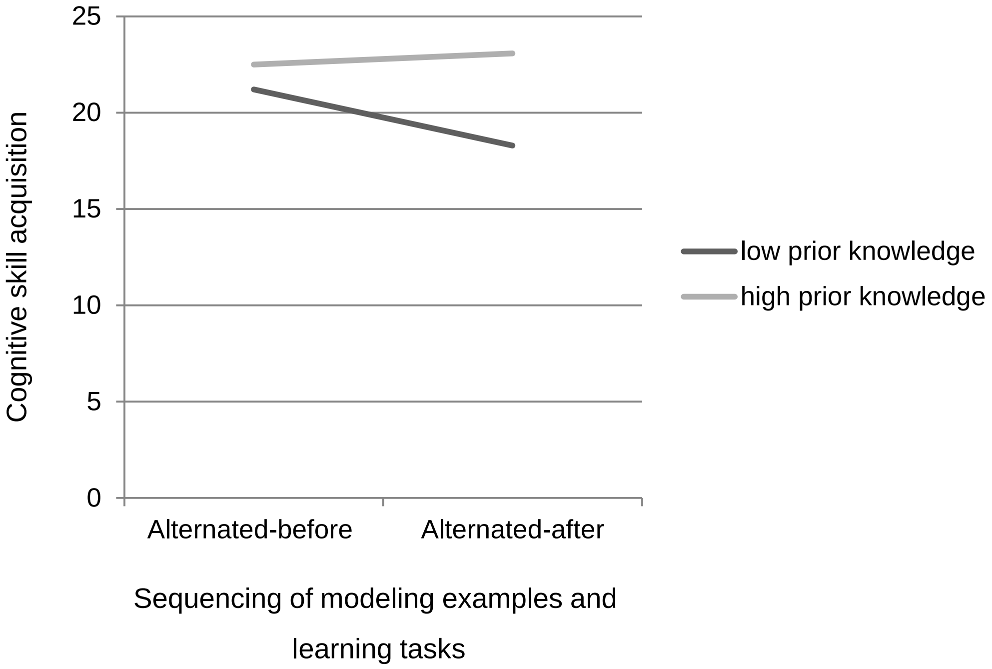


Figure 15. Visualization of the interaction effect of prior knowledge and sequencing, with prior knowledge divided into two categories via median split

The analysis also showed a significant main effect of sequencing on cognitive skill acquisition, $F(1, 131) = 9.96$, $p = .002$, *partial* $\eta^2 = .07$. However, due to the interaction with prior knowledge, this effect is limited to learners with low prior knowledge. There is also a significant medium to large sized effect of prior knowledge on cognitive skill acquisition, $F(1, 131) = 17.19$, $p < .01$, *partial* $\eta^2 = .12$. Furthermore, there is no significant interaction of sequencing and type of learning task on cognitive skill acquisition, $F(2, 131) = .04$, $p = .96$, *partial* $\eta^2 < .01$, no significant interaction of sequencing, type of learning task and prior knowledge on cognitive skill acquisition, $F(2, 131) = .02$, $p = .98$, *partial* $\eta^2 < .01$, no significant interaction of type of learning task and prior knowledge on cognitive skill acquisition, $F(2, 131) = 1.32$, $p = .27$, *partial* $\eta^2 = .02$, and no significant difference in cognitive skill acquisition for the type of learning task, *F*(2, 131) = .27, *p* = .76, *partial* $\eta^2 < .01$. The descriptive values are reported in Table 14.

Table 14. *Means and Standard Deviation of cognitive skill acquisition (max. 25 points)*

	Alternated-before sequence			Alternated-after sequence		
	Problem solving (*n* = 22)	Comparing examples (*n* = 27)	Problem solving + comparing with example (*n* = 25)	Problem solving (*n* = 21)	Comparing examples (*n* = 24)	Problem solving + comparing with example (*n* = 24)
Cognitive skill acquisition	22.50 (3.58)	20.93 (4.81)	21.84 (3.28)	21.24 (4.73)	19.46 (4.76)	20.21 (3.89)

5.3.3 RQ2: Effects of sequencing, type of learning task and the interaction thereof on extraneous cognitive load

We conducted a 2×3 factorial ANOVA with adapted interaction model with prior knowledge, sequencing and type of learning task as between-subject factors and extraneous cognitive load as the dependent variable.

The inferential analysis only showed a significant effect of sequencing with a small to medium effect size on extraneous cognitive load, *F*(1, 126) = 5.32, *p* < .001, *partial* $\eta^2 = .04$; Students in the conditions with the alternated-after sequence perceived higher extraneous cognitive load than students in the conditions with the alternated-before sequence (see Table 15). There are neither significant differences in the type of learning task, *F*(2, 126) = .29, *p* = .748, *partial* $\eta^2 = .01$, nor a significant interaction of sequencing and type of learning task, *F*(2, 126) = .45, *p* = .638, *partial* $\eta^2 = .01$. Prior knowledge had no significant influence on extraneous cognitive load, *F*(1, 126) = 2.51, *p* =.116, *partial* η^2 =. 02. All interactions with prior knowledge were also not significant (*F*´s < 1).

Table 15. *Means and Standard Deviation of extraneous cognitive load (max. 5 points)*

	Alternated-before sequence			Alternated-after sequence		
	Problem solving ($n = 22$)	Comparing examples ($n = 25$)	Problem solving + comparing with example ($n = 25$)	Problem solving ($n = 21$)	Comparing examples ($n = 23$)	Problem solving + comparing with example ($n = 22$)
Extraneous cognitive load	2.58 (0.49)	2.55 (0.60)	2.47 (0.56)	3.18 (0.63)	2.81 (0.58)	3.38 (0.73)

5.3.4 RQ3: Effects of sequencing, type of learning task and the interaction thereof on application quality

We conducted a 2×3 factorial ANOVA with adapted interaction model with prior knowledge, sequencing and type of learning task as between-subject factors and application-quality as the dependent variable.

The inferential analysis only showed a significant effect of sequencing with a high effect size on application-quality, $F(1, 78) = 116.37$, $p < .001$, *partial* $\eta^2 = .24$. Students in the conditions with the alternated-before sequence had higher scores in application-quality than students in the conditions with the alternated-after sequence (see Table 16). There were neither significant differences in the type of learning task-conditions, $F(1, 78) = .35$, $p = .559$, *partial* $\eta^2 < .01$, nor a significant interaction of sequencing and type of learning task, $F(1, 78) = .17$, p=.678, *partial* $\eta^2 < .01$. Prior knowledge had no significant influence on application-quality, $F(1, 78) = 2.34$, p=.131, *partial* η^2 =. 03. All interactions with prior knowledge were also not significant (*F*´s < 1).

Table 16. *Means and Standard Deviation of application-quality (max. 82 points)*

	Alternated-before sequence			Alternated-after sequence		
	Problem solving (*n* = 21)	Comparing examples	Problem solving + comparing with example (*n* = 24)	Problem solving (*n* = 22)	Comparing examples	Problem solving + comparing with example (*n* = 19)
Application-quality	66.29 (11.62)	---	59.50 (10.68)	34.68 (14.04)	---	29.00 (8.96)

5.3.5 RQ4: Mediation effects of extraneous cognitive load and application-quality

I used PROCESS model 5 (Hayes, 2012) for the mediation analyses with cognitive skill acquisition as the criterion variable, extraneous cognitive load and application-quality as the mediators, sequencing as the predictor, and prior knowledge as a moderator.

The mediation analysis showed a significant indirect effect of application quality for the effect of sequencing on cognitive skill acquisition, $\beta = 0.47$, CI_{95} [0.17, 0.79]. There was no significant indirect effect of mental effort for the effect of sequencing on cognitive skill acquisition, $\beta = 0.06$, CI_{95} [-0.03, 0.17]. Table 17 shows the beta- and p-values for the four steps in the mediation analysis.

Table 17. *Mediation analysis with cognitive skill acquisition as criterion variable*

Criterion	**Predictor**	***β***	***p***
Step I			
Cognitive skill acquisition	Sequencing	0.17	.047
	Prior knowledge	0.39	< .001
	Prior knowledge × sequencing	- 0.23	.015
Step II			
Extraneous cognitive load	Sequencing	- 0.54	< .001
Step III			
Application-quality	Sequencing	0.77	< .001
Step IV			
Cognitive skill acquisition	Extraneous cognitive load	- 0.09	.386
	Application-quality	0.49	.001
	Sequencing	- 0.29	.041
	Prior knowledge	0.37	< .001
	Prior knowledge × sequencing	- 0.20	.029

Note. *β* = standardized regression-coefficient, *N* = 83 for steps II, III and IV; *N* = 92 for step I.

5.4 Discussion

5.4.1 Summary of findings

Preliminary analyses revealed different regression slopes for the correlation of prior knowledge and cognitive skill acquisition in the two sequencing conditions. Thus, I had to consider prior knowledge as an interaction term in all further analyses. With respect to RQ 1, the present study showed a medium sized effect of the alternated-before sequence compared to the alternated-after sequence on cognitive skill acquisition. However, due to an interaction with prior knowledge, this effect was limited to learners with low prior knowledge (hypotheses 1.1 was (partly) supported). There was no significant interaction of sequencing and the type of learning task and no significant differences in cognitive skill acquisition with respect to the learning tasks (hypothesis 1.2 and 1.3 were not supported). Furthermore, a small to medium sized effect of sequencing modelling examples alternated-before the learning tasks compared to the alternated-after sequence on extraneous cognitive load was revealed (RQ 2, hypothesis 2 was supported). Analyses for RQ 3 showed a large effect of the alternated-before sequence on application-quality compared to the alternated-after sequence (hypothesis 3 was supported). The mediation analysis (RQ 4) showed a significant indirect effect of application-quality for the effect of sequencing on cognitive skill acquisition (hypothesis 4.2 was supported), but extraneous cognitive load did not significantly mediate the effect of sequencing (hypothesis 4.1 was not supported).

5.4.2 Theoretical implications

In the light of research on sequencing of modeling examples and learning tasks, the results of the present study are again in line with the studies of Van Gog and colleagues (2011) and Leppink and colleagues (2014), which provided evidence in favor of alternated-before sequences of worked example and learning tasks (i.e. problem solving) compared to alternated-after sequences with respect to cognitive skill acquisition. However, the present study revealed that this effect is limited to learners with low prior knowledge, which provides support for the findings of Reisslein and colleagues (2006) who also found an interaction of sequencing and prior knowledge on performance. This finding of my present study is also in line with the expertise-reversal effect (e.g., Kalyuga et al. 2003; Kalyuga & Renkl, 2010), which asserts that worked examples are helpful for novices' skill acquisition, but detrimental for experts' skill acquisition. However, the question remains unresolved in the present study when the sequencing effect might reverse, in other words, how much prior knowledge is needed to benefit more from an alternated-after sequence as opposed to an alternated-before sequence. Future research could further investigate this question by manipulating prior knowledge before sequencing interventions. From a theoretical perspective on cognitive skill acquisition (e.g., Anderson, 1993; Van Lehn, 1996) the interaction-effect of the present study could be explained by the different phases of skill acquisition. Leaners with lower prior knowledge should be in an earlier phase of skill acquisition, where an analogue example before problem solving is important for acquiring a (first) schema through mapping and application of the example features (Van Lehn, 1996). Leaners with higher prior knowledge should be in a later phase of skill acquisition, e.g. when declarative rules are abstracted. In this phase, learners have already constructed a schema. Providing modelling examples before or after problem solving could have helped both students with lower and higher prior knowledge to further abstract (Anderson, 1993) and generalize (Van Lehn, 1996) the acquired schema.

With respect to the effect of sequencing on application-quality, Van Gog (2011) also showed that an alternated-before sequence leads to higher application-quality as opposed to an alternated-after sequence, but in Van Gog's (2011) study application-quality was not analyzed as a potential mediator. My present study closes this gap by analyzing a mediation model and finding a mediation effect of application-quality. Thus, the findings of the present study provide support for Van Lehn's (1996) model, because sequencing modeling examples alternated-before the learning tasks led to higher application-quality in the sense of better retrieval, mapping and application during the problem solving process, which partially mediated the effect of sequencing on cognitive skill acquisition. However, the mediation analysis also showed a negative effect of the alternated-before sequence on cognitive skill acquisition when application-quality is controlled. This means that sequencing modeling examples after problem solving also had a positive effect when controlling application quality. The question is what mechanisms that are not covered in the present study might have contributed to this effect? One possible explanation could be a better connection and integration of the features of the modeling example to one's own prior knowledge, when the learners first attempt to solve a problem before receiving the modeling example (e.g., Loibl, Roll, & Rummel, 2017). A hint for this is provided by the research of Kapur (2008, 2012) on productive failure who showed that learners with productive failure (i.e., problem solving before instruction) significantly outperformed students who received direct instruction with respect to conceptual understanding and transfer before problem solving. Loibl and Rummel (2014) assume that a better activation of prior knowledge acts as one mechanism for the superiority of productive failure. This activation of prior knowledge should lead to different attentional processes. For example, when the learners focus their attention on those critical aspects in the next modeling example they were not sure about during problem solving (Van Gog et al., 2011). In other words, an alternated-after sequence of modeling examples and learning tasks might lead to (a) better elaboration of the example and (b) better interrelation of example and problem features due to the activated prior knowledge; both of which could support construction of a more abstract schema (e.g., Renkl, 2014). Thus, future research should investigate the activation of prior knowledge during the first problem solving attempts and how this affects further attentional processes and the interrelation of concrete and abstract knowledge.

In the present study, sequencing affected extraneous cognitive load, meaning that the alternated-before sequence led to lower extraneous cognitive load than the alternated-after sequence. This is in line with the findings of Van Gog and colleagues (2011), which showed that an example-problem sequence led to lower mental effort as measure for extraneous cognitive load. Cognitive load theory (e.g., Sweller, 2010) assumes that extraneous cognitive load acts as a negative mediator in example-based learning. However, Van Gog et al. (2011) did not analyze it as a potential mediator, which would be important if one aimed to examine extraneous cognitive load as potential underlying mechanisms for sequencing effects. Our findings do not support the view of extraneous cognitive load as a mediator for the effect of sequencing on cognitive skill acquisition. This is in line with findings of Heitzmann (2014), who investigated (extraneous) cognitive load as a mediator for learning with erroneous worked examples, but they also could not show a mediation effect. Likewise, the study of Leppink et al. (2014) did not show a significant correlation of extraneous cognitive load and test performance. It seems that application-quality as a content wise measure is a better and stronger predictor for cognitive skill acquisition than extraneous cognitive load.

One advantage of observational learning is that it allows for monitored enactment (Carroll & Bandura, 1990), which is comparison, detection and correction of discrepancies between conceptions (i.e. models) and behavior. Monitored enactment is regarded as important mechanism in observational learning (Bandura, 1991). Fostering these processes of monitored enactment by a combination of problem solving and comparing with the example as a learning task was assumed to foster cognitive skill acquisition. However, this could not be shown in the present study. One explanation could be that the limited time the learners had to solve the problem and to compare with the example was insufficient to profit adequately from this type learning task. The study of Singley and Anderson (1989) showed, that the number of new to be learned rules predicted the training time. That might be seen as support for this explanation: it might be that the limited time in this study was mostly needed for the problem solving part, leaving insufficient time for the comparing part of the learning task.

Furthermore, the present study assumed that the type of learning task could interact with the sequencing of modeling examples, because case comparison as a learning task may be especially helpful to provide preparation for future learning (Schwartz & Bransford, 1998; Schwartz et al., 2011) in order to profit from the following instruction such as modeling examples when the discovered commonalities and features can be interrelated with the instruction, which supports schema construction (Alfieri et al., 2013). However, the present study did not provide support for an interaction of sequencing and type of learning task. One explanation might be that problem solving activities between example comparisons are needed, as this was the case in the study by Schwartz and colleagues (2011).

5.4.3 Limitations

One limitation of the present study might be that only one modeling example was provided per problem type as opposed to at least two examples per problem type (see Chapter 2.2.1), even though the modeling examples for the three problem types share many principles. A reason for providing two or more examples for the same problem category is that it can support abstracting the structural features of the examples instead of surface-features (Renkl, 2014). However, it is not clear, if the findings of the present study are valid for treatments with two or more modeling examples per problem type, because in the latter case different sequencing strategies can be implemented. Future research could address this issue by manipulation the number of modeling examples as well as the sequence of modeling examples and learning tasks, for example when providing one vs. two examples in an alternated (or blocked)-before sequence vs. an alternated (or blocked)-after sequence.

Another limitation concerns the measurement of application-quality, which could only be assessed in the four conditions with problem-solving and not in the two conditions with comparing examples as learning tasks. This could have affected the results of the mediation analysis, e.g. with respect to the lower predictive value of extraneous cognitive load, because it reduced the statistical power. However, the findings of the study reported in Chapter 4 can be considered as support for the validity of the mediation findings in the present study, because they also showed the same pattern, namely that application-quality mediated the effect of sequencing, whereas extraneous cognitive load did not.

Finally, the present study did not include a delayed posttest and therefore we do not know if the short-term effects also remain on the longer term.

6 General discussion

This dissertation extended prior research on different sequencing strategies in example-based learning by developing and investigating a model on differential sequencing effects (see

Figure *16*). It was assumed that an alternated sequence of modeling examples before learning tasks is more effective in terms of cognitive skill acquisition than a blocked sequence of modeling examples before learning tasks (e.g., Trafton & Reiser, 1993). It was also hypothesized that sequencing modelling examples (alternated) *before* learning tasks should lead to higher cognitive skill acquisition than sequencing modelling examples (alternated) *after* learning tasks (e.g. Van Gog et al., 2011; Leppink et al., 2014, Hsu et al., 2015; Kant et al., 2017). However, both sequencing strategies were assumed to have potentially positive effects on cognitive skill acquisition, but for each strategy different learning mechanisms were assumed. A (potential) positive effect of sequencing modeling examples before learning tasks was supposed to be mediated by extraneous cognitive load (e.g., Van Gog et al., 2011) and application-quality (e.g., Van Lehn, 1996). A (potential) positive effect of sequencing modeling examples after learning tasks was assumed to be moderated by prompting self-explanations and monitoring (Loibl et al., 2017) or by the type of learning task (Alfieri et al., 2013).

6.1 Summary

In this Chapter, the results of both empirical studies of this dissertation are summarized with respect to the general research questions raised in Chapter 3.

General research question I: To what extent do different sequencing strategies for modeling examples and learning tasks have an effect on cognitive skill acquisition?

GRQ Ia: To what extent does an alternated vs. blocked sequence of modeling examples and learning tasks have an effect on cognitive skill acquisition?

It was assumed that an alternated-before sequence of modelling examples and learning tasks would lead to higher cognitive skill acquisition compared to a blocked-before sequence. In line with Trafton and Reiser (1993), the findings of Study I provide support for this. In particular, the alternated-before sequence, i.e. alternating modeling examples and learning tasks led to significantly higher cognitive skill acquisition than the blocked-before sequence, i.e. providing modeling examples blocked before problem solving.

GRQ Ib: To what extent does sequencing of modeling examples before vs. after learning tasks have an effect on cognitive skill acquisition?

In line with studies on sequencing from a cognitive skill acquisition perspective (e.g., Van Gog et al., 2011; Leppink et al., 2014) it was hypothesised that sequencing modelling examples (alternated) before learning tasks would lead to better cognitive skill acquisition compared to an (alternated) sequence of modelling examples after learning tasks. Study I partly supports this: the alternated-before sequence led to significantly higher cognitive skill acquisition than the blocked-before sequence, but the blocked-before sequence did not differ significantly from the blocked-after sequence. Study II enhanced the findings of Study I, because it showed, that the alternated-before sequence also led to significant higher cognitive skill acquisition than the alternated-after sequence.

General research question II: How can the different effects of sequencing of modeling examples and learning tasks on cognitive skill acquisition be explained?

GRQ II a: To what extent is a potential positive effect of sequencing of modeling examples before learning tasks on cognitive skill acquisition mediated by extraneous cognitive load and application-quality during learning?

Extraneous cognitive load was hypothesized to play a mediating role for a potential positive effect of sequencing modelling examples (alternated) before learning tasks (e.g., Sweller & Cooper, 1985, Van Gog et al., 2011). However, neither of the two studies provides support for this, because extraneous cognitive load did not significantly mediate the positive effect of the alternated-before sequence on cognitive skill acquisition.

Application-quality was assumed to play a mediating role for a positive effect of sequencing modeling examples (alternated) before learning tasks (e.g., Van Lehn, 1996). In both studies, the mediation analysis showed a significant indirect effect of application-quality for the positive effect of the alternated-before sequence on cognitive skill acquisition.

GRQ II b: To what extent does a potential positive effect of sequencing modeling examples after learning tasks on cognitive skill acquisition depend on prompting self-explanations and monitoring or on different types of learning tasks?

Prompting self-explanations and monitoring was hypothesized to interact with sequencing modeling examples before or after learning tasks. That is, learners might profit more from modelling examples after learning tasks when they are prompted to self-explain and monitor, because learners who are prompted to self-explain an example and diagnose critical problem solving aspects should be more aware of their knowledge gaps (e.g., Loibl, Roll, & Rummel, 2017) and then focus their attention in the modeling examples on those critical aspects (Bandura, 1986). However, the effect of sequencing was not dependent on prompting self-explanations and monitoring.

An interaction between sequencing and type of learning task for cognitive skill acquisition was assumed. That is, comparing examples as a learning task after modeling examples might be more effective than a modeling example before the learning task, because comparing examples as a learning task may be especially helpful to provide "preparation for future learning" (Schwartz & Bransford, 1998; Schwartz et al., 2011) in order to profit more from subsequent instruction (Alfieri et al., 2013). This kind of interaction might also be the case, when the learning task consists of problem solving and comparing with examples, because when the learners compare their solution with the modeling example's solution, they might become aware of their knowledge gaps (e.g.. Loibl, Roll, & Rummel, 2017) and then can focus their attention in the modeling examples on the critical aspects they were not sure about (Bandura, 1986). However, the effect of sequencing was not dependent on the type of learning task.

Taken together, both studies showed that the alternated-before sequence (i.e. modelling examples alternated before learning tasks) led to lower extraneous cognitive load, higher application-quality as well as higher cognitive skill acquisition, compared to the blocked-before and blocked-after sequence as well as compared to the alternated-after sequence. Furthermore, in both studies application-quality mediated the positive effect of sequencing modelling examples alternated-before the learning tasks on cognitive skill acquisition, whereas extraneous cognitive load did not. Lastly, effects of sequencing modelling examples alternated-after the learning tasks were not dependent on prompting self-explanations and monitoring or on the type of learning task. Table 18 provides an overview for the findings of both studies with respect to the general research questions as well as to the specific hypotheses of the two studies.

Table 18. *Overview of the findings of Study I and Study II*

	Study I	**Study II**
GRQ I a	Alternated-before > blocked-before (H1.1 was supported)	
GRQ I b	Alternated-before > blocked-after (H1.1 was supported) blocked-before = blocked-after (H1.3 was not supported)	Alternated-before > alternated-after (but only for learners with low prior knowledge) (H1.1 was partly supported)
GRQ II	Effect was not dependent on prompting (H1.4 was not supported)	Effect was not dependent on type of LT (H1.2 was not supported)
	Alternated-before > blocked-before = blocked-after for ECL (H2.1 & 2.2 were supported, H2.3 was not supported)	Alternated-before > alternated-after for ECL (H2 was supported)
	Alternated-before = blocked-before > blocked-after for AQ (H3.2 & 3.3 were supported, H3.1 was not supported)	Alternated-before > alternated-after for AQ (H3 was supported)
	ECL no significant mediator, AQ significant mediator (H4.1 was not supported, H4.2 was supported)	ECL no significant mediator, AQ significant mediator (H4.1 was not supported, H4.2 was supported)

Note. ECL stands for extraneous cognitive load, AQ stands for application-quality.

6.2 Theoretical implications

This dissertation investigated the effects of sequencing of modelling examples and learning tasks on cognitive skill acquisition, which learning mechanisms – such as extraneous cognitive load and application-quality – might mediate potential effects on cognitive skill acquisition, and the conditions – such as prompting self-explanations and monitoring or the type of learning task – under which specific sequences have specific effects on cognitive skill acquisition. The next sections will discuss the theoretical implications of the findings of both empirical studies of this dissertation.

6.2.1 GRQ I: Effects of sequencing on cognitive skill acquisition

In line with models of cognitive skill acquisition (e.g., Anderson, 1982; Van Lehn, 1996; Renkl, 2014), both empirical studies of this dissertation provide support for explanations of the positive effect of sequencing modeling examples before learning tasks based on schema construction indicated by application-quality: the alternated-before sequence increased cognitive skill acquisition more than blocked-before sequence (e.g., Trafton & Reiser, 1993) as well as compared to the blocked- and alternated-after sequences (e.g., Van Gog et al., 2011), at least for learners with low prior knowledge (e.g., Reisslein et al., 2006). From a cognitive skill acquisition perspective, learning in early phases of skill acquisition heavily relies on examples that can support constructing a basic schema that guides later problem solving (e.g., Renkl, 2014). When the learner has more prior knowledge (i.e., cognitive schema) that can guide problem solving, the sequence seems not that important any more. Taken together, the results of my dissertation confirm the models on cognitive skill acquisition.

6.2.2 GRQ II: Explanations for the effects of sequencing on cognitive skill acquisition

The findings of both empirical studies provide (some) evidence in favor of application-quality and extraneous cognitive load as explanatory factors. During problem solving, an example is retrieved, mapped and applied, which can further generalize the acquired schemata (e.g., Van Lehn, 1996). Both studies provide support for this, that is, application quality mediated the superiority of the alternated-before sequence. At a first glance, the results of the two present studies also provide support that extraneous cognitive load has explanatory power (e.g., Van Gog et al., 2011), because the alternated-before sequence led to lower extraneous cognitive load and better cognitive skill acquisition compare to an (alternated- and blocked-) after sequence. Van Gog and colleagues (2011) also argued in this direction and concluded that this sequencing effect can be explained by extraneous cognitive load. However, the mediation analyses in both studies revealed that extraneous cognitive load is not only less important compared to application-quality; it was also not a significant mediator. When we would not have analyzed extraneous cognitive load as mediator, we would have come to a wrong conclusion, as stated by Van Gog and colleagues (2011). The findings with regard to extraneous cognitive load of both studies are in line with criticism on cognitive load theory by De Jong (2010), who questioned the theory's validity and generalizability. In particular, De Jong raised the question of how likely it is to experience cognitive overload in realistic learning settings, when learning time is not only ten minutes as is the case in many experimental studies (e.g., Leppink et al., 2014), and when there is more learner control (such as making notes, manipulating videos and so on) which is quite limited in many experimental studies to avoid confounding effects. In addition to concerns about the applicability of cognitive load theory to authentic learning settings, results of mediation analysis in this dissertation show that application-quality as a qualitative (i.e., content-wise) indicator of the learning process is more important than extraneous cognitive load. This is underlining the argument that questions the validity aspect of cognitive load theory with respect to the role of extraneous cognitive load.

The explanations from the existing literature for a (potential) superiority of sequencing modeling examples after learning tasks (see Chapter 2) were not supported by the two empirical studies of this dissertation. Prompting self-explanations and monitoring (in Study I) and problem solving and comparing with examples (in Study II) were assumed to support a greater awareness of knowledge gaps which should help to better integrate the subsequent instruction (Loibl & Rummel, 2014; Loibl, Roll, & Rummel, 2017). However, both studies did not show that an effect of sequencing is dependent on supporting the awareness of knowledge gaps (i.e., by prompting self-explanations and monitoring in Study I or by problem solving and comparing with examples in Study II). This might be explained by the finding of Loibl and Rummel (2014) that it is also important to compare and contrast the student solution to the canonical (i.e., correct) solution, which was not done in Study I. Therefore, only prompting self-explanations and monitoring might not have been enough to support the connection and integration of the subsequent modelling examples with the identified knowledge gaps. Although the problem solving and comparing with examples as learning tasks in Study II included – to some degree – contrasting the own solution to the canonical solution, the procedure in Study II differed from Loibl and Rummel (2014) in at least two aspects: (1) the students did not compare their solution to the identical correct solution, but to the same solved example that was also provided at the end of the respective modeling example, and (2) the students had to compare their solution by themselves, and it was not compared and contrasted by a teacher, as it was the case in Loibl and Rummel (2014). A comparison by the teacher might be superior, because learners with little prior knowledge often fail to monitor their performance correctly (Dunning et al., 2003). An indicator for this could be the interaction of sequencing and prior knowledge in Study II, where the alternated-before sequence was important for cognitive skill acquisition of learners with low prior knowledge, but not for learners with high prior knowledge. Another possibility for supporting awareness of knowledge gaps and also the connection and integration of the subsequent instruction with the identified knowledge gaps might be regarded in a more self-regulated use of the examples during problem solving. Foster et al. (in press) recently showed that worked examples were studied more often after failing to solve the problem compared to successfully solving the problem. Future research could compare the role of a teacher-regulated comparison phase with a self-regulated comparison phase for the effectiveness of different sequences.

The implementation of specific design features, especially the use of comparing examples as learning tasks, was considered as another explanation for findings that show a superiority of sequences with modeling examples after learning tasks (Alfieri et al., 2013; Loibl, et al., 2017). However, Study II did not show that an effect of sequencing is dependent on the type of learning tasks (i.e., comparing examples). To some extent, this is in line with the findings of Nokes-Malach and colleagues (2013) which showed that conditions with comparing examples outperformed students with examples alternated before problem solving on far transfer when transfer is assessed with multiple choice items that require qualitative reasoning. However, this advantage disappeared for intermediate transfer that is the application of procedures in a new combination. For near transfer (i.e., application of the same procedure as during learning), examples alternated before problem solving were even better than comparing examples (Nokes-Malach et al., 2013). However, Study II did not address far transfer. The skill acquisition posttest used in Study II only required the application of the same procedures as during learning, once with a different combination, which is similar to near and intermediate transfer. The measurement of near transfer in Nokes-Malach and colleagues (2013) study was interleaved with three problem solving tasks throughout the learning phase. In Study II, learners in the comparing examples conditions had only one opportunity for problem solving in the learning phase; this was after the first set of modeling example and learning task. From a perspective on cognitive skill acquisition (e.g., Van Lehn, 1996; Anderson, 1983) the different numbers of practice opportunities could have influenced the degree of cognitive skill acquisition, because practice is important for the transition from a declarative phase to a compilation phase.

The present studies further differed from studies that showed a benefit of sequences with modeling examples after learning tasks (e.g., Kapur, 2012; Schwartz et al., 2011) in at least three aspects: (1) the focus in the learning tasks in Kapur’s (2012) study or in Schwartz and colleagues’ (2011) study was on invention activities, whereas this was not the case in the present studies. Learners had already had a lecture on the topic one month before the data collection of the studies and thus could not invent the to be learned principles anymore. It might be possible that invention activities are important as preparation for learning from subsequent instruction. A study by Roelle and Berthold (2015) which showed that learners with comparing examples as invention activity learned more from subsequent instructional explanations than students with no preparation activity (i.e., without comparing examples) could be seen as evidence for this assumption. (2) A second difference regards the social form of working on the learning tasks that is in a cooperative setting in Kapur's and Schwartz and colleagues' studies compared to individual learning in my studies. Collaboration can have the advantage of fostering elaboration processes, for example by explaining to others or by inducing cognitive conflicts (Nastasi & Clements, 1992; King, 2007). This in turn could have affected the differentiation of the learners' prior knowledge, when they experience cognitive conflicts or that they are not able to explain to others. However, a study of Mazziotti and colleagues (2015) showed that modeling examples after problem solving is similarly effective for both individual and collaborative learning settings. Therefore, a potential positive effect of after-sequences should not depend on the social form. (3) A third difference lies in the length of learning time. Studies on productive failure (e.g., Kapur, 2012; Mazziotti et al., 2015) and preparation for future learning (e.g., Schwartz & Martin, 2004; Schwartz et al., 2011) have learning times over several hours. In contrast, our studies had a learning time of around 40 minutes, and studies of Leppink and colleagues (2014) or Van Gog (2011) only had ten minutes of learning time. It might be the case, that providing modeling examples after learning tasks is more effective, when the learners have enough time trying to solve the learning tasks. Future research could vary the learning time the students have to solve the learning tasks before they receive the modeling example(s) to reveal, whether a benefit of after-sequences only occurs for longer learning times.

6.2.3 A model for the effectiveness of different sequencing strategies on cognitive skill acquisition

Taken together, this dissertation's model of different sequencing strategies in example-based learning provides more support for an explanation based on application-quality according to cognitive skill acquisition theories (e.g., Anderson, 1982; Van Lehn, 1996, Renkl, 2014) than for an explanation based on extraneous cognitive load from cognitive load theory (e.g., Sweller, 2010) for the positive effect of the alternated-before sequence. Explanations based on prompting the awareness of knowledge gaps for productive failure (e.g., Kapur, 2012; Loibl, Roll, & Rummel, 2017) or integrating comparison tasks as preparation for future learning (e.g., Schwartz & Bransford, 1998; Alfieri et al., 2013) for a (potential) effectiveness of sequencing modeling examples after learning tasks were not supported in this dissertation.

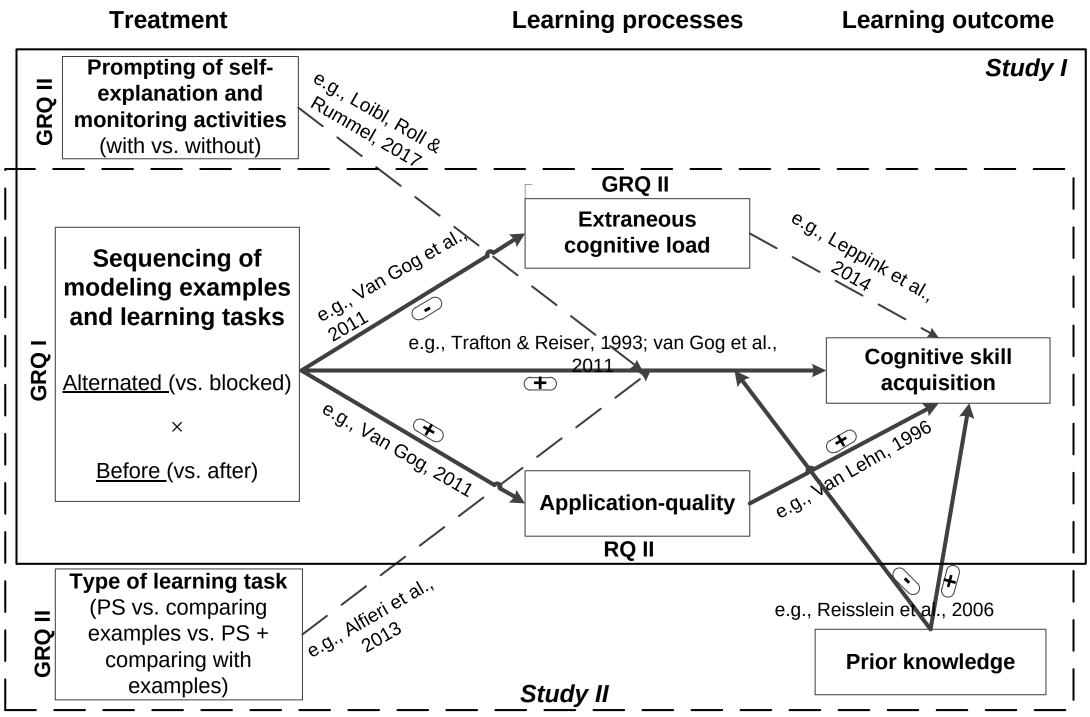


Figure 16. Model for the effectiveness of different sequencing strategies in EBL

Note. The bold-lines illustrate the effects and +/- the direction of the effect, the dashed-lines illustrate that there was no significant effect.

However, the two empirical studies also have some limitations with respect the validity of the investigated model. These will be discussed in the next section.

6.3 General limitations and implications for future research

Limitations of the present studies can be derived from at least four perspectives: (1) internal and external validity of the intervention, (2) internal and external validity of the process measure, (3) internal and external validity of the outcome measure, as well as (4) the validity of the constructed model.

6.3.1 Internal and external validity of the intervention

Regarding the internal validity of the intervention, it is a limitation that I had no traditional control group such as problem solving only or examples only (e.g., Van Gog et al., 2011). Thus, I cannot conclude whether students are better in terms of cognitive skill acquisition when they learned with different sequences of modelling examples and learning tasks compared to modelling examples or learning tasks only. However, the specific interest of my dissertation aimed at different sequences of modelling examples and learning tasks, and not whether modelling examples or problem solving only are more effective for cognitive skill acquisition compared to different sequences of modelling examples and learning tasks. Furthermore, there is yet a large body of research that compared a combination of worked examples and problem solving with problem solving only (e.g., Cooper & Sweller, 1987; Crippen & Earl, 2007; Hilbert & Renkl, 2009; Paas & Van Merriënboer, 1994). Thus, my dissertation enhanced research on combining worked examples with problem solving by investigating different sequencing strategies.

A limitation with respect to the external validity of the interventions is the sample. In both of my studies, the sample consisted of educational science students of which the majority were female. De Jong (2010) also criticized this aspect for research on cognitive load theory. However, De Jong questions not so much whether the findings are transferable to older or younger, male or female learners, but rather whether conclusions drawn from studies with student samples are valid when learning materials (such as the frog leave-problem in Van Gog, 2011) that are irrelevant for such a sample are being used. Although the conclusions of this dissertation are limited because the sample consisted of mainly female university students, external validity is strengthened with using authentic material. The learning material used in this dissertation is considered a difficult “topic” by undergraduate educational sciences students and is also important for their exam. Thus, the material used can be considered relevant for the sample that was utilized because it was taken from the actual curricula, provided additional practice opportunity and exam preparation. One the one hand, the fact that there was an exam can also be regarded as limitation, because the students might have been extrinsically motivated by the exam and acted differently as they would have had when been intrinsically motivated. On the other hand, as the learning content was not solely important for the exam, but also elementary for future courses as well as for future work as professional, more intrinsically motivated learner attitudes are – even though unlikely – also possible.

Another aspect that might limit the external validity is the decision, that the studies of my dissertation provided video-based modelling examples, because there was no interaction with a teacher who could be asked questions as is the case in settings where the teacher is physically present and models the procedure. However, the provision of video-based modelling examples instead of ‘live’ modelling examples has the advantage of standardization between the conditions, which is important for internal validity. Furthermore, research has shown that video-based modelling examples are also effective (Van Gog & Rummel, 2010).

The problem-type (i.e., well- or ill-structured) can also be regarded as a limitation with respect to the external validity, because whether problems are well- or ill-structured (e.g., Jonassen, 1997; Schraw, Dunkle, & Bendixen, 1995) might influence the processes of example retrieval, mapping and application to the learning task (e.g., Van Lehn, 1996). The problems I used in this dissertation can be classified as well- structured problems that have a clear solution in contrast to ill-structured problems with no clear, but multiple possible solutions (e.g., Jonassen, 1997; Schraw, Dunkle, & Bendixen, 1995). Within well-structured problems, retrieval, mapping and application of example features to problem solving might be easier, because there is only one typical solution strategy. Whether the findings of my dissertation can be transferred to ill-structured problems is not clear yet. Future research could address this by a between-within-subjects design, where the sequence is manipulated between groups of learners as an alternated-before vs. an alternated-after sequence and the structure of the problem situation is manipulated within the groups. Students first learn with an alternated-before or after-sequence that is first based on well-structured problems and then based on ill-structured problems.

6.3.2 Internal and external validity of the process measures

A limitation with respect to the internal validity of the process measure is that the processes, which the moderators (i.e. prompting and type of learning task) aimed to encourage (i.e., monitored enactment and mapping processes) were not investigated. Due to the manipulation, only a portion of the subjects provided data on those processes. Evidence from Study II on the negative effect of the alternated-before sequence of modelling examples and learning tasks when controlling for application-quality in the mediation-analysis supports the view that there might be other mechanisms, such as monitored enactment (i.e., comparison, detection and correction of discrepancies between the model's and the learners' behavior), that were not captured. Future research could design experiments that aim at investigating the assumed mechanisms for a superiority of modelling examples after learning tasks, followed by the conduction of experiments that aim to support these processes (e.g., Bannert, 2009).

Regarding the content validity of the measure for extraneous cognitive load, it might be questioned whether subjective rating scales are an appropriate way to assess extraneous cognitive load due to the subjective nature of the rating scales (e.g., De Jong, 2010). For example, it is not clear whether students interpret the answer options of the rating-items in the same manner and whether that might explain the mixed results for the relation of extraneous cognitive load with cognitive skill acquisition (e.g., Leppink et al., 2013; Leppink et al., 2014). However, my dissertation showed consistently over both studies that extraneous cognitive load – assessed with the mental effort rating scale by Paas (1992) and with a subscale by Opfermann (2008) – was significantly related with cognitive skill acquisition, but only when application quality was not included. These findings also imply a methodological suggestion for future research: When the aim is to investigate extraneous cognitive load as an explanation (i.e., mediator) for the effectiveness of an intervention (such as different sequencing strategies), a mediation analysis is recommended instead of calculating mental efficiency, which is done in many studies on investigating extraneous cognitive load as explanation for an effect (see Chapter 2.2.1). When mental efficiency is calculated, the expected mechanism is combined with the learning outcome and thus not distinguishable any longer (e.g., De Jong, 2010). From a mediation-perspective, it would make more sense to combine extraneous cognitive load (indicated with mental effort) and application-quality (or training performance) and to investigate this as mediator. However, the conceptual problem of combining two measures that are based on different theoretical concepts (i.e., cognitive load theory and cognitive skill acquisition theories) in one score remains. This is, that the two measures are not distinguishable any longer which questions the content validity (De Jong, 2010). Furthermore, when two assumed mediators are combined in one score (i.e., extraneous cognitive load and application-quality) it is no longer possible to compare the predictive value of the two mediators, which is another argument for a mediation analysis with distinct variables.

6.3.3 Internal and external validity of the outcome measures

The content validity of the outcome measure was limited with respect to the spectrum that was assessed by the post-test on cognitive skill acquisition. For example, the post-test did not address all sub-productions (Anderson, 1992) of the overall skill to describe designs according to the Campbell and Stanley scheme.
This limits the content validity of the measurement because not the full range of the to be learned cognitive skills was assessed. Also, far transfer was not assessed, which somehow limits the generalization of the findings with respect to the type of learning tasks of ‘comparing examples’, because it seems that comparing examples especially influences far transfer (e.g., Nokes-Malach et al., 2012). Thus, if my

second study would have assessed far transfer, it might have had shown different effects for the type of learning task of ‘comparing examples’.

6.3.4 Validity of the model

A limitation with respect to the validity of the model is that we did not vary all sequencing conditions (i.e., alternated – blocked × before – after) in both studies. Thus, this dissertation cannot provide conclusions with respect to a comparison of an alternated-after sequence with a blocked-after sequence. Future research could close this gap by examining the different sequences in a 2×2 design. This could reveal, whether a sequencing effect (before vs. after) is dependent on the alternating sequencing and whether an alternated-after sequence is better as a blocked-after sequence for cognitive skill acquisition as it might be expected based on cognitive skill acquisition models (see Chapter 2.1).

6.4 Practical implications

What can we conclude from the findings of this dissertation for the practical problem raised in Chapter I? What sequence can we recommend to Michael for his mathematics class?

Based on the results of my empirical studies I can recommend that modeling examples should be provided alternated-before learning tasks instead of a blocked-before sequence of modeling examples and learning tasks as well as instead of an alternated- and blocked-after sequence, at least for learners with low prior knowledge. This also suggests, that when teachers provide lessons for a class where they do not know exactly each leaner's prior knowledge, an alternated-before sequence is recommended. With respect to the type of learning task no concrete conclusion can be drawn, because there were no statistically significant differences between the different learning tasks. However, the findings of both studies show that application-quality mediates the positive effect of alternating modeling examples before problem solving. Therefore it seems worthwhile to provide modeling examples alternated-before problem solving and also to support application-quality during example-based learning.

Taken together, when Michael wants to provide three modeling examples, each for one different probability calculation and to combine the three examples with three problem solving tasks (i.e., each for one calculation type), he should sequence the three modeling examples alternated-before the problem solving tasks in order to support the quality of the application of the calculation principles. Thus, he should implement the following sequence: Modeling example on calculation type 1, problem solving task on calculation type 1, modeling example on calculation type 2, problem solving task on calculation type 2, and modeling example on calculation type 3, problem solving task on calculation type 3. With this sequence, the learners can best apply the calculation principles during problem solving and thus fosters their cognitive skill acquisition.

7 Outlook

Taken together, my dissertation advanced empirical research and theory on different sequencing strategies in example-based learning by investigating a model on differential sequencing effects that examined mediating processes as well as moderating conditions. In two empirical studies I have shown, that an alternated-before sequence of providing modelling examples before learning tasks leads to lower extraneous cognitive load, higher application-quality as well as higher cognitive skill acquisition compared to blocked-before and blocked-after sequences as well as compared to an alternated-after sequence, at least for learners with low prior knowledge. Furthermore, in both studies application-quality mediated the effect of sequencing on cognitive skill acquisition, whereas extraneous cognitive load did not. Lastly, the effects of sequencing were neither dependent on prompting self-explanations and monitoring nor on the type of learning task.

In conclusion, for explaining the effectiveness of the alternated-before sequence the results of my two empirical studies provide more support for an explanation based on cognitive skill acquisition theories (e.g., Anderson, 1982; Van Lehn, 1996, Renkl, 2014) than for an explanation based on cognitive load theory (e.g., Sweller, 2010). Explanations based on supporting the awareness of knowledge gaps for productive failure (e.g., Loibl, Roll, & Rummel, 2017) and on comparing tasks as preparation for future learning (e.g., Schwartz et al., 2011 were not supported in this dissertation. However, both studies most likely were not able to capture the relevant processes for the (potential) effectiveness of sequences with modeling examples after learning tasks, which is a limitation of this dissertation. An indicator for this is the finding in Study II that the alternated-after sequence had a positive effect on cognitive skill acquisition after controlling for application-quality. Future research could focus on the processes of (prior) knowledge activation and the awareness of knowledge gaps as well as integration of the new knowledge (e.g., Loibl et al., 2017) when comparing the effects of (alternated- or blocked-) before and after sequences.

The activation of prior knowledge could be investigated by measuring declarative prior knowledge before and after the first modeling example or learning task (depending on the condition), e.g. with ABC-lists (Birkenbihl, 2011) which is an open association task. Thereby, the learners would write down their associations they have to the letters of the alphabet with respect to a specific topic (i.e., describing designs according to the Campbell & Stanley-scheme). When providing learning tasks before modeling examples will actually lead to a better activation of prior knowledge compared to modelling examples before learning tasks, there should be an improvement in the numbers of associations made in the ABC-list dependent on the sequencing condition. The awareness of knowledge gaps could be measured by a scale developed by Glogger-Frey and colleagues (2013) on global awareness of knowledge gaps. This scale was also used by Loibl and Rummel (2014) and would allow for a direct comparison with their study on sequencing effects. The knowledge integration, which can be compared to monitored enactment (e.g., Carroll & Bandura, 1990) might be addressed by analyzing, how the application-quality developed over the different problem solving attempts. That means, for example, to examine the errors that are made in each task and whether they were corrected or not.

A further question that can be raised for future research is whether it is an either-or decision for either one sequence or whether it is a matter of (appropriate) combination. From a cognitive skill acquisition-perspective (e.g., Anderson, 1992; Van Lehn, 1996; Renkl, 2014), investigating effects of fading might also be fruitful, because fading seeks to incorporate the different stages of cognitive skill acquisition (e.g., Renkl & Atkinson, 2003). From this perspective it would be interesting for future research to compare the following sequencing-combinations: modeling example - faded modeling example, modeling example - problem solving and problem solving – modeling example vs. alternated modeling examples before problem solving vs. alternated problem solving before modeling example pairs vs. fading only.

In this dissertation, the examples were implemented as video-based modeling examples that were provided to all students. From private communication is known, that this was very appreciated by the students, who used the modeling examples and the learning tasks to learn and prepare for the exam, in which proposing a design for a specific research scenario is commonly asked. From this more self-regulated learning perspective (e.g. Zimmermann, 2002) it would be interesting to compare a self-regulated use of modeling examples and learning tasks with a prescribed sequence of modeling examples and learning tasks. Such an investigation could outline an advantage of a self-regulated use of modeling examples during problem solving, because the modeling examples could be studied in the right time the learners need them. Furthermore, a self-regulated use of modeling example might better allow for monitored enactment (e.g., Carroll & Bandura, 1990) as well as for regulation activities (e.g., Zimmermann, 2002). However, there might also be the risk that students will use the modeling examples in a max-analogy way (Muldner & Conati, 2010), which means that they copy as much as they can instead of first trying to solve the problem on their own and refer to the example only when they experience impasses or struggle (i.e., min-analogy). Therefore it might be necessary to support a min-analogy strategy by training or by prompts (e.g., Renkl, 2014).

Lastly, it can be asked, how educational professionals such as teachers can be supported to use the evidence of this dissertation – and from educational psychology in general – for their daily practice. By example-based learning, my dissertation concludes.

8 References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons. A meta-analytic review *Educational Psychologist*, *48*(2), 87–113. doi: 10.1080/00461520.2013.775712

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369–404.

Anderson, J. R. (1993). Problem solving and learning. *American Psychologist*, *48*(1), 35–44. doi:10.1037/0003-066X.48.1.35

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences*, *4*(2), 167–207. doi: 10.1207/s15327809jls0402_2

Anderson, J. R. & Fincham, J. M. (1994). Acquisition of procedural skills from examples. *Journal of experimental psychology: learning, memory, and cognition*, *20*(6), 1322–1340. doi: 10.1037/0278-7393.20.6.1322

Anderson, J. R., Greeno, J. G., Kline, P. J., & Neves, D. M. (1981). Acquisition of problem-solving skill. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum, 191-230.

Anderson, J. R., Fincham, J. M., & Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of experimental psychology: learning, memory, and cognition*, *23*(4), 932–945. doi: 10.1037/0278-7393.23.4.932

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, *95*(4), 774–783. doi: 10.1037/0022-0663.95.4.774

Baars, M., van Gog, T., Bruin, A. de, & Paas, F. (2017). Effects of problem solving after worked example study on secondary school children’s monitoring accuracy. *Educational Psychology*, *37*(7), 810–834. doi: 10.1080/01443410.2016.1150419

Baggett, P. (1987). Learning a procedure from multimedia instructions: The effects of film and practice. *Applied Cognitive Psychology*, *1*(3), 183-195.

Bandura, A. (1971). *Psychological modeling: Conflicting theories*. Chicago, IL: Aldine Atherton.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs: Prentice Hall.

Bandura, A. (1991). Social cognitive theory of moral thought and action. In W.M. Kurtines, J.L. Gewirtz (eds.), *Handbook of moral behavior and development*. Hillsdale, NJ: Erlbaum, 45–103.

Bannert, M. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie*, *23*(2), 139–145. doi: 10.1024/1010-0652.23.2.139

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, *17*(5), 564–577. doi: 10.1016/j.learninstruc.2007.09.007

Berthold, K., Röder, H., Knörzer, D., Kessler, W., & Renkl, A. (2011). The double-edged effects of explanation prompts. *Computers in Human Behavior*, *27*(1), 69–75. doi: 10.1016/j.chb.2010.05.025

Birkenbihl, V. F. (2011). *Trotzdem lehren* (5. Aufl.). München: mvg-Verl.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII. Cognitive* *regulation of performance: Interaction of theory and application*. Cambridge: MIT Press, 435-459.

Brünken, R., Moreno, R., & Plass, J. L. (2010). *Cognitive load theory.* New York: Cambridge University Press.

Campbell, D. T. & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally & Company.

Carroll, W. R. & Bandura, A. (1987). Translating cognition into action: The role of visual guidance in observational learning. *Journal of Motor Behavior, 19*(3), 385-398.

Catrambone, R. & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of experimental psychology: Learning, memory, and cognition, 15*(6), 1147–1156. doi: 10.1037/0278-7393.15.6.1147

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182. doi: 10.1016/0364-0213(89)90002-5

Chi, M. T. H., De Leeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439–477. doi: 10.1207/s15516709cog1803_3

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. L. Erlbaum Associates.

Cooper, G., & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, *79*(4), 347-362. doi: 10.1037/0022-0663.79.4.347

Conati, C., & VanLehn, K. (2000). Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education, 11*, 398–415.

Crippen, K. J., & Earl, B. L. (2007). The impact of web-based worked examples and self-explanation on performance, problem solving, and self-efficacy. *Computers and Education, 49*(3), 809–821.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science*, *38*(2), 105-134. doi: 10.1007/s11251-009-9110-0

Doyle, W. (1983). *Managing Academic Tasks: An Interim Report of the Junior High School Study*. R & D Rep. 6186.

Dunning, D., Johnson, K., Erlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12*, 83–87. doi: 10.1111/1467-8721.01235

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self-evaluations undermine students' learning and retention. *Learning and Instruction, 22*(4), 271-280. doi: 10.1016/j.learninstruc.2011.08.003

Durkin, K. (2011). The self-explanation effect when learning mathematics: A meta-analysis. Paper presented at the *Society for Research on Educational Effectiveness*, Washington DC.

Faries, J. M., & Reiser, B. J. (1988). Access and use of previous solutions in a problem solving situation. In V. L. Pate1 & G. L. Groen (Ed.), *Proceedings of the Tenth Annual Cot&rence of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates, 433-439.

Fitts, P. M. (1964). Perceptual-motor skill learning. In Melton, A. W. (Ed.), *Categories* of *human learning*. New York: Academic Press, 243–285.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191. doi: 10.3758/BF03193146

Foster, N. L., Rawson, K. A., & Dunlosky, J. (in press). Self-regulated learning of principle-based concepts: Do students prefer worked examples, faded examples, or problem solving? *Learning and instruction.* doi: 10.1016/j.learninstruc.
2017.10.002

Gentner, D. (2003). Psychology of analogical reasoning. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science*. London: Nature Publishing Group, Vol. 1, 106–112.

Germ, M. (2009). *Lernaufgaben als kohärenzbildende Elemente in der naturwissenschaftlichen Lehrerbildung*. Kiel: Dissertation.

Gick, M. L. (1986). Problem-solving strategies. *Educational Psychologist*, *21*(1-2), 99–120. doi: 10.1080/00461520.1986.9653026

Gick, M. L. & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive psychology*, *12*(3), 306-355. doi: 10.1016/0010-0285(80)90013-4

Glogger, I., Kappich, J., Schwonke, R., Holzäpfel, L., Nückles, M., & Renkl, A. (2013). Inventing prepares learning motivationally, but a worked-out solution enhances learning outcomes. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, 2416-2421.

Gräsel, C. & Mandl, H. (1993). Förderung des Erwerbs diagnostischer Strategien in fallbasierten Lernumgebungen. *Unterrichtswissenschaft, 21*(4), 355– 370.

Große, C. S. & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction, 17*(6), 612–634. doi: 10.1016/j.learninstruc.2007.09.008

Hartwig, M. K., Was, C. A., Isaacson, R. M., & Dunlosky, J. (2012). General knowledge monitoring as a predictor of in-class exam performance. *The British journal of educational psychology*, *82*(3), 456–468. doi: 10.1111/j.2044-8279.2011.02038.x

Hausmann, R. G. M. & Chi, M. T. H. (2002). Can a computer interface support self-explaining? *Cognitive Technology, 7*(1), 4–14.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling. http://www.afhayes.com/public/process2012.pdf

Hilbert, T. S., & Renkl, A. (2009). Learning how to use a computer-based concept-mapping tool: Self-explaining examples helps. *Computers in Human Behavior, 25*(2), 267–274. doi: 10.1016/j.chb.2008.12.006

Holyoak, K. J. (2005). Analogy. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning*. New York: Cambridge University Press, 117–142.

Holyoak, K. J. (2012). Analogy and relational reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning*. New York: Oxford University Press, 234–259.

Hsu, C.-Y., Kalyuga, S., & Sweller, J. (2015). When should guidance be presented in physics instruction? *Archives of Scientific Psychology*, *3*(1), 37–53. doi: 10.1037/arc0000012

Huang, X. (2007). *The effect of different types of worked examples on student learning and transfer of a problem-solving task*. Electronic Theses, Treatises and Dissertations. Paper 3679.

Ireson, J. (2008). *Learners, learning and educational activity*. Routledge.

Jonassen, D. H. (1997). Instructional design models for well-structured and iII-structured problem-solving learning outcomes. *Educational Technology Research and Development*, *45*(1), 65–94. doi: 10.1007/BF02299613

Kalyuga, S. (2011). Cognitive load theory. How many types of load does it really need? *Educational Psychology Review*, *23*(1), 1–19. doi: 10.1007/s10648-010-9150-7

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational psychologist*, *38*(1), 23-31. doi: 10.1207/S15326985EP3801_4

Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications: Introduction to the special issue. *Instructional Science*, *38*(3), 209-215. doi: 10.1007/s11251-009-9102-0

Kant, J. M., Scheiter, K., & Oschatz, K. (2017). How to sequence video modeling examples and inquiry tasks to foster scientific reasoning. *Learning and Instruction*, *52*, 46-58. doi: 10.1016/j.learninstruc.2017.04.005

Kapur, M. (2008). Productive failure. *Cognition and instruction*, *26*(3), 379-424. doi: 10.1080/07370000802212669

Kapur, M. (2012): Productive failure in learning the concept of variance. *Instructional Science*, *40*(4), 651–672. doi: 10.1007/s11251-012-9209-6

Kapur, M. (2014). Productive failure in learning math. *Cognitive Science*, *38*(5), 1008-1022. doi: 10.1111/cogs.12107

King, A. (2007). Scripting collaborative learning processes: A cognitive perspective. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational, and educational perspectives*. New York: Springer, 13-37.

Kirschner, P.A. & Kirschner, F. (2012). Mental effort. In N. Seel (Ed.), *Encyclopedia of the Sciences of Learning*. New York: Springer, 2182-2184.

Kitchner, K.S. (1983). Cognition, metacognition, and epistemic cognition:A three-level model of cognitive processing. *Human Development, 26*(4), 222-232. doi: 10.1159/000272885

Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods, 45*(4), 1058-1072. doi: 10.3758/s13428-013-0334-1

Leppink, J., Paas, F., Van Gog, T., Van Der Vleuten, C. P., & Van Merrienboer, J. J. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, *30*, 32-42. doi: 10.3758/s13428-013-0334-1

Lim, J., Reiser, R. A., & Olina, Z. (2009). The effects of part-task and whole-task instructional approaches on acquisition and transfer of a complex cognitive skill. *Educational Technology Research and Development, 57*(1), 61-77.

Loibl, K., & Rummel, N. (2014). The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instructional Science*, *42*(3), 305–326. doi: 10.1007/s11251-013-9282-5

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied, 11*(4), 256–265. doi: 10.1037/1076-898X.11.4.256

Mazziotti, C., Loibl, K. & Rummel, N. (2015). Collaborative or individual learning within Productive Failure. Does the social form of learning make a difference? In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.). *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning 24 (CSCL) Conference 2015*. International Society of the Learning Sciences, Inc., Vol 2, 570-575.

Mugford, R., Corey, S., Bennell, C., & Martens, C. (2009). A meta-analysis of the self-explanation effect. Poster presented at the *3rd International Cognitive Load Theory Conference, Heerlen, Netherlands.*

Muldner, K., & Conati, C. (2010). Scaffolding meta-cognitive skills for effective analogical problem solving via tailored example selection. *International Journal of Artificial Intelligence in Education*, *20*(2), 99-136.

Mwangi, W. & Sweller, J. (1998). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognition and Instruction*, *16*(2), 173–199.

Nastasi, B. K. & Clements, D. H. (1992). Social-cognitive behaviors and higher-order thinking in educational computer environments. *Learning and Instruction*, *2*(3), 215–238. doi: 10.1016/0959-4752(92)90010-J

Newell, A. & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nokes-Malach, T. J., Van Lehn, K., Belenky, D. M., Lichtenstein, M., & Cox, G. (2013). Coordinating principles and examples through analogy and self-explanation. *European Journal of Psychology of Education*, *28*(4), 1237–1263. doi: 10.1007/s10212-012-0164-z

Nückles, M., Hübner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and instruction*, *19*(3), 259-271. doi: 10.1016/j.learninstruc.2008.05.002

Opfermann, M. (2008). *There's more to it than instructional design: The role of individual learner characteristics for hypermedia learning.* Wissensprozesse und digitale Medien: Bd. 11. Berlin: Logos-Verlag.

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of educational psychology*, *84*(4), 429-434. doi: 10.1037/0022-0663.84.4.429

Paas, F., Ayres, P., & Pachman, M. (2008). Assessment of cognitive load in multimedia learning: theory, methods and applications. In D. H. Robinson, & G. Schraw (Eds.), *Recent innovations in educational technology that facilitate student* learning. (Current Perspectives on Cognition Learning and Instruction). Charlotte, NC: Information Age Publishing, 11-35.

Paas, F. & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The cambridge handbook of multimedia learning* (2. ed.). Cambridge, MA: Cambridge University Press, 27-42.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist*, *38*(1), 63-71. doi: 10.1207/S15326985EP3801_8

Paas, F. & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive load approach. *Journal of Educational Psychology*, *86*(1), 122–133. doi: 10.1037/0022-0663.86.1.122

Paas, F. G., Van Merriënboer, J. J., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and motor skills*, *79*(1), 419-430. doi: 10.2466/pms.1994.79.1.419

Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist, 27*(1), 91-109. doi: 10.1207/s15326985ep2701_7

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., ... Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences, 13*(3), 337-386. doi: 10.1207/s15327809jls1303_4

Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction, 16*(2), 92-103. doi: 10.1016/j.learninstruc.2006.02.008

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

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist*, *38*(1), 15–22. doi: 10.1207/S15326985EP3801_3

Renkl, A., Atkinson, R. K., Maier, U. H., & Staley, R. (2002). From example study to problem solving: Smooth transitions help learning. *Journal of Experimental Education, 70*(4), 293–315. doi: 10.1080/00220970209599510

Renkl, A., Gruber, H., Weber, S., Lerche, T., & Schweizer, K. (2003). Cognitive Load beim Lernen aus Lösungsbeispielen. *Zeitschrift für Pädagogische Psychologie, 17*(2), 93-101. doi: 10.1024//1010-0652.17.2.93

Renkl, A. (2014). Toward an instructionally oriented theory of example-based learning. *Cognitive Science*, *38*(1), 1-37. doi: 10.1111/cogs.12086

Rittle-Johnson, B. & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, *99*(3), 561–574. doi: 10.1037/0022-0663.99.3.561

Roelle, J., Nowitzki, C., & Berthold, K. (2017). Do cognitive and metacognitive processes set the stage for each other? *Learning and Instruction*, *50*, 54–64. doi: 10.1016/j.learninstruc.2016.11.009

Rummel, N. & Spada, H. (2005). Learning to collaborate: an instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences, 14*(2), 201–241. doi: 10.1207/s15327809jls1402_2

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning*, *4*(1), 69-92. doi: 10.1007/s11412-008-9054-4

Schabram, K. (2007). *Lernaufgaben im Unterricht: Instruktionspsychologische Analysen am Beispiel der Physik.* Duisburg-Essen: Dissertation.

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology, 9*(6), 523-538. doi: 10.1002/acp.2350090605

Schunk, D. H. (1981). Modeling and attribution effects on children´s achievment: A self-efficacy analysis. *Journal of Educational Psychology, 73*(1), 93-105. doi: 10.1037/0022-0663.73.1.93

Schwartz, D. L. & Bransford, J. D. (1998). A time for telling. *Cognition and instruction, 16*(4), 475-5223.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*(4), 759-775. doi: 10.1037/a0025140

Schwartz, D. L. & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*(2), 129–184. doi: 10.1207/s1532690xci2202_1

Schworm, S. & Renkl, A. (2006). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers & Education*, *46*(4), 426-445. doi: 10.1016/j.compedu.2004.08.011

Schworm, S. & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology, 99*(2), 285–296. doi: 10.1037/0022-0663.99.2.285

Shebilske, W., Gildea, K., Freeman, J., & Levchuk, G. (2006). A benchmarked experiential system for training (BEST) and dynamic systems theory. Paper presented at the *50th Annual Meeting of the Human Factors and Ergonomics Society*.

Simon, H.A. (1999): *Problem solving*. In: R. A. Wilson & F. C. Keil (Eds.): The MIT encyclopedia of the cognitive sciences. Cambridge, MA: MIT Press, 674-676.

Singley M. & Anderson J.R. (1989). *Transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Stark, R. (1999). *Lernen mit Lösungsbeispielen*. Göttingen: Hogrefe-Verlag.

Stark, R., Gruber, H., Renkl, A., & Mandl, H. (2000). Instruktionale Effekte einer kombinierten Lernmethode. Zahlt sich die Kombination von

Lösungsbeispielen und Problemlöseaufgaben aus? *Zeitschrift für Pädagogische Psychologie, 14*(4), 206–218. doi: 10.1024//1010-0652.14.4.206

Stark, R., Tyroller, M., Krause, U.-M., & Mandl, H. (2008). Effekte einer metakognitiven Promptingmaßnahme beim situierten, beispielbasierten Lernen im Bereich Korrelationsrechnung. *Zeitschrift für Pädagogische Psychologie, 22*(1), 59–71. doi: 10.1024/1010-0652.22.1.59

Stegmann, K., Pilz, F., Siebeck, M., & Fischer, F. (2012). Vicarious learning during simulations: is it more effective than hands-on training? *Medical education, 46*(10), 1001-1008. doi: 10.1111/j.1365-2923.2012.04344.x

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive science*, *12*(2), 257-285. doi: 10.1207/s15516709cog1202_4

Sweller, J. (2010). Cognitive Load Theory: Recent Theoretical Advances. In J. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive Load Theory*. New York: Cambridge University Press, 29-47.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Sweller, J. & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59-89. doi: 10.1207/s1532690xci0201_3

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–295. doi: 10.1023/A:1022193728205

Trafton, J. G. & Reiser, B. J. (1993). The contributions of studying examples and solving problems to skill acquisition. In M. Polson (Ed.), *The Proceedings of* *the Fifteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum, 1017–1022.

Van der Meij, H., Rensink, I., & Van der Meij, J. (2017). Effects of practice with videos for software training. *Computers in Human Behavior*. doi: 10.1016/*j*.chb.2017.11.029

Van Gog, T. (2011). Effects of identical example–problem and problem–example pairs on learning. *Computers & Education*, *57*(2), 1775-1779. doi: 10.1016/j.compedu.2011.03.019

Van Gog, T. & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*, *43*(1), 16-26. doi: 10.1080/00461520701756248

Van Gog, T., Paas, F., & Van Merriënboer, J. J. (2006). Effects of process-oriented worked examples on troubleshooting transfer performance. *Learning and Instruction, 16*(2), 154-164. doi: 10.1016/j.learninstruc.2006.02.003

Van Gog, T. & Kester, L. (2012). A test of the testing effect: Acquiring problem-solving skills from worked examples. *Cognitive Science, 36*(8), 1532–1541. doi: 10.1111/cogs.12002

Van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem and problem-example pairs on novices' learning. *Contemporary Educational Psychology, 36*(3), 212-218. doi: 10.1016/j.cedpsych.2010.10.004

Van Gog, T., Kirschner, F., Kester, L., & Paas, F. (2012). Timing and frequency of mental effort measurement: Evidence in favor of repeated measures. *Applied Cognitive Psychology, 26*(6), 833-839. doi: 10.1002/acp.2883

Van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social cognitive research perspectives. *Educational Psychology Review*, *22*(2), 155-174. doi: *10.1007/s10648-010-9134-7*

Van Lehn, K. (1996). Cognitive skill acquisition. *Annual review of psychology*, *47,* 513–539. doi: 10.1146/annurev.psych.47.1.513

Van Lehn, K. (1998). Analogy events: How examples are used during problem solving. *Cognitive Science, 22*(3), 347-388. doi: 10.1207/s15516709cog2203_4

Van Lehn, K., & Jones, R. M. (1993). Better learners use analogical problem solving sparingly. In P. E. Utgoff (Ed.), *Proceedings of the Tenth International Conference on Machine Learning*. San Mateo, CA: Morgan Kaufmann, 338-345.

Van Merriënboer, J. J., Clark, R. E., & De Croock, M. B. (2002). Blueprints for complex learning: The 4C/ID-model. *Educational technology research and development, 50*(2), 39-61. doi: 10.1007/BF02504993

Wichmann, A. & Leutner, D. (2009). Inquiry learning: Multi-level support in terms of inquiry, explanations and regulation during an inquiry cycle. *Zeitschrift für Pädagogische Psychologie*, *23*, 117–127. doi: 10.1024/1010-0652.23.2.117

Wood, P.K. (1983). Inquiring systems and problem structures: Implications for cognitive development. *Human Development, 26*(5), 249-265. doi: 10.1159/000137808

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2), 64-70. doi: 10.1207/s15430421tip4102_2

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology*, *94*(4), 660-668. doi: 10.1037/0022-0663.94.4.660

Tables

Figures

Versicherung an Eides statt

(gemäß § 8 Abs. 2 Nr. 4 Promotionsordnung für die Fakultäten 09, 10, 11, 12 und 13 (2016) vom 15. September 2016)

Name, Vorname: Murböck, Julia

Hiermit versichere ich an Eides statt, dass die Dissertation mit dem Titel „The order matters: Sequencing strategies in example-based learning“ selbständig und ohne unerlaubte fremde Hilfe angefertigt und keine anderen, als die von mir angegebenen Schriften und Hilfsmittel benutzt wurden. Die den benutzten Werken wörtlich und inhaltlich entnommenen Stellen sind kenntlich gemacht.

Julia Murböck

München, 20.03.18___________________________________

Ort, Datum, Unterschrift Doktorandin