Learning by Explaining

How Implementation- and Student-Related Boundary Conditions Determine the Effectiveness of Generating an Explanation to a Fictitious Peer

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Leonie Jacob
aus Friedberg

Tübingen, 2021

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	23.06.2021
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Andreas Lachner
2. Berichterstatter:	Prof. Dr. Stephan Schwan

To teach is to learn twice.

Joseph Joubert

Acknowledgements

I would like to thank everyone who supported me during my dissertation. First, I would like to thank my first supervisor, Prof. Dr. Andreas Lachner, whose determination has always inspired me to pursue my own goals. Thank you, Andi, for all your time, expedient feedback, support, and for sharing your knowledge with me. Secondly, I would like to thank my second supervisor Prof. Dr. Katharina Scheiter, whose constructive and strategic way of thinking significantly improved my work. Thank you, Katharina, for your ideas, input, and quick feedback, through which I could understand the larger context of my work.

I would like to thank the Leibniz-Institut für Wissensmedien which provided me excellent working conditions and helped me in conducting my studies. In this context, I would also like to thank the LEAD Graduate School and Research Network and the Tübingen School of Education, which provided me with great networking opportunities. Additionally, I would also like to thank all the participating students and pupils who took part in my studies, and the school administrators, teachers, and parents who made these participations possible. I would also like to use this opportunity to thank all research assistants who helped me in conducting the studies and coding the data.

Furthermore, I would like to thank all my colleagues who supported me during my PhD. I would especially like to thank Iris Backfisch for sharing her time both inside and outside of work. Thank you, Iris, for the great years together, all your support, and the friendship that has developed. Additionally, I would also like to thank Nora Umbach, Samuel Merk, Thérése Eder, and Jürgen Schneider who were always enthusiastic in discussing analyses with me. I would also like to thank Richard Fair and Dr. Piotr Małysz who provided excellent feedback on my thesis.

Finally, I would like to say a special thank you to Scott Sibley, who supported me emotionally through my PhD but also through constructive feedback. Thank you, Scott, for your unconditional love, all your time, and your ability to build me up in difficult moments. Also, I would like to thank my family and all my friends who supported me through my PhD with their love, care packages, time, and by reminding me of my life outside of academia.

Thank you all!

Table of Contents

Summary

Gaining conceptual knowledge in science is regarded as a fundamental goal in education (BMBF, 2019; KMK, 2009). However, several studies have shown only rudimentary and superficial conceptual knowledge in STEM (Science, Technology, Engineering, Mathematics) subjects (MINT Nachwuchsbarometer, 2020; Reiss et al., 2019). Conceptual knowledge is considered a crucial requirement to deeply understand new information and to transfer it to other contexts (Gruber et al., 2000; Jong & Ferguson-Hessler, 1996; Mandl et al., 1994). Generative learning strategies may trigger students to deeply process new content and to monitor their learning, which has been shown to enhance students' conceptual knowledge and their monitoring accuracy (Brod, 2020; Fiorella & Mayer, 2016; Fukaya, 2013). In this context, one strategy received attention in both research and practice: learning by explaining to fictitious peers (Hoogerheide, Visee, et al., 2019). Generating an explanation to a fictitious peer is regarded as a beneficial learning strategy, as students engage in cognitive and metacognitive processes, which should result in higher levels of conceptual knowledge (Fiorella & Mayer, 2014; Lachner et al., 2021). However, prior research demonstrated mixed findings regarding the effectiveness of explaining to fictitious peers (e.g., Fiorella & Mayer, 2013, 2014; Fukaya, 2013; Hoogerheide et al., 2014; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019; Lachner et al., 2020). Additionally, large variances among studies were reported, which might highlight that the effectiveness of explaining to fictitious peers depends on boundary conditions (Kobayashi, 2019; Lachner et al., 2021). Such boundary conditions, however, have not yet been investigated systematically. In addition, little is known about the underlying mechanism of explaining to fictitious others. This dissertation aims to close both research gaps. In a first step, a theoretical framework model was generated based on prior research that considered implementation-related (i.e., explanatory modality, text complexity, social presence) and student-related (i.e., prior knowledge, academic self-concept) boundary conditions. Second, three underlying mechanisms of explaining were explored and included in the framework model, which were derived from previous research (i.e., retrieval practice hypothesis, generative learning hypothesis, social presence hypothesis). Lastly, the framework model was empirically tested within three studies.

In Study 1, I investigated whether text complexity moderates the effectiveness of generating oral versus written explanations regarding students' conceptual knowledge (i.e., factual knowledge, transfer knowledge) and their monitoring accuracy. University students (*N* = 115) studied a complex versus a simple text and then explained the content to a fictitious peer in either oral or written form. In a control group, students engaged in a retrieval practice activity. Results revealed that explaining is only beneficial when the provided text is complex, but not when it entails a low level of complexity regarding both students' transfer knowledge and their monitoring accuracy. Additionally, results showed that oral explaining was more beneficial than writing explanations. The explanatory effect was mediated by students' perceived social presence as students who explained orally showed higher levels of perceived social presence, which triggered students to generate more comprehensive explanations and resulted in higher transfer knowledge. In Study 2, I conducted a replication study with an additional writing condition in which perceived social presence of the fictitious peer was induced. In contrast to Study 1, results revealed no differences among conditions regarding students' conceptual knowledge and monitoring accuracy. In Study 3, I investigated the impact of students' prerequisites and tested whether school students' prior knowledge and academic self-concept moderated the effectiveness of explaining in oral versus written form. Results showed that academic self-concept, but not prior knowledge moderated the explanatory effect, as only students with low academic self-concept benefited from explaining regarding their factual knowledge. Finally, I discussed whether explaining to fictitious others is effective by conducting two additional meta-analyses. Results revealed a small positive effect of explaining on students' factual knowledge but not on their transfer knowledge. Additionally, results of the dissertational studies showed that the explaining effect on students' comprehension was mediated by students' perceived social presence during explaining. However, simply inducing social presence did not result in higher learning outcomes. Finally, results revealed that oral explaining was more beneficial than writing explanations, and that students' academic self-concept but not their prior knowledge moderated the explaining effect.

In summation, this dissertation provides a systematic investigation of potential boundary conditions of learning by explaining to fictitious peers and additionally investigated underlying mechanisms of explaining. Results of this dissertation revealed that explaining is a beneficial learning strategy to enhance students' conceptual knowledge in STEM subjects, however, it is crucial to consider explanatory modality and students’ academic self-concept as boundary conditions of explaining.

Zusammenfassung

Der Erwerb von konzeptionellem Wissen in den Naturwissenschaften wird als grundlegendes Ziel im Unterricht erachtet (BMBF, 2019; KMK, 2009). Mehrere Studien zeigten jedoch lediglich rudimentäres Wissen der Lernenden in naturwissenschaftlichen Fächern, wie den MINT-Fächern (MINT Nachwuchsbarometer, 2020; Reiss et al., 2019). Dabei gilt konzeptuelles Wissen als Voraussetzung, um neue Informationen tiefgreifend verstehen und in andere Kontexte übertragen zu können (Gruber et al., 2000; Jong & Ferguson-Hessler, 1996; Mandl et al., 1994). Generative Lernstrategien können Lernende anregen, neue Inhalte tiefgehend zu verarbeiten und ihr Lernen zu überwachen, was zu konzeptuellem Wissen und besserer Überwachungsgenauigkeit führen soll (Brod, 2020; Fiorella & Mayer, 2016; Fukaya, 2013). In diesem Zusammenhang hat vor allem eine Strategie sowohl in der Forschung als auch in der Praxis Beachtung gefunden: die Generierung einer Erklärung für fiktive Personen (Hoogerheide, Visee, et al., 2019). Das Generieren einer Erklärung wird als effektive Lernstrategie angesehen, da Lernenden zu kognitiven und metakognitiven Prozessen angeregt werden, welche zu höherem konzeptuellen Wissen führen sollten (Fiorella & Mayer, 2014; Lachner et al., 2021). Die bisherige Forschung zeigte jedoch gemischte Ergebnisse hinsichtlich der Effektivität des Erklärens gegenüber fiktiven Personen (z. B. Fiorella & Mayer, 2013, 2014; Fukaya, 2013; Hoogerheide et al., 2014; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019; Lachner et al., 2020). Zudem wurden große Varianzen zwischen den Studien berichtet, was darauf hinweisen könnte, dass Erklären nicht zwangsläufig effektiv ist, sondern von weiteren Bedingungen abhängt (Kobayashi, 2019; Lachner et al., 2021). Solche Bedingungen sind jedoch noch nicht systematisch untersucht worden. Darüber hinaus ist noch wenig über den zugrundeliegenden Mechanismus des Erklärens gegenüber fiktiven Personen bekannt. Diese Dissertation zielt darauf ab, beide Forschungslücken zu schließen. In einem ersten Schritt wurde hierfür auf Basis vorheriger Forschung ein theoretisches Rahmenmodell generiert, das implementationsbezogene (Erklärungsmodalität, Textkomplexität, soziale Präsenz) und Lerner bezogene (Vorwissen, akademisches Selbstkonzept) Bedingungen berücksichtigt. Zweitens wurden drei potenzielle Erklärungsmechanismen aus der vorherigen Forschung abgeleitet und in das Rahmenmodell integriert (Retrieval-Practice-Hypothese, Generative-

Lern-Hypothese, Soziale-Präsenz-Hypothese). Schließlich wurde das Rahmenmodell innerhalb drei Studien empirisch getestet.

In Studie 1 wurde untersucht, ob die Textkomplexität die Effektivität der Generierung von mündlichen oder schriftlichen Erklärungen hinsichtlich des konzeptuellen Wissens (Faktenwissen, Transferwissen) der Lernenden und ihrer Überwachungsgenauigkeit moderiert. Universitätsstudierende (N = 115) lasen einen komplexen oder einfachen Text und erklärten den Inhalt anschließend einer fiktiven Person entweder in mündlicher oder schriftlicher Form. In einer Kontrollgruppe waren die Studierenden mit einer Abrufaufgabe beschäftigt. Die Ergebnisse zeigten, dass das Erklären nur dann effektiv war, wenn der Lerntext komplex war, nicht aber, wenn er eine niedrige Komplexität aufweist. Dieser Effekt zeigte sich sowohl hinsichtlich des Transferwissens als auch der Überwachungsgenauigkeit der Studierenden. Zudem war mündliches Erklären effektiver als schriftliches Erklären. Dieser Erklärungseffekt wurde durch die wahrgenommene soziale Präsenz mediiert. Studierende, die mündlich erklärten, nahmen die soziale Präsenz der fiktiven Person stärker wahr, was die Studierenden dazu veranlasste, umfassendere Erklärungen zu generieren und schlussendlich zu höheren Transferwissen führte. Studie 2 (N = 137) war eine Replikationsstudie, mit einer zusätzlichen Schreibbedingung, in der die wahrgenommene soziale Präsenz der fiktiven Person stärker induziert wurde. Im Gegensatz zu Studie 1 zeigten die Ergebnisse keine Unterschiede zwischen den Bedingungen. In Studie 3 (N = 129) wurde der Einfluss der Voraussetzungen von Schülerinnen und Schülern der achten Klassenstufe untersucht und getestet, ob das Vorwissen und das akademische Selbstkonzept die Effektivität des Erklärens in mündlicher oder schriftlicher Form moderierten. Die Ergebnisse zeigten, dass das akademische Selbstkonzept, aber nicht das Vorwissen den Erklärungseffekt moderierte, da nur Schülerinnen und Schüler mit niedrigem akademischen Selbstkonzept vom Erklären hinsichtlich ihres Faktenwissens profitierten. Abschließend wurden innerhalb der Diskussion zwei Meta-Analysen durchgeführt, um die Effektivität des Erklärens gegenüber fiktiven Personen zu überprüfen. Die Ergebnisse zeigten einen kleinen positiven Effekt des Erklärens auf das Faktenwissen, aber nicht auf das Transferwissen. Zudem zeigten die Ergebnisse der Dissertationsstudien, dass der Effekt des Erklärens durch die wahrgenommene soziale Präsenz mediiert wurde. Die bloße Induktion sozialer Präsenz führte jedoch nicht zu höheren Lernergebnissen. Weiterhin zeigten die Ergebnisse, dass mündliches Erklären effektiver war als schriftliches Erklären und dass das akademische Selbstkonzept der Lernenden, nicht aber ihr Vorwissen, den Erklärungseffekt moderierte.

ZUSAMMENFASSUNG

Zusammenfassend lässt sich sagen, dass diese Dissertation eine systematische Untersuchung der Bedingungen des Lernens durch Erklären gegenüber fiktiven Personen aufzeigt und zusätzlich die zugrundeliegenden Mechanismen untersuchte. Die Ergebnisse zeigten, dass das Erklären eine effektive Lernstrategie sein kann, um das konzeptuelle Wissen der Lernenden in MINT-Fächern zu fördern, jedoch ist es entscheidend, die Erklärungsmodalität und das akademische Selbstkonzept zu berücksichtigen.

List of Publications and Contributions

The content of Study 1 was published in the journal *Learning and Instruction* in 2020. The proportional contributions of the different co-authors to the manuscript are presented in the subsequent table.

Author	Author position	Scientific ideas %	Data generation %	Analysis & interpretation %	Paper writing %
Leonie Jacob	first author	35 %	90 %	60 %	60 %
Andreas Lachner	second author	55 %	10 %	35 %	30 %
Katharina Scheiter	third author	10 %	0 %	5 %	10 %

Title of paper: Learning by Explaining Orally or in Written Form? Text Complexity Matters

Publication process: Published: *Learning and Instruction*, 68, 101344
https://doi.org/10.1016/j.learninstruc.2020.101344

The content of Study 2 was published in the journal *PLOS ONE* in 2021. The proportional contributions of the different co-authors to the manuscript are presented in the subsequent table.

Author	Author position	Scientific ideas %	Data generation %	Analysis & interpretation %	Paper writing %
Leonie Jacob	first author	65 %	90 %	80 %	70 %
Andreas Lachner	second author	25 %	5 %	15 %	20 %
Katharina Scheiter	third author	10 %	5 %	5 %	10 %

Title of paper: Does Increasing Social Presence Enhance the Effectiveness of Writing Explanations?

LIST OF PUBLICATIONS

Publication process: Published: PLOS ONE, 16(4), e0250406
https://doi.org/10.1371/journal.pone.0250406

The content of Study 3 received a major revision from the journal *Computers & Education*. The proportional contributions of the different co-authors to the manuscript are presented in the subsequent table. This article may not exactly replicate the final version published in the journal.

Author	Author position	Scientific ideas %	Data generation %	Analysis & interpretation %	Paper writing %
Leonie Jacob	first author	65 %	80 %	65%	60 %
Andreas Lachner	second author	20 %	10 %	30 %	30 %
Katharina Scheiter	third author	15 %	10 %	5 %	10 %

Title of paper:	What’s App? School Students’ Self-Concept Moderates the Effectiveness of Generating Technology-Mediated Explanations
Publication process:	Major revision: Computers & Education

1

INTRODUCTION

1. Introduction

1.1 Problem Statement

Scientific breakthroughs consistently impact humans lives (KMK, 2009). For instance, technical inventions, such as electricity and digitalization, have increased humans' life expectancy and quality of life over the past centuries (BMBF, 2019). Science has made these achievements possible and is therefore considered as a crucial requirement for society's health (BMBF, 2019). Hence, the political interest in science education, technology, engineering, and mathematics – also known as *STEM* subjects – increased drastically in recent decades. For instance, several foundations and institutions began to support STEM education in schools and universities in Germany by providing financial support and educational programs (BMBF, 2019; see Köller et al., 2019, for an example). To successfully engage in STEM-related education, such as reasoning or problem solving, students need to build high levels of scientific literacy, which is defined as "the ability to engage with science-related issues" (OECD, 2018a, p. 4) and contains three domain-specific competencies: explaining phenomena scientifically, interpreting data and evidence scientifically, and evaluating and designing scientific enquiry (OECD, 2017, 2018b). Thus, students need to understand the scientific phenomena on a deep conceptual level to be able to draw inferences between the provided processes and systems and to transfer and apply them in different settings (KMK, 2020). Unfortunately, international comparative studies such as PISA (Program for International Student Assessment) reported that every fifth student showed only rudimentary and superficial conceptual knowledge in STEM subjects (MINT Nachwuchsbarometer, 2020). For example, students were able to reproduce learned content but could not build inferences between different concepts. This rudimentary conceptual knowledge does not meet the requirements for higher education in scientific areas. These alarming findings highlight that many students are not prepared to study STEM subjects at university. Dropout rates at university confirm this concern as every third student who started to study a STEM subject left university before graduating (Heublein et al., 2017). None of the other domains showed such high dropout rates as STEM domains. Students reported that the main reason to leave university was that they were not able to meet the high requirements (Heublein et al., 2017).

Students' superficial conceptual knowledge in schools and the high dropout rates at universities confront the educational system with an urgent need to act. The first approach to handle this problem is to analyze why STEM subjects are challenging for students. One reason might be that students are confronted with very difficult learning materials; STEM subjects mainly contain complex processes and difficult systems, which are challenging to process adequately (IET, 2008). Learning difficult content – especially from texts – challenges students and reveals only low levels of comprehension (Ozuru et al., 2010). Additionally, students are required to not only reproduce the content on a superficial level but also to deeply process the content and gain meaningful conceptual knowledge. However, students are unable to achieve deep conceptual knowledge through memorization since they are not forced to think beyond the given information. Therefore, they need to apply additional learning strategies. However, students fail in choosing and applying adequate learning strategies (Brod, 2020): First, students show few competencies in judging their current comprehension correctly (Brod, 2020; Schleinschok et al., 2017),which is a crucial ability to choose adequate learning strategies to enhance their comprehension. Second, even when they choose a learning strategy, they often fail in implementing it properly (Brod, 2020; Jairam & Kiewra, 2010). Therefore, they depend on external support which helps them to implement efficient learning strategies. In this context, *generative learning strategies* are regarded to help students to efficiently process new learning content (Brod, 2020; Fiorella & Mayer, 2016).

Generative learning strategies have been shown to enable students to think beyond the provided learning content and to increase their comprehension (see Brod, 2020; Fiorella & Mayer, 2016, for comprehensive reviews). These strategies trigger students to actively elaborate new content by forcing students to connect new information with their prior knowledge, which has been shown to result in higher learning outcomes (Chi et al., 1994; Chi, 2009; Chi & Wylie, 2014; Fiorella & Mayer, 2016). Additionally, these strategies help students to monitor their current learning process, which has been shown to enhance students' learning as well (Anderson & Thiede, 2008; Fiorella & Mayer, 2016; Fukaya, 2013; Schleinschok et al., 2017; Thiede et al., 2003b, 2003a). Therefore, generative learning strategies might be efficient strategies to support students during learning challenging STEM content. In recent years, one generative learning strategy has received significant attention in educational settings and research: *explaining* learned information. Explaining learned content can be implemented as a learning strategy in distinct ways. For instance, students can explain new information in an interactive setting. Thus, they interact with other students during the explanation. Interactive settings are very promising learning settings as they enable students to exchange ideas, to profit from each other's thoughts, and to reflect their own knowledge. Several studies provided empirical evidence of positive effects of explaining in interactive settings on students' comprehension (Chi et al., 1989; Chi et al., 1994; Chi & VanLehn Kurt, 1991; Dillenbourg, 1999; Duran, 2016; McNamara, 2004, 2017; McNamara & Scott, 1999; Plötzner et al., 1999; Roscoe, 2014; Topping, 2005; Webb et al., 1995; Webb et al., 2009). Interestingly, more recent research started to investigate the potential effect of explaining to a fictitious peer: an absent and non-active peer (Fiorella & Kuhlmann, 2020; Fukaya, 2013; Hoogerheide et al., 2014; Hoogerheide et al., 2016; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019; Lachner et al., 2018; Lachner et al., 2020). However, these studies reported mixed findings regarding both students' comprehension and monitoring accuracy (e.g., Fukaya, 2013; Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021). Therefore, it is still unclear whether generating an explanation to fictitious peers enhances students' judgment of their current understanding and their comprehension.

Interestingly, prior research on explaining to fictitious others showed a high variance among studies (see Kobayashi, 2019; Lachner et al., 2021, for meta-analytic evidence). This variance across studies can be regarded as an indicator that generating an explanation to fictitious others depends on further boundary conditions. These boundary conditions might be implementation-related, such as the complexity of the learning material (e.g., Ozuru et al., 2010) or student-related, such as students' prior knowledge (see McNamara & Scott, 1999). But potential boundary conditions have not been systematically analyzed in prior research. Additionally, still little is known about the underlying mechanism that causes the explaining effect on students' comprehension. In this context, recent research discussed three different hypotheses. First, researchers stated that the explaining effect only occurs because students need to retrieve the information from memory during explaining (i.e., *retrieval practice hypothesis*; see Koh et al., 2018). Second, researchers argued that the effect of explaining results from combining new information with prior knowledge, which results in the generation of new mental knowledge structures that enhances learning (i.e., *generative learning hypothesis*). Third, researchers discussed that the explaining effect is strongly determined by students' perceived presence of the (fictitious) peer during explaining, which may trigger students to adapt their knowledge to the audience's needs (i.e., adaptive and elaborative processes) and may be more motivating than restudying content (i.e., *social presence hypothesis*; Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021). Prior research has not yet solved the question regarding the responsible underlying processes. Knowing students' underlying processes while applying learning strategies would have a strong impact on how tasks should be designed to explicitly trigger these processes and, therefore, needs to be examined in detail.

In summation, gaining deep conceptual knowledge is essential to develop scientific literacy in science. However, students showed only rudimental and superficial conceptual knowledge in STEM subjects (MINT Nachwuchsbarometer, 2020). Therefore, students should be provided with effective learning strategies to help them process complex learning content more easily (Brod, 2020). In this context, learning by explaining to fictitious peers is considered a beneficial strategy to enhance students' monitoring accuracy and their comprehension (Brod, 2020; Fiorella & Mayer, 2016). However, prior research only demonstrated mixed results regarding its effectiveness. In this context, potential boundary conditions of learning by explaining might play a crucial role in explaining these mixed findings, as they may determine its effectiveness. Additionally, little is known about the underlying mechanism, which complicates the understanding of learning by explaining.

1.2 Aims of the Dissertation

The aim of this dissertation is twofold. First, potential boundary conditions of learning by explaining are subjected to empirical analysis to investigate the effectiveness of learning by explaining. Second, potential underlying mechanisms are explored. Therefore, the dissertation focusses on two main objectives:

Aim 1 of the dissertation is to systematically analyze potential boundary conditions of learning by explaining to fictitious others. In the first step, I generated a theoretical framework model (i.e., ISEO model: **I**mplementation- and **S**tudent-related Boundary Conditions of **E**xplaining on Students' **O**utcomes) based on reviewing prior research regarding potential boundary conditions. On the one hand, this model contains implementation-related boundary conditions which concern either the learning strategy itself (i.e., the modality of explaining), the learning material (i.e., text complexity), or the learning environment (i.e., perceived social presence during explaining). On the other hand, the model includes student-related boundary conditions such as performance-oriented (i.e., prior knowledge) and belief-oriented (i.e., academic self-concept) prerequisites. In the second step, these boundary conditions were systematically analyzed in three laboratory experiments with university students (i.e., Study 1 and Study 2) and school students (i.e., Study 3).

Aim 2 of the dissertation was to systematically investigate potential underlying processes of explaining to fictitious others. In this context, recent research discussed three possible underlying processes which could cause the explaining effect (a: *retrieval practice hypothesis*; b: *generative learning hypothesis*; c: *social presence hypothesis*). To obtain information about the processes in which students were engaged during learning, it is essential to analyze their generated explanations. For this purpose, step one was to develop a coding system, which was derived from prior literature. An indicator was determined for each hypothesis (a: *concepts*; b: *elaborations*; c: *personal references*) which provides evidence for or against the corresponding hypothesis. In a second step, all generated explanations across all studies were coded based on the generated coding system and analyzed in relation to students' learning outcomes.

Thus, this dissertation provides a theoretical framework model and empirical evidence about the effectiveness of explaining that considers both potential boundary conditions and potential underlying processes of learning by explaining to a fictitious peer.

2

GENERATIVE LEARNING DURING KNOWLEDGE ACQUISITION

2. Generative Learning During Knowledge Acquisition

Generative learning is considered to result in deep conceptual knowledge, as required for scientific literacy. To follow the theory of generative learning, it is necessary to understand how students learn and how knowledge is acquired, stored, and connected in memory. Therefore, I first explain in some detail theoretical foundations and levels of knowledge acquisition, and, related to this, metacognitive monitoring during learning, as fundamental theories of generative learning.

2.1 The Role of Working Memory and Long-Term Memory in Knowledge Acquisition

Knowledge is defined as a mental representation of information reflected in a hierarchical and associative net in which nodes represent concepts or propositions (Ericsson, 2006; Kintsch, 1988; Renkl, 2009). To build such mental nets, students need to cognitively encode, retain, transform, and retrieve new information (Renkl, 2009). During this process, new information is processed in different memory stores; those are *sensory memory*, *working memory*, and *long-term memory* (Atkinson & Shiffrin, 1968; Mayer, 1996). First, new information is captured in the sensory memory for a very short period (see Figure 1). In a second step, when attention is drawn to the captured information, the information will be processed in the working memory (Atkinson & Shiffrin, 1968; Mayer, 1996). The working memory represents a crucial part of knowledge acquisition, since information is coordinated within the working memory (Baddeley, 1992). The capacity of the working memory is highly limited (see Miller, 1956; Peterson & Peterson, 1959; see also relatedly cognitive load theory, e.g., Sweller, 1994; Sweller et al., 2019; Sweller & Chandler, 1991). Thus, information may be forgotten easily without any action of memorizing (Atkinson & Shiffrin, 1968). Through retrieving the information, it might be transferred to the long-term memory where it then represents lasting knowledge which is accessible to the student in the future (Baddeley, 2010; Mayer, 1996). Retrieved information that have been transferred into the long-term memory might include isolated information, such as remembering a telephone number (see Atkinson & Shiffrin, 1968). Or, instead, connecting new information to prior knowledge, which is stored in the long-term memory, may integrate it into existing knowledge structures (Mayer, 1996; Wittrock, 1974). Such integrated information represents integrated and organized knowledge that generally occurs in hierarchical and associative nets, thus, a mental presentation of knowledge (Ericsson, 2006; Kintsch, 1988).

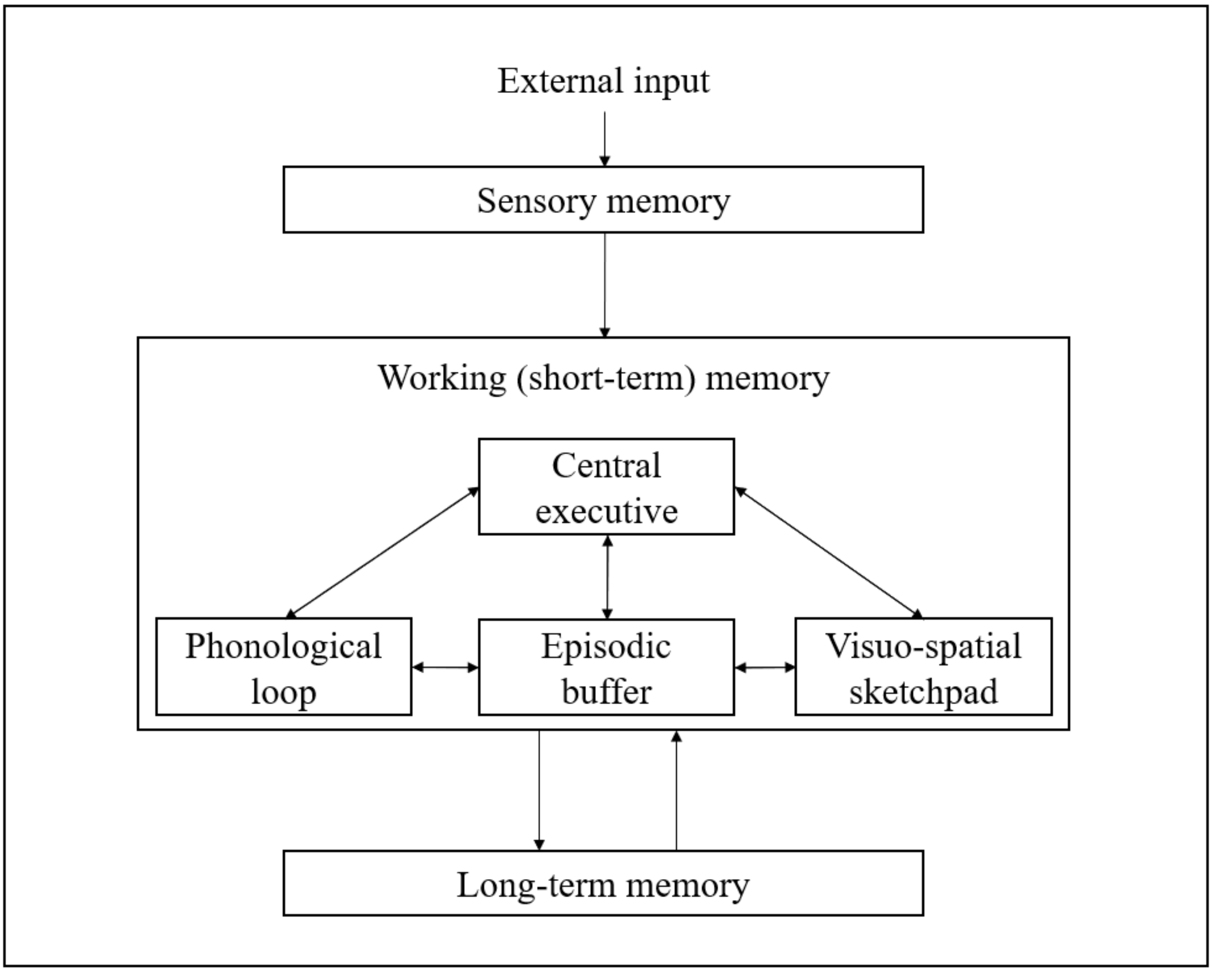

Figure 1. Graphical illustration of the structure of the memory system based on Atkinson and Shiffrin (1968, p. 17) and Baddeley (2010, p. R138).

2.2 Levels of Knowledge Acquisition

Wharton and Kintsch (1991) specified the process of acquiring knowledge in the form of mental presentations in their Construction-Integration Model, which provides more detailed insights into the overall comprehension process (see Figure 2). The authors distinguish between two ordered steps, namely, the *knowledge construction step* and the *knowledge integration step*. The *knowledge construction step* is the process of transforming new information into a crude mental representation network; it contains three parts, namely, *linguistic representation*, *propositional network*, and *elaborated propositional network* (see Figure 2). First, through word recognition, written and spoken words are transformed into a linguistic representation (see also Moreno & van Orden, 2001). Second, these words are connected within sentences, which results in a propositional network. This propositional network represents a construction of a textbase, thus, lower-order knowledge. Prior knowledge, such as past experiences, that are stored in the long-term memory allows the elaboration of these propositions in more detail. Through such elaborative processes, such as the connection between new information and prior knowledge, the elaborated propositional network is built. At this point, comprehension is presented as a “crude mental representation of discourse in form of an associative network of propositional nodes” (Wharton & Kintsch, 1991, p. 169). This elaborated propositional network reflects the situation model which contains elaborated and connected information, thus representing higher-order knowledge (van Dijk & Kintsch, 1992). In a last step, according to Wharton and Kintsch, the *knowledge integration step* is necessary to refine this network (see Figure 2). This integration step is implemented through structuring the elaborated propositional network into a coherent and interpretable whole network, which reflects the final text presentation and might allow drawing inferences between the provided information and related content.

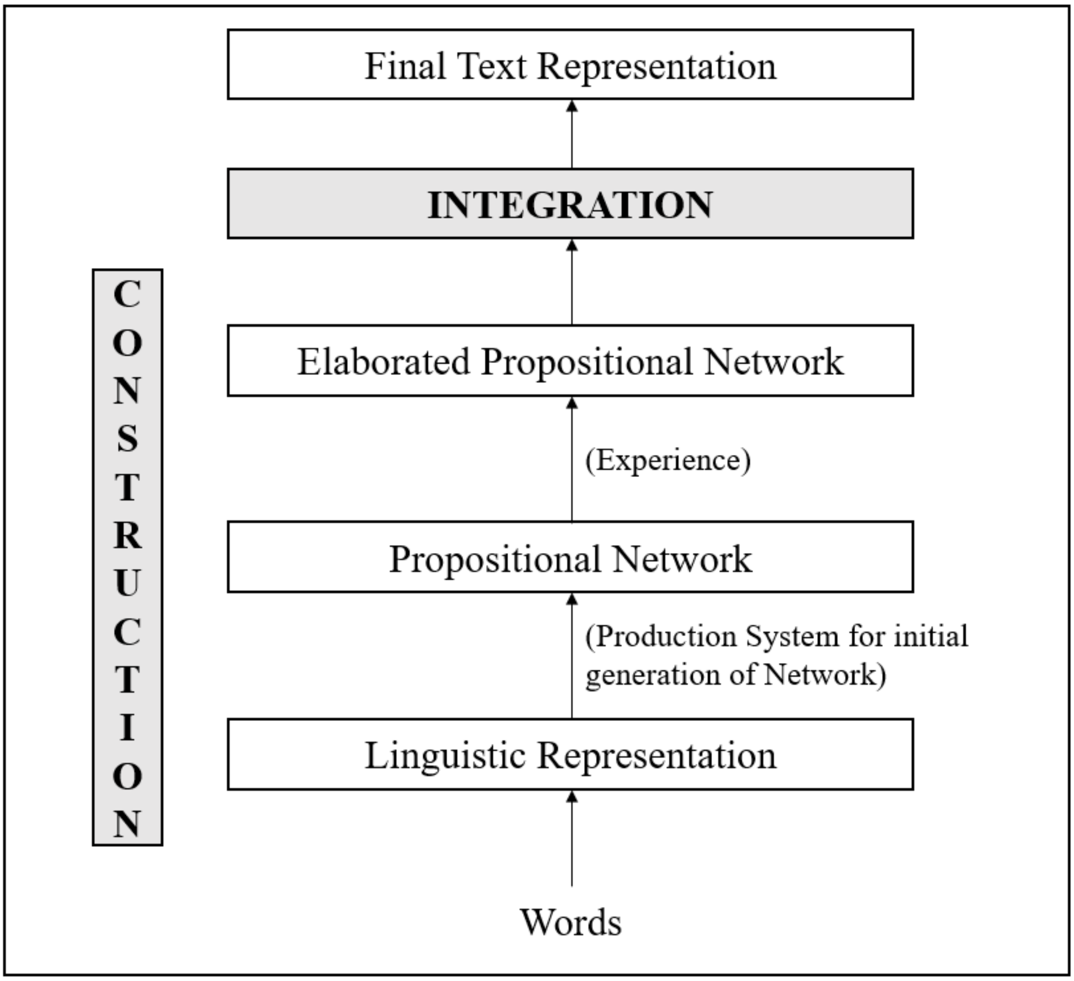

Figure 2. Construction-Integration Model by Wharton and Kintsch (1991, p. 170). Reprinted from ACM SIGART Bulletin, Wharton & Kintsch, *An overview of construction-integration model: A theory of comprehension as a foundation for a new cognitive architecture*, p. 170, Copyright (1991), with permission from the author Walter Kintsch.

In line with Wharton and Kintsch (1991), Jong and Ferguson-Hessler (1996) claim that conceptual knowledge represents different levels of quality that are determined by the level of depth in processing new information (i.e., lower-order versus higher-order conceptual knowledge). Lower-order conceptual knowledge reflects superficial knowledge that might enable students to reproduce learned information (Gräsel et al., 1997; Mandl et al., 1994; Renkl, 1996). For instance, they might recognize the right answer in a multiple choice test without understanding the deeper meaning of the content (Gruber et al., 2000; Mandl et al., 1994). Prior research has defined this lower-order conceptual knowledge in several ways (e.g., factual knowledge, declarative knowledge, inert knowledge, theoretical knowledge; see

Jong & Ferguson-Hessler, 1996; Renkl, 1996, for examples). In this dissertation it is defined as *factual knowledge*. By contrast, higher-order conceptual knowledge represents *meaningful* knowledge (Jong & Ferguson-Hessler, 1996), which results from elaborating new information by connecting it with prior knowledge, which in turn, might enable students to draw inferences between the provided information and related content. They thus build knowledge that goes beyond the provided information. Higher-order conceptual knowledge was defined in diverse ways in prior research (e.g., inference knowledge, transfer knowledge). In this dissertation, it is defined as *transfer knowledge*. Transfer knowledge includes factual knowledge, as students first need to understand the new information from provided learning materials, in order then to understand their relations and purpose. For instance, students who generated transfer knowledge should not only be able to reproduce results of a conducted experiment but also to interpret the data regarding its meaning, significance, and impact on further research. As scientific literacy refers to the ability to think beyond the given information, for instance, to be able to interpret provided data correctly (OECD, 2017, 2018b), students are required to achieve higher-order conceptual knowledge, namely, transfer knowledge.

2.3 Metacognitive Monitoring During Knowledge Acquisition

For acquiring comprehension, metacognitive monitoring skills are essential, as they determine the success of learning new information (Anderson & Thiede, 2008; Schleinschok et al., 2017; Thiede et al., 2003a, 2003b). In their framework, Nelson & Narens (1990) described metacognitive monitoring processes in detail (see Figure 3). The authors argued that learning contains three phases, namely, the *acquisition phase*, the *retention phase*, and the *retrieval phase*. During these phases, students need to apply monitoring and control strategies (in the working memory) to learn new information successfully, in order to be able to store it in their long-term memory. Monitoring one’s own learning process is necessary to (re-)act accordingly by applying adequate learning strategies. For instance, after setting a learning goal, students first need to estimate the required learning time and effort by judging the difficulty of the provided learning content (Nelson & Leonesio, 1988). As a reaction (i.e., control strategy) students then decide which learning strategy to apply to be able to memorize the information. During the acquisition and retention phase, it is crucial that students monitor their understanding accurately to be able to reflect whether learning was successful, that is, whether the information has been stored in the long-term memory, or whether they should restudy the provided materials. This type of monitoring is reflected in students' *Judgment of Learning* (JOL), which is defined as the prospective prediction about future performance (Nelson & Narens, 1990). Ideally, students continue studying until their judgement of learning indicates that they reached their learning goals. Precondition for this statement is an accurate judgement (Thiede et al., 2003b). Accurate judgements of learning are regarded as the ability to judge one's own comprehension correctly (Baars et al., 2017). Correct judgements of their quality of learning may allow students to detect knowledge gaps and to adapt learning strategies according to their needs, for instance, by restudying specific parts of the provided learning material (Baars et al., 2017; Fiorella & Mayer, 2016; Nückles et al., 2009).

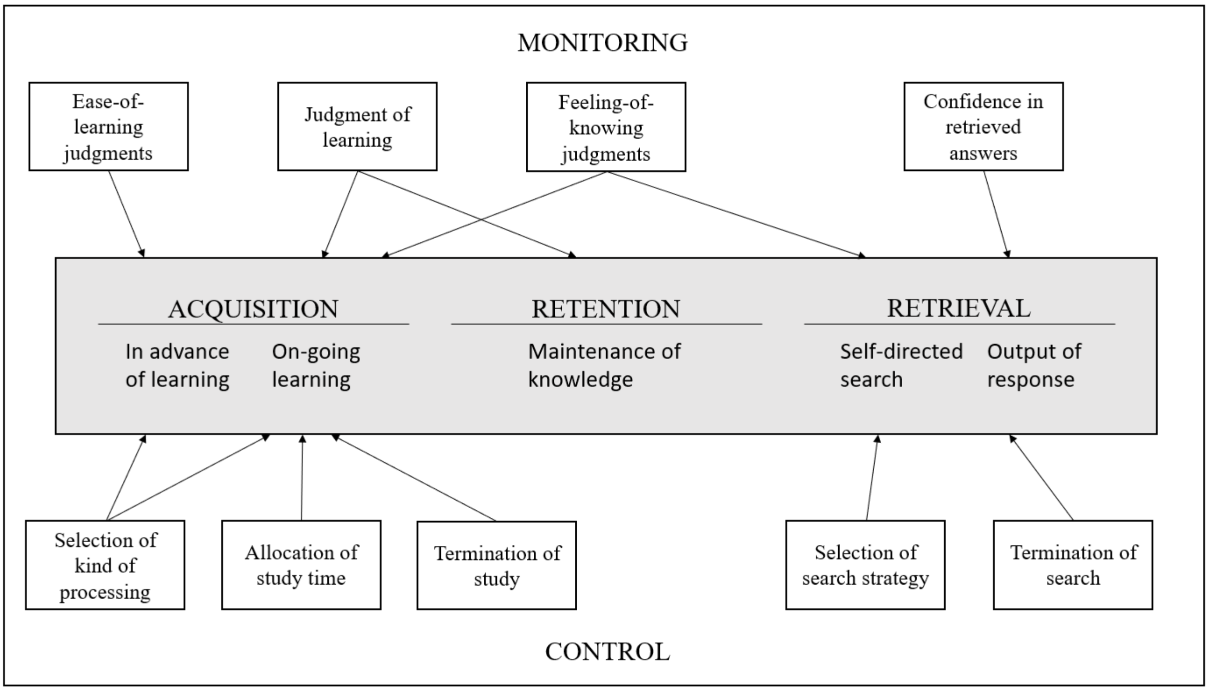

Figure 3. Theoretical memory framework by Nelson & Narens (1990, p. 129). Reprinted from Psychology of Learning and Motivation, Nelson & Narens, *Metamemory: A theoretical framework and new findings*, p. 129, Copyright (1990), with permission from Elsevier.

However, prior studies demonstrated that students often inaccurately judge their own comprehension (Griffin et al., 2008; Kiewra, 2005; Maki et al., 1994; Maki & Berry, 1984;

Maki & McGuire, 2009). On the one hand, they might underestimate their current knowledge (i.e., by judging their current understanding lower as their actual performance). This, however, is rather unproblematic as students might spend unnecessary time and effort for studying the provided content (Dentakos et al., 2019) but may nevertheless reach their learning goals. On the other hand, students may also overestimate their current knowledge (i.e., by judging their current understanding higher than their actual performance). This type of inaccurate judgement of learning predominantly occurs among students (e.g., Eitel, 2016; Mayer et al., 2007; Prinz et al., 2020). Especially students with low prior knowledge tend to overestimate their current knowledge. This phenomenon is known as the double curse of incompetence (Dunning et al., 2003; Kruger & Dunning, 1999) as it presents a dual burden: First, students with prior knowledge tend to show poor performance and additionally are not aware of their lack of understanding, thus, show low metacognitive monitoring skills (Prinz et al., 2018).

In summation, deep conceptual knowledge (i.e., transfer knowledge) and metacognitive monitoring (measured by judgement of learning) are crucial components enabling one to acquire deep conceptual knowledge and to achieve high levels in science literacy. However, prior research reveals that students show mainly basic factual knowledge (MINT Nachwuchsbarometer, 2020) and display inaccurate judgements of their own learning (Griffin et al., 2008; Kiewra, 2005; Maki et al., 1994; Maki & Berry, 1984; Maki & McGuire, 2009). Students' inability to build deep conceptual knowledge (i.e., transfer knowledge) and their insufficient metacognitive monitoring abilities conflict with the aim of science education to enable students to achieve high levels in science literacy and highlight the need for additional learning support which engages students in deep learning processes. In this context, *generative learning* is regarded as engaging students in acquiring deep knowledge and to monitor their understanding more accurately.

2.4 Generative Learning Theory

The generative learning theory claims that learning is a generative process since students need to actively connect new knowledge to their prior knowledge through elaborations; this results in new generated knowledge that goes beyond the provided content (Chi, 2009; Fiorella & Mayer, 2016; Renkl, 2008). The generative learning theory is based on and combines several fundamental theories of cognitive processes during learning (cf.

Atkinson & Shiffrin, 1968; Wharton & Kintsch, 1991). In this context, Wittrock's *Generative Model of Learning* is considered as a pioneer work and fundamental theory in the field of generative learning (Figure 4). Based on prior research (see Barlett, 1932; Piaget, 1926), Wittrock developed a model of generative learning which is based on the premise that students need to *actively* build a connection between new information and their prior knowledge to deeply understand the provided content (Wittrock, 1974, 1985, 2010). This process is based on two steps: First, students must generate relations or structures between the components of the to-be-learned content (see Figure 4). This process is comparable with building the propositional network by Wharton and Kintsch (1991). Second, they must connect new information with their prior knowledge by elaborating (e.g., using past experiences, see Wharton & Kintsch, 1991) the content. Through these elaborative processes, students may be able to draw inferences that go beyond the provided learning content (Chi, 2009; Chi & Wylie, 2014).

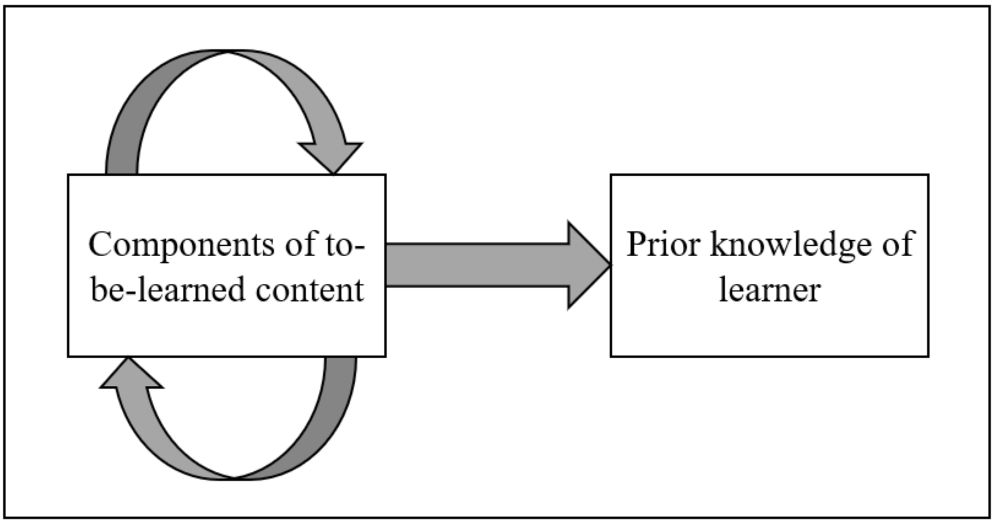

Figure 4. Graphical illustration of Wittrock's Generative Model of Learning (Wittrock, 1974).

Wittrock's perspective on learning, which involves the student as an active part of the learning process, can be regarded as a milestone in research on learning, since prior research regarded learning as an act of memorizing (see Katona, 1940, as an example for theory of memorizing). As a limitation, his model is documented rather superficially since it does not comprise a cognitive perspective on underlying processes during learning. Mayer (1996) presented an extended model in which he included detailed cognitive processes. His *Select-Organize-Integrate (SOI) Model* (see Figure 5; see also Kiewra, 2005, for related model *SOAR*). According to Mayer, meaningful learning (that results in deep knowledge) contains three cognitive processes (i.e., selecting, organizing, and integrating) and takes place in three memory stores (i.e., sensory memory, working memory, and long-term memory). Similarly to Atkinson and Shiffrin (1968), new information is first captured in the sensory memory. Students are then required to select the most relevant information and to mentally organize it in a coherent structure in their working memory. As in Wittrock's and the Wharton-Kintsch model, students then activate their prior knowledge to elaborate the content more deeply, doing so by integration the new information with past experiences. Through this integration, students build a new mental knowledge representation. Following a successful integration, the generated new knowledge is stored in the long-term memory (e.g., in form of new generated schemas or principles).

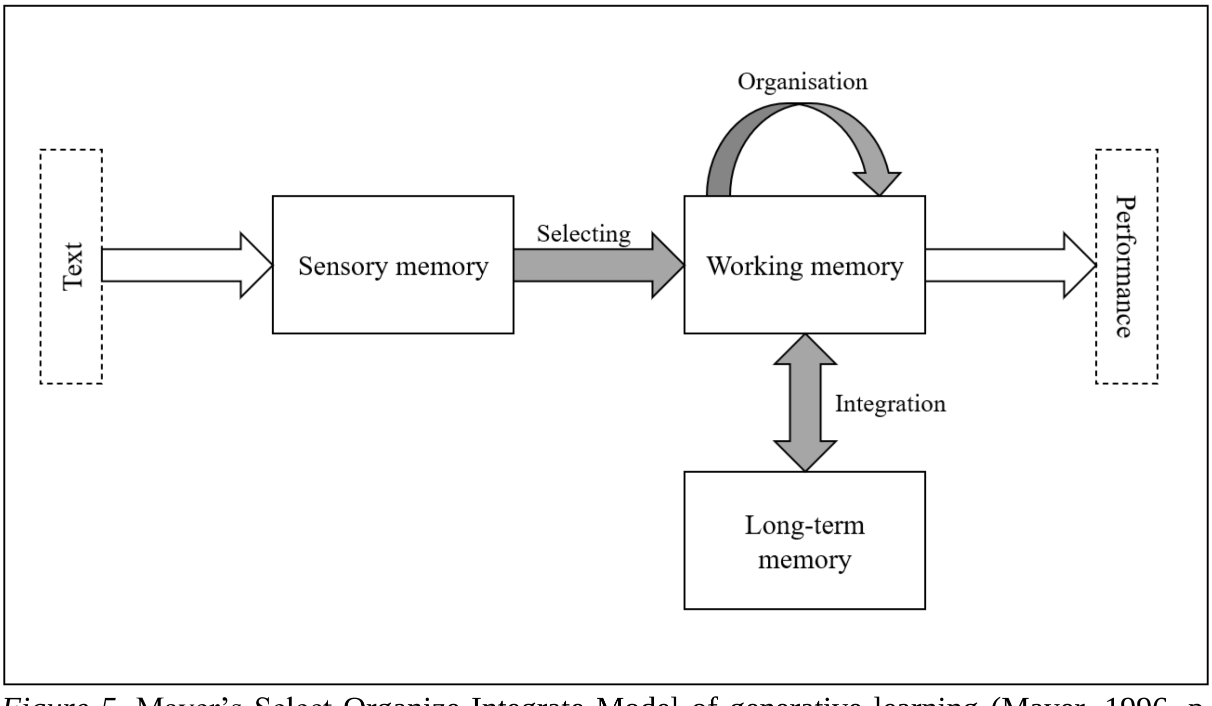

Figure 5. Mayer's Select-Organize-Integrate Model of generative learning (Mayer, 1996, p. 365). Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, *Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction*, Mayer, copyright, 1996.

In sum, elaborating new information by activating prior knowledge allows students to think beyond the provided content which may result in deep conceptual knowledge (i.e., transfer knowledge). In this context, students' prior knowledge plays a crucial role in knowledge acquisition, as it might determine the generative learning process (Mayer, 2014; Wharton & Kintsch, 1991). Students with low prior knowledge, for instance, might have more problems with connecting new information to their prior knowledge since their prior knowledge is limited, and so only limited connections are possible (Brod, 2020; Fiorella & Mayer, 2014; Mayer, 2014; Renkl, 2009).

2.5 Fostering Generative Learning

Prior research has demonstrated that students often fail in engaging in generative learning processes (Brod, 2020). Students' difficulty in engaging in generative learning processes highlights the need of providing adequate instructional learning strategies. Chi (2009) has outlined different learning strategies in her *Interactive-Constructive-Active-Passive (ICAP)* model, which may offer adequate learning strategies to foster generative learning. In her model, she additionally provides a classification of the strategies of students' cognitive engagement during learning. Chi argues that students deepen their understanding and increase their learning outcome the more cognitively engaged they are with the learning content. Chi focused on four different levels of cognitive engagement during learning, namely, *passive learning*, *active learning*, *constructive learning*, and *interactive learning* (see Figure 6). Whereas passive (i.e., using no learning strategy; e.g., listening) and active learning (i.e., lower-order learning strategy that involves a physical action, such as highlighting signal words or summarizing the content) are considered to enhance rather lower-order knowledge (i.e., factual knowledge), constructive and interactive learning is considered to trigger generative learning (Chi & Wylie, 2014).

According to Chi (2009), constructive learning is the process of generating new knowledge. In line with the generative learning theory, students need to elaborate new information (by connecting it with their prior knowledge) to generate new mental knowledge structures which are stored in long-term memory (see also Mayer, 1996, 2014). Constructive learning can be initiated through learning strategies such as self-explaining or taking notes in one's own words (Chi, 2009; Chi & Wylie, 2014). For such learning strategies, students need to restructure and elaborate the new information by activating their prior knowledge to be able to reformulate the content in their own words (Fiorella & Mayer, 2016). This integration of new knowledge to existing knowledge structures and the reformulation of the content in one's own words might enable students to additionally detect missing knowledge gaps (Brod, 2020; Plötzner et al., 1999; van de Pol et al., 2020).

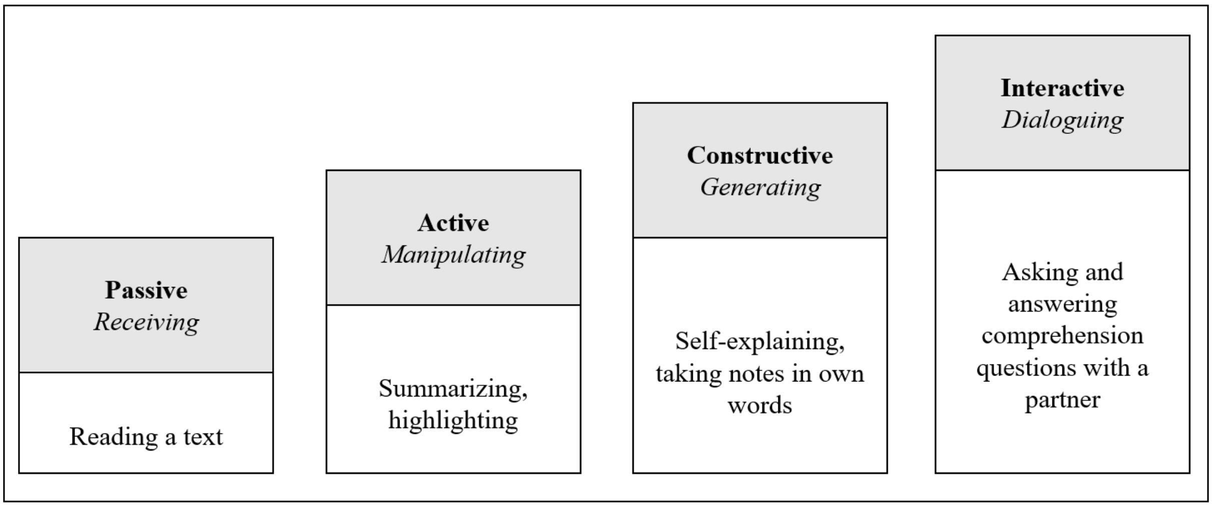

Figure 6. Graphical Illustration of Chi's Interactive-Constructive-Active-Passive Model (Chi, 2009, 2014).

Interactive learning, by contrast, represents learning in settings in which students learn together and react to each other's input (e.g., by providing feedback or asking questions; see Figure 6). Through discussing ideas and perspectives, learners should be able to generate new ideas "that neither individual knew initially nor could generate alone" (Chi & Wylie, 2014, p. 223). These processes can be initiated through learning strategies such as explaining or asking and answering comprehension questions to each other. Based on the ICAP model, interactive learning should result in higher learning outcomes compared to constructive learning since students generate new knowledge together that no individual learner could achieve alone without such interactions (Chi & Wylie, 2014). However, prior research indicated only a subtle benefit of interactive strategies instead of constructive strategies (e.g., Wiggins et al., 2017). This slightly higher benefit raises the question of whether interactive strategies are indeed more beneficial since they are more time and effort consuming to implement, for instance, as homework assignments. Another limitation of the ICAP refers to the strict classification of learning strategies regarding their effectiveness. Ainsworth and Scheiter (2021), for instance, showed that the same learning strategy (i.e., learning by drawing) might be implemented as an active, a constructive, and an interactive method to improve students' learning. Thus, Chi's classification of learning strategies should rather be considered as examples.

In summation, students have difficulties in engaging in generative learning processes (Brod, 2020) which would help them to generate deep conceptual knowledge (Chi, 2009; Fiorella & Mayer, 2016; Mayer, 1996; Renkl, 2008) and to enhance their metacognitive monitoring (Dunlosky et al., 2013; van de Pol et al., 2020). Hence, students need to be provided with learning strategies which help them to engage in generative learning processes. Both constructive and interactive learning strategies (e.g., explaining) are regarded as generative learning strategies (Wylie & Chi, 2014) which may foster generative learning with a goal to enhance both students' deep conceptual knowledge (Chi, 2009; Fiorella & Mayer, 2016; Mayer, 1996; Renkl, 2008) and their monitoring accuracy (Dunlosky et al., 2013; van de Pol et al., 2020).

3

FORMS OF EXPLAINING

3. Forms of Explaining

Explaining is considered to be a learning strategy that may engage students in generative learning processes (Brod, 2020; Fiorella & Mayer, 2016). Explaining is the act of generating an explanation to present learned content in an understandable way (Fiorella & Mayer, 2016). While generating an explanation, students are involved in several cognitive (e.g., activation of prior knowledge) and metacognitive (e.g., metacognitive monitoring) processes which are considered to deepen their understanding. Regarding cognitive processes and in line with Mayer’s SOI model (1996), students need to detect the most relevant information and to organize it in a coherent way for generating an explanation. Additionally, they have to activate their prior knowledge to integrate new information into their existing mental knowledge structures (i.e., elaborative processes; see Brod, 2020; Fiorella & Mayer, 2016; Mayer, 1996). Through this elaborative connection, they might be able to draw inferences from their new generated knowledge to other related domains. Besides cognitive processes, explaining may also engage students in metacognitive processes. For instance, generating an explanation might induce students to reflect on their current understanding more carefully, thus demonstrating whether they have correctly understood all relevant content (i.e., monitoring accuracy, cf. Atkinson & Shiffrin, 1968). This monitoring, triggered by explaining, may enable them to detect knowledge gaps or unclarities which they could solve either through elaborating the misunderstanding by activating more prior knowledge or through restudying the given content (Fukaya, 2013; Lachner et al., 2020). These (meta-)cognitive processes might be challenging for younger students as their cognition is not fully developed; meaning that their prior knowledge and their ability to monitor their current understanding is limited, which may prevent applying generative learning strategies, such as explaining successfully (see Brod, 2020). Thus, students age and their level of cognitive development should be considered when explaining as a generative learning strategy (see Brod, 2020, for more details).

Generating an explanation as a learning strategy may be implemented in different formats which follow distinct pragmatic operations (e.g., Brod, 2020; Fiorella et al., 2017; Fonseca & Chi, 2011; O'Reilly et al., 2004; Plötzner et al., 1999). In this context, Leinhardt (2002) distinguishes two types of explaining that may be implemented in educational settings, namely, *self-explaining* and *instructional explaining*.

3.1 Self-Explaining as a Generative Learning Strategy

Self-explaining is considered to be a "constructive or generative learning activity that facilitates deep and robust learning" (Wylie & Chi, 2014, p. 415); it is generally defined as the process of explaining the meaning of learned content to oneself (McNamara, 2004). Self-explaining may trigger five strategies during learning, namely, *paraphrasing*, *elaborating*, *using logic*, *predicting*, and *bridging inferences* (cf. McNamara, 2017). Paraphrasing represents the act of restating learned content, or preferably reformulating it in own words (McNamara, 2017). Paraphrasing affects learning positively, especially regarding students with low prior knowledge, as it enables the students to develop a better understanding of the provided content (McNamara et al., 2006; O'Reilly & McNamara, 2007a). Elaborating indicates the process of activating related prior knowledge so as to connect new information into existing knowledge structures (McNamara, 2017). Hence, prior knowledge seems to play a crucial role in effective learning. In this context, the next strategy, namely using logic, might serve as an activity, especially to support students with low prior knowledge (McNamara, 2004). Encouraging the use of logic, or in other words general knowledge or common sense (see McNamara, 2017), might enable students to make sense out of content even if it is unfamiliar. Forth, self-explaining may lead students to predict further steps or possible outcomes which help students to think ahead and more generally (McNamara, 2017). Lastly, generating bridging inferences between individual sentences represents an important activity during learning. Bridging inferences are fundamental to organizing and integrating single sentences into a coherent and combined mental representation (cf. Kintsch, 1988).

Besides these cognitive processes that might be initiated by self-explaining, generating an explanation to oneself may also engage students in metacognitive processes. VanLehn et al. (1992) developed the Cascade model in which the authors presented unobservable single steps during explaining examples aimed at supporting students' skills in problem solving (see Figure 7). The authors stated that self-explaining enhances both students' comprehension (i.e., problem solving skills) and their monitoring accuracy through recognizing unclarities during learning.

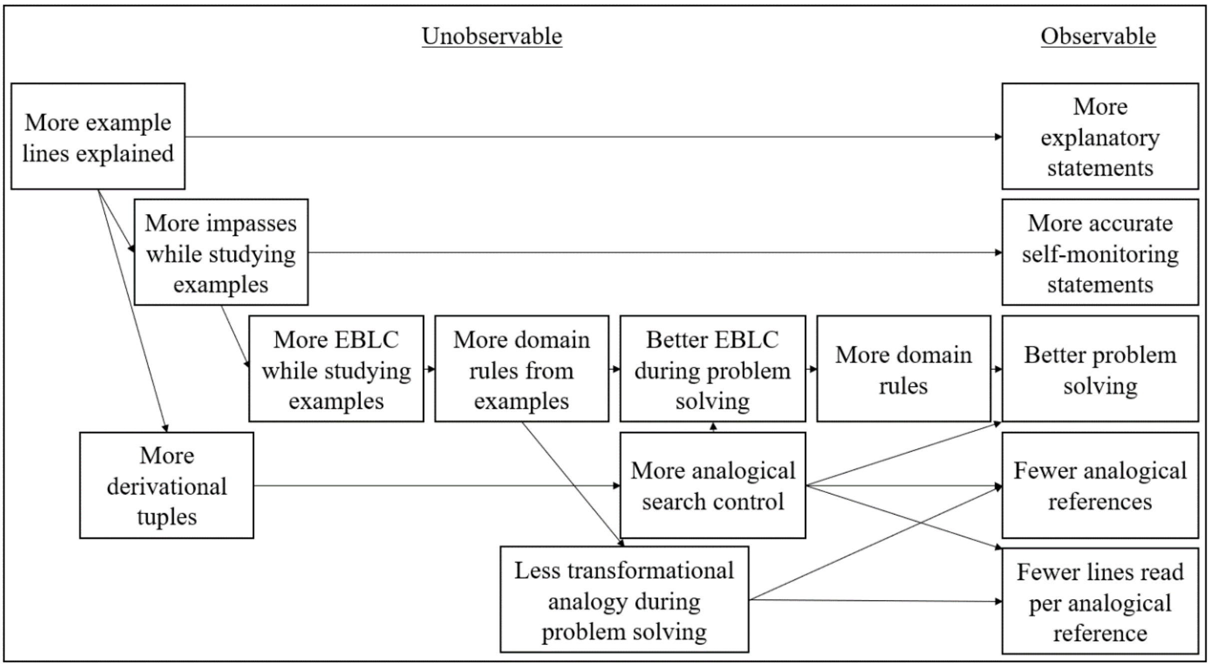

Figure 7. Cascade model of the self-explaining effect by VanLehn, Jones, & Chi, (1992, p. 25). Reprinted from Journal of the Learning Sciences, VanLehn, Jones, & Chi, *A model of the self-explaining effect*, p. 25, Copyright (1992), with permission from Taylor & Francis. The acronym "EBLC" stands for "explanation-based learning of correctness".

In experimental studies, self-explaining as a generative learning strategy is generally implemented as explaining learned content aloud during learning new information, for example, after reading a text (McNamara, 2004, 2017). O'Reilly et al. (2004), for instance, conducted an experiment with nine biology classes (*N* = 136 school students) to investigate the effectiveness of self-explaining compared to another learning strategy, namely previewing. The authors randomly assigned the students to one of three conditions. They either read a chapter of a science textbook and, following each paragraph, explained the content out loud to themselves (self-explanation condition), or they were asked to preview subsections of the text before reading it (preview condition), which is regarded to be a beneficial reading strategy that supports student learning (O'Reilly et al., 2004). A control group simply read the text without being engaged in additional learning strategies. Results showed that students who self-explained the content outperformed not only the control group but also the students in the preview condition regarding their comprehension. Additionally, especially low prior knowledge students seemed to benefit from self-explaining (see McNamara, 2004, 2017; McNamara & Scott, 1999, for similar results). These results are in line with the meta-analysis on self-explaining by Bisra et al. (2018). In total, the authors included 64 studies in which the effectiveness of self-explaining on students' learning was investigated. Results revealed an overall effect of self-explaining (g = .55, medium effect size) which provides empirical evidence for the beneficial effects of this strategy.

3.2 Instructional Explaining as a Generative Learning Strategy

Instructional explanations are designed to "specifically communicate some portion of the subject matter to others" (Leinhardt, 2002, p. 340). Leinhardt defined instructional explanations as part of the teaching process of teachers; however, the process might also be transferred to explaining as a learning strategy (cf. *learning by teaching*, e.g., Roscoe, 2014; Topping, 2005). In line with the generative learning theory, instructing students to explain learned content to others is viewed as engaging them in generative learning (Chi & Wylie, 2014). Prior research commonly investigated explaining in interactive settings; thus, students provide *explanations to each other*. However, researchers started to investigate potential effects of *explaining to fictitious peers* as well.

3.2.1 Explaining to Each Other

Commonly, explaining is implemented as *explaining to each other* in interactive learning situations (e.g., Dillenbourg, 1999; Duran, 2016; McNamara & Scott, 1999; Plötzner et al., 1999; Roscoe, 2014; Topping, 2005; Webb et al., 1995; Webb et al., 2009). Interactive learning settings, meaning that students work and learn together, allow students to respond to each other's input, to learn from each other's ideas, and deepen their own understanding (Chi & Wylie, 2014). For instance, finding supporting facts to defend one's own position requires a profound elaboration of the provided content (Chi, 2009; Chi & Wylie, 2014). Through interacting with each other during explaining, students are able to generate knowledge "that neither individual knew initially nor could generate alone" (Chi & Wylie, 2014, p. 223).

Prior research has mainly focused on investigating the effectiveness of instructional explaining in such interactive settings and shown positive effects on students' comprehension of explaining learned content to a present and active peer. For instance, Annis (1983) conducted a study to investigate the effectiveness of learning by explaining to a peer. In her study, university students (N = 130) were provided with a text and were assigned to one of five experimental conditions. Students in the first condition only read the text (read condition), in the second condition they did not read the text but were taught about its content (taught condition), in the third condition students read the text and additionally were also taught its content (read and taught condition), in the fourth condition they read the text with the expectation to explain the content to a peer afterwards but did not actually explain (explaining expectancy condition), and in the fifth condition students read the text and actually explained the content in an interactive tutoring setting (explaining condition). Results showed that students who explained the content outperformed students who were taught. More importantly, students in the explaining condition also outperformed students who prepared to explain (explaining expectancy condition). These results highlight that the actual act of explaining to others seemed to be the crucial factor that enhanced students' comprehension (see Duran, 2016; Palinscar & Brown, 1984; Plötzner et al., 1999; Roscoe, 2014; Topping, 2005; Webb et al., 1995; Webb et al., 2009, for similar findings).

Kobayashi (2019) conducted a meta-analysis to investigate effects of explaining to each other across published studies. In his analysis, he included 28 experiments that either compared a preparing-to-explain condition with a control condition (28 comparisons) or that compared an explaining-to-each-other condition with a control condition (16 comparisons). Results showed medium positive effects of preparing to explain in contrast to the control groups (g = .56). Regarding explaining to each other, results even revealed large positive effects (g = .84). With his meta-analysis, Kobayashi (2019) provided empirical evidence for a beneficial effect of learning by explaining on students' comprehension across studies.

3.2.2 Explaining to Fictitious Peers

In the last decade, researchers started to raise the questions whether students profit from generating an explanation to fictitious peers as well (Fiorella & Mayer, 2013; Hoogerheide, Renkl, et al., 2019; Lachner et al., 2021). When explaining to fictitious peers, students are asked to imagine a peer during explaining. In this scenario, the audience (e.g., peer) is neither present nor active and, therefore, not able to react to the generated explanation. Thus, students are confronted with a situation in which they do not receive any feedback or further input. Nevertheless, having a partner in mind to whom they provide the explanation might lead them to build a profile of the fictitious peer to be able to generate an adequate explanation (Lachner et al., 2021; Nückles et al., 2006). For instance, they might consider the fictitious peer's prior knowledge which may serve as a baseline in communication to exchange and understand ideas (Clark & Brennan, 1991; Nückles et al., 2006). Adapting learned content to the fictitious peer's needs might lead students to explain the content in more detail, for instance, to provide more examples which could help the fictitious peer to understand the explanation more easily (cf. Clark & Brennan, 1991; Lachner et al., 2021).

Prior research indicated positive effects of learning by explaining to fictitious peers (e.g., Fiorella & Mayer, 2013, 2014; Hoogerheide et al., 2014). For instance, Hoogerheide, Renkl, et al. (2019) conducted an experiment in which university students (N = 61) first studied provided learning materials. Afterwards, they were randomly assigned to one of two conditions: They either generated an explanation to a fictitious peer (explaining condition) or restudied the content (control condition). Results showed that students in the explaining condition outperformed students in the control condition regarding their comprehension. Even though positive effects of explaining were demonstrated in several studies (Fiorella & Mayer, 2013; Hoogerheide et al., 2014; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019), there is a growing counter current of research revealing null or even negative findings of explaining. For instance, together with Lachner and Hoogerheide, I conducted two experiments in which university students either explained provided content to themselves or to a fictitious peer (see Lachner et al., 2021).[1] A control group retrieved the content. Results of both experiments revealed no differences between explaining to a fictitious peer and retrieval practice regarding comprehension (see also Fiorella et al., 2017; Fukaya, 2013; Hoogerheide et al., 2016, for similar findings). These contradictory findings raise the question of whether explaining to a fictitious peer is an effective learning strategy to enhance learning and highlight the need for further research.

[1] This published article is not part of the dissertation studies and, therefore, constitutes an additional article independent of the dissertation.

4

EVIDENCE OF EXPLAINING TO A FICTITIOUS PEER

4. Evidence of Explaining to A Fictitious Peer

Learning by explaining to fictitious peers has risen in importance among the generative learning strategies in recent years. This increase of popularity is mainly due to the advent of digitalization in schools. Digitalization has promoted student-centered learning in the last years as the power of digital media to increase students' cognitive engagement during learning has become recognized (Jonassen, 2005). Therefore, many teachers started to implement student-centered digital learning activities in their classrooms, such as generating a digital explanation. Whether explaining to a fictitious peer is effective, however, is still an open question. Prior research only reported mixed findings of explaining to fictitious others on both students' comprehension (e.g., Hoogerheide et al., 2016; Lachner et al., 2020) and monitoring accuracy (Fukaya, 2013; Lachner et al., 2020).

4.1 Effects on Students' Comprehension

To date, twelve studies, which included a total of twenty experiments, have been published in the last decade in which the effects of explaining to fictitious peers on students' comprehension were investigated (see Table 1 for summarized findings). Reviewing this research revealed that several studies documented positive effects of explaining to fictitious others on factual knowledge. For instance, Fiorella and Mayer (2013) conducted an experiment to investigate whether the expectation to explain or the actual act of explaining promotes learning. They conducted two experiments with three conditions (i.e., explaining condition, expect to explain condition, control condition). First, university students (N = 93) studied a lesson about the doppler effect with the intention to either answer questions about the content afterwards (control condition) or to explain the content afterwards. Those who studied the text with the intention to explain then randomly either indeed explained its content to a fictitious peer (explaining condition) or answered questions as the control group did (expect to explain condition). Results showed that both conditions, the explaining (d = 0.82, large effect) and the expect-to-explain condition (d = 0.59, medium effect), outperformed the control group regarding students' factual knowledge. There was no difference between the expect-to-explain and the explaining condition. The authors replicated their study with a slight change regarding the posttest. This time, students (N = 75) did not answer the posttest immediately after the learning activity, but instead waited one week to measure long term effects. Results revealed that only explaining resulted in higher learning outcomes than restudy (*d* = 0.79, large effect) but not the expectation to explain (*d* = 0.24, small effect). These results provide first evidence that, although the expectation to explain also showed a positive impact on learning, the actual act of explaining is the crucial component of the explaining effect, as it resulted in long-term effects.

The authors supported their findings with an additional experiment with a 2 × 2 design with expectation (to be tested versus to explain) and explaining (explaining versus no explaining) as between-subject factors (Fiorella & Mayer, 2014, Experiment 2). University students (*N* = 104) again studied a text about the doppler effect with the intention to either explain or to be tested afterwards and then randomly either generated an explanation to a fictitious peer or restudied the content. In line with prior findings, results showed that generating an explanation increased students' factual knowledge more than restudying the content (*d* = 0.56, medium effect). Additionally, the authors reported an interaction effect of expectancy and explaining: Students who prepared to explain, and generated an explanation, outperformed the control group (*d* = 0.56, medium effect), however, students with the same explaining expectation who did not explain scored lower than the control group regarding students' factual knowledge (*d* = -.27, small effect).

With their research, Fiorella and Mayer (2013, 2014) highlighted the crucial role of the actual act of explaining to promote learning. These beneficial effects of explaining to fictitious peers are in line with more recent research. For instance, Hoogerheide, Visee, et al. (2019) conducted an experiment in which school students (*N* = 131) were asked to study a text about photosynthesis as often as they wanted (i.e., restudy condition), and then to either summarize the text (i.e., summarizing condition), or generate an explaining video in which they explained the content to a fictitious peer (i.e., explaining condition). Results revealed that only explaining to a fictitious peer but not summarizing resulted in higher learning outcomes compared to the restudy condition. These results provide evidence that students benefit more from explaining compared to restudying and, additionally, more than summarizing the learning content.

These studies provided empirical evidence of the beneficial effects of explaining to fictitious peers. In this context, it is worth noting that the presented studies only measured students' factual knowledge, but not their transfer knowledge. As generative learning strategies, such as explaining, are considered to especially enhance students' transfer knowledge, research on the explanatory effect on students' transfer knowledge are particularly interesting. Reviewing prior research revealed twelve experiments in which students' transfer knowledge was measured in addition to their factual knowledge (see Table 1). For instance, Hoogerheide, Renkl, et al. (2019) recently conducted an experiment to investigate the effectiveness of explaining on students' factual and transfer knowledge with two conditions (i.e., explaining condition, restudying condition). In their experiment, university students (N = 61) studied a text about troubleshooting electrical circuits, and either explained the content to a fictitious peer (explaining condition) or restudied the content for the same amount of time (restudy condition). Afterwards, they answered two posttests: As factual questions, they first had to solve four troubleshooting problems, which were based on the provided learning material. As transfer questions, they then had to solve two problems which contained a different fault. Results showed that students in the explaining condition outperformed students in the control condition regarding both students' factual knowledge (d = 0.54, medium effect) and their transfer knowledge (d = 0.53, medium effect). In line with the generative learning theory (Mayer, 1996, 2009; Wittrock, 2010), these results indicate that explaining indeed seems to be an effective learning strategy to engage students in deep learning processes (see both experiments in Hoogerheide et al., 2014, for similar large effects).

In total, thirteen experiments reported beneficial effects of learning by explaining to fictitious peers (see Table 1). Nine additional experiments showed that explaining was more effective than restudying or retrieving the content regarding students' factual knowledge. Additionally, four experiments indicated positive effects on students' transfer knowledge.

Table 1

Effects of Explaining on Students' Factual and Transfer Knowledge

	Hypotheses			
Experiments	Explaining > Restudy / Retrieval	Explaining > Expect to Explain	Expect to Explain > Restudy / Retrieval	Explaining with support > Explaining
Factual knowledge				
Fiorella & Mayer, 2013, Ex. 1	✓	✗	✓	
Fiorella & Mayer, 2013, Ex. 2	✓	✗	✗	
Fiorella & Mayer, 2014, Ex. 1			✓	
Fiorella & Mayer, 2014, Ex. 2	✓	✓	✗	
Fiorella et al., 2017, Ex. 2	✗			
Fiorella & Kuhlmann, 2020	✗			✓
Fukaya, 2013, Ex. 1	✗	✗	✗	
Fukaya, 2013, Ex. 2	✗	✗	✗	
Hoogerheide et al., 2014, Ex. 1	✗	✗	✗	
Hoogerheide et al., 2014, Ex. 2	✓	✗	✓	
Hoogerheide et al., 2016, Ex. 1	✗	✗	✗	
Hoogerheide et al., 2016, Ex. 2	✓ | ✗**			
Hoogerheide, Visee, et al., 2019	✓			
Hoogerheide, Renkl, et al., 2019	✓			
Koh et al., 2018	✓ | ✗*			
Lachner & Neuburg, 2019				✗
Lachner, et al., 2020, Ex. 1	✗			
Lachner, et al., 2020, Ex. 2	✓ | ✗***			
Lachner et al., 2021, Ex. 1	✗			
Lachner et al., 2021, Ex. 2	✗			
Transfer knowledge				
Fiorella & Kuhlmann, 2020	✓			✓
Fukaya, 2013, Ex. 2	✗	✗	✗	
Hoogerheide et al., 2014, Ex.1	✓	✓	✗	
Hoogerheide et al., 2014, Ex. 2	✓	✗	✗	
Hoogerheide et al., 2016, Ex. 1	✗	✗	✗	
Hoogerheide et al., 2016, Ex. 2	✗			
Hoogerheide, Renkl, et al., 2019	✓			
Lachner & Neuburg, 2019				✓
Lachner, et al., 2020, Ex. 1	✗			

Lachner, et al., 2020, Ex. 2	✗
Lachner et al., 2021, Ex. 1	✗
Lachner et al., 2021, Ex. 2	✗

Note. Checkmarks (✓) symbolize significant results in favor of the corresponding hypothesis. Crosses (✗) symbolize results against the hypothesis. Regarding explaining conditions with additional support, students were either provided with concept maps (Lachner & Neuburg, 2019) or with an additional drawing activity (Fiorella & Kuhlmann, 2019). *Explaining resulted in higher learning outcomes when students explain without provided notes but did not result in higher learning outcomes when students explain with provided notes. **Explaining resulted in higher learning outcomes when students explained orally but not in written form. ***Explaining was only significant when students explained in-between but not after passages.

Even though these findings provide empirical evidence for the effectiveness of explaining to fictitious peers, there are several studies which reported contradictory results and, therefore, limit the promising effects of explaining to fictitious peers. Fiorella et al. (2017), for instance, conducted an experiment with a 2 × 2 design with the form of an instructional video (i.e., first-person perspective versus third-person perspective) and explaining (i.e., explaining versus no explaining) as between-subject factors. In their study, university students (N = 121) first watched a video either with first-person perspective or with third-person perspective about electric circuit. Afterwards, they randomly explained how to build an electric circuit (explaining condition), or they built it for themselves without an explaining activity (no explaining condition). In contrast to the previously mentioned studies, results showed no effect of explaining among conditions regarding students' factual knowledge (see also Fukaya, 2013; Hoogerheide et al., 2016; Lachner et al., 2020; Lachner et al., 2021, for further null findings). In a similar study, Hoogerheide et al., (2016) investigated whether the intention to explain or the actual act of explaining had an effect on students' factual and transfer knowledge. The authors conducted an experiment with a 2 × 2 design with intention (i.e., to explain versus to be tested) and learning activity (i.e., explaining versus restudying) as between-subject factors. University students (N = 123) read a learning text with the intention to either explain the content to a fictitious peer afterwards or to be tested. Then, they randomly indeed generated an explanation or restudied the content. Results revealed no effects across conditions regarding both students' factual and transfer knowledge. Thus, neither the intention to explain nor the actual act of explaining seemed to enhance students' learning. These conflicting results demonstrate that the effectiveness of explaining to fictitious others might depend on further boundary conditions which have not been examined yet.

In summation, nineteen experiments reported null or negative findings of explaining to fictitious peers on students' comprehension compared to restudying or retrieving the learning content (see Table 1). Regarding students' factual knowledge, twelve experiments had to reject the hypothesis of a benefit of explaining. Additionally, seven experiments reported null or negative findings regarding students' transfer knowledge. Thus, reviewing prior research revealed that prior experiments resulted in equally positive and negative findings of learning by explaining to fictitious peers, which highlights the unclear pattern regarding the effectiveness of explaining.

4.2 Effects on Students' Monitoring Accuracy

Along with students' comprehension, prior research highlighted the crucial role of monitoring accuracy (Fiorella & Mayer, 2016; Fukaya, 2013; Schleinschok et al., 2017), as it helps students to learn new content (Anderson & Thiede, 2008; Schleinschok et al., 2017; Thiede et al., 2003a, 2003b). However, students often fail to monitor their current understanding correctly (Griffin et al., 2008; Kiewra, 2005; Maki et al., 1994; Maki & Berry, 1984; Maki & McGuire, 2009). Generative learning strategies, such as explaining, are regarded to enhance students' monitoring accuracy (Brod, 2020; Fiorella & Mayer, 2016). However, empirical evidence regarding the effectiveness of generating an explanation to fictitious peers on students' monitoring accuracy is strongly limited. In total, only two researchers investigated the effects of explaining on students' monitoring accuracy.

First, Fukaya (2013) conducted two experiments to address this issue. In his first experiment, university students (N = 39) studied five texts about daily devices with the intention to explain the content afterwards. Then, they randomly indeed generated an explanation (i.e., explaining condition) or proceeded to the knowledge test (i.e., expect to explain condition). A control group studied the text with the intention to generate keywords after studying the text and then actually wrote keywords about the content. Afterwards, all students judged their current comprehension for each text and answered a comprehension test. Fukaya (2013) determined students' monitoring accuracy as the correspondence between students' judgement of their own current comprehension and their actual performance on a comprehension test. Results revealed a main effect on monitoring accuracy across conditions (η^2 = .17, large effect). Students in the explaining condition judged their comprehension more accurately than both students in the control group and students in the expect-to-explain condition. Fukaya (2013) replicated these findings in a second experiment in which university students (*N* = 48) again either only had the expectation to explain afterwards or actually explained the learned content to a fictitious person. This time, Fukaya (2013) implemented a control group in which students were not engaged in any activity, and thus, did not generate keywords. Students again judged their comprehension after the study phase and answered a comprehension test. In line with the prior findings, results showed that students who explained judged their performance more accurately than both students who only had the expectation to explain and students who were in the control group (η^2 = .27, large effect). These findings provided the first empirical evidence that explaining indeed promotes students' monitoring accuracy and, additionally, is more beneficial than studying with the expectation to explain but without generating an explanation. Nevertheless, the results should be interpreted with caution as the sample sizes of both studies were small and, therefore, do not meet the requirements of sufficient statistical power.

In a similar study, Lachner et al. (2020) investigated the effect of explaining on students' comprehension and monitoring accuracy. The authors conducted two experiments to test whether explaining in-between paragraphs is more beneficial than explaining after all paragraphs. In their first study, university students (*N* = 91) read a text about combustion engines and explained the content either in-between the text phases or after all text phases. A control group recalled the content after reading the text. Similar to Fukaya (2013), all students then judged their comprehension and answered a comprehension test. Results revealed no significant differences among conditions (*d* = 0.14, small effect). Thus, in contrast to Fukaya (2013), explaining did not result in more accurate judgements. As a replication study, Lachner et al. (2020) conducted a second experiment (*N* = 126) with a 2 × 2 design with timing of explaining (i.e., in-between versus after the text phases) and learning activity (i.e., explaining versus retrieval practice) as between-subject factors. Again, students judged their comprehension comparably accurate across conditions (*d* = 0.18, small effect).

These contradicting results reveal the lack of research regarding the effect of explaining on students' monitoring accuracy. Further research is needed to solve this unclear pattern and to investigate the effects on monitoring accuracy in more detail, including the investigation of boundary conditions and providing well-designed studies with sufficient power to be able to draw well-founded conclusions.

5

TEST OF UNDERYLING MECHNISMS

5. Test of Underlying Mechanisms

Prior research showed mixed findings regarding the effectiveness of explaining. Nevertheless, several studies indicated beneficial effects of explaining to fictitious peers on students' comprehension and monitoring accuracy. However, little is known about the underlying mechanism which drives the explaining effect. Reviewing prior research revealed that different hypotheses were discussed which might explain the mechanism of explaining, namely, *retrieval practice hypothesis*, *generative learning hypothesis*, and *social presence hypothesis* (see Lachner et al., 2021; Fiorella & Mayer, 2016; Hoogerheide, Visee, et al., 2019; Lachner et al., 2020, for related approaches). These hypotheses are not mutually exclusive but rather build on each other. Additionally, they provide rather theoretical and explorative assumptions on potential underlying mechanisms of explaining since empirical evidence is rare. Therefore, it is essential to systematically test these hypotheses in more detail in future experiments.

5.1 Retrieval Practice Hypothesis

The most basic hypothesis about the underlying mechanism of explaining is the *retrieval practice hypothesis*. According to this hypothesis, the process of explaining results in higher learning outcomes as it is a form of *retrieval*: students are forced to spend time on retrieving the learned content during explaining (Koh et al., 2018; Lachner et al., 2021). Retrieval practice is a mnemonic technique to maintain "access to knowledge in memory" (Bjork, 1988). During retrieving, students are forced to "reactivate and operate on memory traces either by elaborating mnemonic representations or by creating multiple retrieval routes to them" (Roediger & Karpicke, 2006a, p. 197). Retrieval can be initiated in different forms. For instance, within an experimental study, students first might be asked to study new content in the learning phase. Afterwards, they may be provided with a test which students answer without seeing the leaning materials. To answer the test correctly, they need to reactivate the learned content. Retrieval tests can be provided either as recognition tests (e.g., multiple choice) or as recall tests (e.g., open-ended questions). Research demonstrated that recall tests resulted in higher learning outcomes (especially in delayed tests) than recognition tests as they "require greater retrieval effort or depth of processing" (Roediger & Karpicke, 2006a, p. 198). These results can be regarded as evidence that active retrieval is a successful learning strategy to memorize new content (Roediger & Karpicke, 2006a). Another form of recall testing is to ask students to retrieve the learned content in mind (i.e., free recall test). In their comprehensive review, Roediger and Karpicke (2006a) analyzed research from experimental and educational psychology on the testing effect. They highlighted that retrieving new learned content through a (free recall) test resulted in higher learning outcomes than to restudying the materials (Roediger & Karpicke, 2006a). These results are promising since students who retrieved the material spent the same amount of time with the material than students who restudied it. Thus, the retrieval effect is not a result of duration, but of the active process or reactivating learned content.

Regarding learning by explaining, however, little is known about whether it is the act of retrieval which drives the explaining effect or whether there are further underlying processes. This research deficit might occur because prior research did not test retrieval processes during explaining. Some studies, however, contrasted an explaining condition to a retrieval condition to investigate whether explaining is more beneficial than retrieving (see Koh et al., 2018; Lachner et al., 2020; Lachner et al., 2021). For instance, Koh et al. (2018) conducted a study with four conditions (i.e., explaining with retrieval, explaining without retrieval, retrieval practice, control condition). In total, 124 students read a text. Afterwards, students in the explaining conditions explained the content to a fictitious other, either without seeing the materials anymore (i.e., explaining and retrieving condition), or with seeing the learning material (i.e., explaining without retrieving condition). Students in the retrieval practice condition only retrieved the content in form of a free recall test (i.e., retrieval condition), and students in the control condition neither explained nor retrieved the content but solved arithmetic problems instead. After one week, students answered the comprehension posttest. Results showed a main effect of the learning strategies ($\eta_p^2 = .13$, large effect). In more detail, students in the explaining with retrieval and in the retrieval practice condition outperformed the control condition ($d = 0.84$, large effect). More importantly, students who explained with retrieval and students who were engaged in retrieval practice outperformed students who explained learned content without retrieval regarding their comprehension ($d = 0.58$, medium effect). There was no difference between students who explained without the material (i.e., explaining and retrieving condition) and students who retrieved the content (i.e., retrieving practice condition). These results are in line with other studies that showed that students who explained did not outperform students who retrieved the content (Lachner et al., 2020; Lachner et al., 2021).

In contrast, Lachner et al. (2020, Experiment 2) conducted an experiment which contradicts the findings by Koh et al. (2018). The authors aimed at investigating whether explaining to a fictitious peer in-between study phases was more beneficial than explaining after the study phases. University students (N = 126) read a text and then explained either in-between paragraphs or at the end of the text. Additionally, the authors implemented a control group in which students retrieved the content in form of a free recall test. After this learning activity, students answered factual and transfer questions as posttests. Results showed that in-between explaining was more beneficial than explaining after the study phases. More importantly both explaining conditions resulted in higher learning outcomes than the retrieval condition regarding students' factual knowledge (d = 0.45, medium effect). There was no difference among groups regarding transfer knowledge. Thus, in contrast to Koh et al. (2018) explaining resulted in higher learning outcomes than retrieval practice.

In summation, in both studies, an explaining condition was contrasted to a retrieval practice condition and results provided evidence for and against the effectiveness of explaining in contrast to retrieval practice. However, the authors did not directly test retrieval as an underlying mechanism of explaining. To systematically test the *retrieval practice hypothesis*, underlying mechanisms during the learning activity need to be analyzed. In this context, students' generated explanations might reveal interesting information about the underlying mechanism of explaining. As an indicator for retrieval processes, for instance, students' mentioned concepts within their explanations might be calculated. Concepts are of lexical nature and contain both the word itself and its meaning in form of a mental representation. The concept "immunology", for instance, contains the knowledge about the word itself (i.e., study of immune systems in all organisms) but also the knowledge about its risks and applications (e.g., vaccination). Retrieving provided content and its containing concepts (e.g., in form of a free recall test) should help students to consolidate the concepts in their long-term memory. Thus, successful retrieval practice should result in high numbers of mentioned concepts. Prior research, however, has not provided details about indicators for students' active retrieval process during explaining (e.g., in form of concepts). Therefore, further studies are needed that continue to investigate the retrieval practice effect as a potential underlying mechanism of learning by explaining, for instance, by systematically analyzing students' generated explanations. In this context, the number of concepts might be an indicative factor for students' retrieval processes during explaining.

5.2 Generative Learning Hypothesis

The most prominent hypothesis about the underlying process of explaining is presumably the *generative learning hypothesis*. This hypothesis is in line with the generative learning theory (Mayer, 1996; Wittrock, 1974) and therefore also presents the same theoretical assumptions: Explaining is regarded as a learning strategy that triggers students to engage in generating new knowledge which results in a deeper level of comprehension than restudying or retrieving the learned content (Chi & Ohlsson, 2005; Wittrock, 2010). The *generative learning hypothesis* does not contradict the *retrieval practice hypothesis*, as retrieval might still be a part of the explanatory process. It rather extends the previous hypothesis, as the *generative learning hypothesis* states that students reach meaningful transfer knowledge as they are engaged in elaborative processes during explaining. Prior research mainly investigated such elaborative processes through counting the number of elaborations which occurred in students' explanations. Elaborations are regarded as "inferences that go beyond the text, excluding monitoring statements, paraphrases, comprehension" (Chi et al., 1994, p. 455). For elaborating processes, students have to connect the provided information with their prior knowledge (Chi, 2009; Chi & Wylie, 2014; Lachner et al., 2018). Thus, also personal experiences and analogies, which are mentioned within an explanation are elaborations (Fiorella & Kuhlmann, 2020; Lachner et al., 2018).

Empirical research on elaborative processes as the underlying mechanism of explaining to fictitious others is limited. In a related research area, namely self-explaining, Chi et al. (1994) provided evidence for the positive impact of elaborations. The authors conducted a study in which university students (N = 24) read a text and either self-explained the content after each paragraph (explaining condition) or restudied the text (control condition). Results revealed that students in the explaining condition not only outperformed students in the control condition regarding the knowledge test, but also made more inferences and integrated more information in their answers. Additionally, students who elaborated more within their explanations showed higher comprehension than students who made less inferences (see Chi et al., 1989; Roscoe & Chi, 2008; Webb, 1989; Webb et al., 2009, for similar research and findings). Whether these effects hold true for explaining to fictitious others, however, is still an open research question.

As an exception, Lachner et al. (2018) provided the first evidence of a mediating effect of elaborating on the effectiveness of explaining to fictitious peers on students' comprehension. In their experiment, university students (*N* = 48) read a text about the four-stroke engine and then either explained the content to a fictitious peer in oral versus in written form. Results revealed that students who explained orally outperformed students who wrote an explanation regarding their transfer knowledge (*d* = 0.67, large effect). This explaining effect was mediated by the number of elaborations within students' generated explanations (κ^2 = .21, medium effect). Students who explained orally elaborated more compared to students who wrote an explanation, which resulted in higher learning outcomes regarding students' transfer knowledge. These findings provide first direct evidence for the *generative learning hypothesis* as elaborations were the underlying mechanism of – at least oral – explaining. However, as empirical evidence is limited to this study, further research is essential to provide more evidence for the *generative learning hypothesis*.

5.3 Social Presence Hypothesis

Another assumption about the underlying mechanism of explaining is the *social presence hypothesis* (Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021). Social presence is defined as the "degree of salience of the other person in the interaction" (Short et al., 1976, p. 65). The concept of social presence not only refers to real interactive settings, such as in learning contexts in which students explain to each other, but also to settings with mediated conversations, such as in online learning settings (Atkinson, 2002; Kim, 2013; Wang & Antonenko, 2017). Prior studies demonstrated that perceived social presence of others enhanced important motivational and cognitive concepts of learning in mediated settings and is therefore considered as a crucial factor of learning (Atkinson, 2002; Kim, 2013; Wang & Antonenko, 2017; Weidlich & Bastiaens, 2017, 2019). For instance, Richardson et al. (2017) conducted a meta-analysis to investigate the correlation between social presence in mediated settings and students' perceived learning. The authors included 25 empirical studies in their analysis. As expected, results showed a strong correlation between perceived social presence and students' perceived learning. Thus, it might be that higher levels of perceived social presence enhanced students' actual learning.

In this context, researchers started transferring these findings to contexts of learning by explaining and began to investigate students' perceived social presence during explaining to fictitious others. Hoogerheide et al. (2016, Experiment 2) provided evidence that students perceived the social presence even of a fictitious peer during explaining. In their study, university students (N = 129) read a text which they either explained to a fictitious peer in oral or written form or restudied the content. Results showed that explaining in oral form (d = 0.43, small effect) but not in written form (d = 0.19, small effect) outperformed students who restudied the content. The authors additionally estimated students' perceived social presence of the fictitious peer during explaining by calculating their mentioned personal references. Personal references are personal pronouns that directly address the receiver by mentioning second pronouns (i.e., "you", "your", "yours") or that includes the student him- or herself with first pronouns (i.e., "I", "me", "mine"). Prior research revealed that personal reference within produced language are an indicator of perceived social presence (Akinnaso, 1985; Chafe, 1982; Chafe & Tannen, 1987; Einhorn, 1978). Thus, determining students' mentioned personal references within their explanations might indicate their level of perceived social presence of the fictitious peer. In the study by Hoogerheide et al. (2016, Experiment 2) results revealed that students who explained orally mentioned more personal references in their explanation than those who explained in written form (d = 1.54, large effect). As oral explaining was more beneficial than writing explanations, the authors argued that this modality effect was a result of different levels of perceived social presence during explaining. Following this assumption, students who perceived higher levels of social presence might show higher learning gains compared to students who perceived lower levels of social presence. Why perceived social presence results in higher learning outcomes, however, is still unclear. Reviewing prior research revealed mainly two different perspectives on the underlying mechanism of social presence, namely, the *audience design perspective* and the *motivational perspective*. Both perspectives are not mutually exclusive but may rather complement each other.

From the *audience design perspective*, the benefit of social presence is a result of adaptive processes during explaining (Hoogerheide et al., 2016; Lachner et al., 2021). In line with the audience design theory, it claims that students adjust their explanations to the audience's needs (Clark & Schaefer, 1993; Fussell & Krauss, 1992). In other words, students try to consider the audience's background, such as prior knowledge, to adjust their provided explanations to the audience's level of comprehension (Nickerson, 1999). Having the audience's prior knowledge in mind helps students to adapt their knowledge to the audience's needs (Nückles et al., 2006). These adaptive processes require students to provide further details and examples that go beyond the learned content to ensure that the audience is able to follow their thoughts and understand the information as the students intended (Clark &

Brennan, 1991). For providing information that goes beyond the given learning content, students need to deeply elaborate its content which is regarded to result in meaningful learning (Chi, 2009; Chi & Wylie, 2014; Mayer, 1996; Wittrock, 2010). Thus, having a partner in mind during explaining may trigger students to engage in adaptive and elaborative processes during explaining which might positively affect learning.

From the *motivational perspective*, perceiving high levels of social presence during explaining triggers students to be more motivated to complete the task (cf. social agency theory, Bandura, 1989) which might enhance students' learning (Hoogerheide, Visee, et al., 2019; Lachner et al., 2021). Motivation is defined as a "temporal and dynamic state" which represents the "desire and willingness" of a student to complete a task (Brown & Acosta, 2007, p. vii). Motivation is a multidimensional construct that included further constructs, such as enjoyment and interest (Neighbors et al., 2007). Thus, from a *motivational perspective*, the act of explaining to another person is considered as more enjoyable, interesting, and motivating compared to other learning strategies such as summarizing (Hoogerheide, Visee, et al., 2019). This assumption is in line with prior research that showed beneficial effects of motivation regarding students' learning (Deci & Ryan, 2000; Eccles & Wigfield, 2002; Hidi & Harackiewicz, 2000; Lazowski & Hulleman, 2016; Orhan Özen, 2017). Higher levels of motivation (e.g., enjoyment) might motivate students to spend more time and effort in explaining which might result in more comprehensive explanations, for instance, explaining more concepts (see also Deci & Ryan, 2000). Thus, having a partner in mind may motivate students, for instance, to generate more comprehensive explanations that might be indicated by more concepts.

Empirical evidence for the *social presence hypothesis* as the underlying mechanism of explaining to fictitious peers is rare. Hoogerheide, Visee, et al. (2019) provided first evidence by conducting a study in which school students (N = 131) read a text and either explained or summarized the content. In a control condition, students restudied the learning content. Results showed that explaining the content to a fictitious peer was more beneficial than restudying the content regarding students' comprehension. Interestingly, this effect was mediated by students' level of enjoyment. Students perceived explaining as more enjoyable than summarizing and restudying. Higher levels of enjoyment, in turn, resulted in higher learning outcomes. The authors argued, in line with the *social presence hypothesis*, that perceived social presence of the audience motivated students to spend more time and effort for generating an explanation (see also Deci & Ryan, 2000). However, as perceived social presence was not measured (e.g., in form of personal references), the authors' assumption was speculative and needs further investigations for empirical evidence.

6

BOUNDARY CONDITIONS OF EXPLAINING TO A FICTITIOUS PEER

6. Boundary Conditions of Explaining to a Fictitious Peer

Explaining learned content to fictitious peers mainly resulted in mixed findings regarding students' comprehension (i.e., factual knowledge, transfer knowledge; e.g., Fiorella & Mayer, 2013, 2014; Fukaya, 2013; Hoogerheide et al., 2014; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019; Lachner et al., 2020). In this context, meta-analyses on learning by explaining to fictitious peers revealed a large variance among studies (Kobayashi, 2019; Lachner et al., 2021). These results highlighted that learning by explaining might not be effective in general but rather depends on further boundary conditions. To systematically investigate potential boundary conditions that might moderate the explanatory effect, I reviewed prior research to determine potential boundary conditions. Five potential conditions were observed that were discussed in prior research, namely, *explanatory modality*, *text complexity*, *perceived social presence* during explaining, and students' *prior knowledge* and their *academic self-concept*.

6.1 Does Explanatory Modality Moderate Learning by Explaining?

Although several studies demonstrated that generating an explanation to a fictious peer enhanced students' comprehension (e.g., Fiorella & Mayer, 2013, 2014; Hoogerheide et al., 2014; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019), it has to be noted that the majority of prior studies only investigated the effectiveness of *oral* explaining (see Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2020; Lachner et al., 2021, for exceptions). Oral explaining, however, might significantly differ from implementing an explaining task in written form. Evidence for this assumption can be found in applied linguistic research which highlights that producing oral language is a different process compared to producing language in written form (Akinnaso, 1985; Chafe, 1982; Chafe & Tannen, 1987; Sindoni, 2013; Tannen, 1982, 1983). Producing oral language (i.e., speaking) is a spontaneous and automatic process (Chafe, 1981, 1982; Chafe & Tannen, 1987; Lakoff, 1982; Liberman, 1994; Sindoni, 2013). Speakers are normally not aware of their discourse before they have uttered it (Sindoni, 2013). As a result, this process yields a less coherent and less organized language outcome compared to writing (Chafe, 1981; Lakoff, 1982; Sindoni, 2013). Although producing oral language results in less coherent and organized outcomes, it may trigger students to find knowledge gaps and missing connections in their non-coherent explanations which may enhance elaborative processes (Bourdin & Fayol, 2002; Chafe & Tannen, 1987). In contrast, writing is a forethought and planned process in which writers rather focus on the structure and organization of their language (Chafe, 1981; Lakoff, 1982; Sindoni, 2013). Due to the externalization of thoughts into a writing system, the writing process is regarded as mentally effortful which has been shown in empirical studies (e.g., Bourdin & Fayol, 1994, 2002; Liberman, 1994). Another essential difference between producing oral versus written language is the feeling of an audience's presence during the activity. Research showed that speaking triggered higher levels of social involvement as compared to writing, expressed by more personal references in their language (i.e., first-person and second-person pronouns, see Akinnaso, 1985; Chafe, 1982; Chafe & Tannen, 1987; Einhorn, 1978). These differences might have an impact on the effectiveness of learning by explaining. Therefore, the explanatory modality can be a crucial boundary condition which has rarely been investigated in prior research.

So far, only two studies focused on investigating the effectiveness of writing an explanation to a fictitious peer on students' comprehension. First, Hoogerheide et al. (2016) conducted a study in which they investigated whether students' expectancy to explain or the actual act of explaining determine the effectiveness of explaining (similar to Fiorella & Mayer, 2014, see Chapter 4). However, in contrast to prior research, Hoogerheide et al. (2016) asked students to write an explanation instead of generating an oral explanation. The authors conducted an experiment with a 2 × 2 design with students' expectation (i.e., expect to be tested versus expect to explain) and learning activity (restudying versus writing an explanation) as between-subject factors. University students (N = 123) answered a prior knowledge test and read an instructional text about syllogistic reasoning with the expectancy to either answer questions or to explain the content afterwards. After this study phase, students were randomly assigned to one of two learning activities: they either restudied the provided content or they wrote an explanation to fictitious peers. All students had eight minutes for the learning activity and rated their mental effort after completing this task. Finally, they answered the immediate posttest, which contained factual questions and transfer questions. One week later, the students were asked to complete the delayed posttest which was a parallel version of the first immediate posttest. In contrast to prior research (see Fiorella & Mayer, 2014), the authors reported no differences among conditions: neither regarding learning expectation nor regarding learning activity. However, students in the explaining condition reported that they invested more mental effort during writing an explanation compared to students who restudied the learning content (*d* = 0.46, medium effect). This increased mental effort, however, did not positively affect students' comprehension. Thus, Hoogerheide et al. (2016) could not replicate prior findings, as there was no effect of generating an explanation to a fictitious peer. These mixed findings might be due to the fact that the authors asked students to generate an explanation in written form and not in oral form as in prior research. To solve that open research question regarding potential differential effects of the explaining modality, Hoogerheide et al. (2016) directly compared the effectiveness of oral explaining to writing an explanation on students' comprehension in a second study. Again, university students (*N* = 129) first answered a prior knowledge test and studied the provided learning content. Then, they generated an explanation either in oral or in written form. A control group restudied the content. Afterwards, they rated their mental effort for completing the learning activity and answered a posttest (i.e., factual knowledge, transfer knowledge). Regarding students' factual knowledge, results again showed no difference between students who restudied the content and students who wrote an explanation. However, and in line with prior research on oral explaining, the authors reported that students who explained orally outperformed students who restudied the content (*d* = 0.43, medium effect). Results indicated no differences between generating an oral versus a written explanation. Regarding the transfer posttest, results revealed no differences among conditions. Additionally, students who explained orally (*d* = 1.96, large effect) and in written form (*d* = 0.93, large effect) reported that they invested higher levels of mental effort during explaining than students who restudied the content. These results provided first evidence that only explaining in oral but not in written form is more beneficial than restudying the content.

Lachner et al. (2018) replicated this study by conducting a similar experiment. University students (*N* = 48) were randomly assigned to one of two conditions: they explained learned content either orally or in written form to a fictitious peer. Similar to the procedure by Hoogerheide et al. (2016), students first answered a prior knowledge test and then were provided with a learning text about combustion engines. In contrast to Hoogerheide et al. (2016), this text was more complex as it included more specific terms and a complex test structure. Then, students randomly generated an explanation in either oral or in written form to a fictitious peer. Finally, they answered a posttest (i.e., factual questions and transfer questions). Regarding students' factual knowledge, results indicated no differences among conditions. Regarding students' transfer knowledge, results revealed that students who explained orally outperformed students who wrote an explanation (d = 0.67, large effect). Although these results must be interpreted with caution as the sample size was rather small and, therefore, statistical power was low, they may indicate that oral explaining is indeed more beneficial than writing explanations.

The results are in line with the findings by Hoogerheide et al. (2016) which demonstrated that oral explaining was more beneficial than restudying the content. However, they conflict with the results by Hoogerheide et al. (2016), as Lachner et al. (2018) additionally found a significant difference between oral explaining and writing an explanation. These findings might seem contradictory at first glance, however, analyzing both studies in more detail reveal a difference in the complexity of the used learning material. In contrast to Hoogerheide et al. (2016), Lachner et al. (2018) used rather complex learning material. Thus, these conflicting results may indicate that explaining orally versus in written form does not only depend on the act of explaining itself but on further conditions, such as the complexity of the learning materials, which have not been systematically investigated so far.

6.2 Does Linguistic Complexity of the Learning Material Moderate Learning by Explaining?

Since prior research about explaining in written and in oral form to fictitious peers resulted in mixed findings (see Hoogerheide et al., 2016; Lachner et al., 2018), it is essential to investigate further boundary conditions to solve this unclear pattern. Reviewing prior research revealed that the given learning material differed regarding its complexity among studies. Hoogerheide et al. (2016) used an instructional text with concrete terms which was rather easy to process, whereas Lachner et al. (2018) provided students with a complex Wikipedia article which contained many specific terms and a complex sentence structure. Different levels of complexity of the learning material might strongly impact the effectiveness of learning strategies, such as explaining (e.g., Lachner & Nückles, 2015; McNamara et al., 1996; McNamara & Knitsch, 2009). According to McNamara (2013), text complexity is a multi-dimensional construct which contains different text features (see Figure 8), namely, genre, syntax, word concreteness, referential cohesion, and deep cohesion. As the genre did not differ in prior studies (instructional texts in STEM fields), I focus on the remaining constructs, namely, syntactic complexity, word concreteness, and the cohesion of a text in the following.

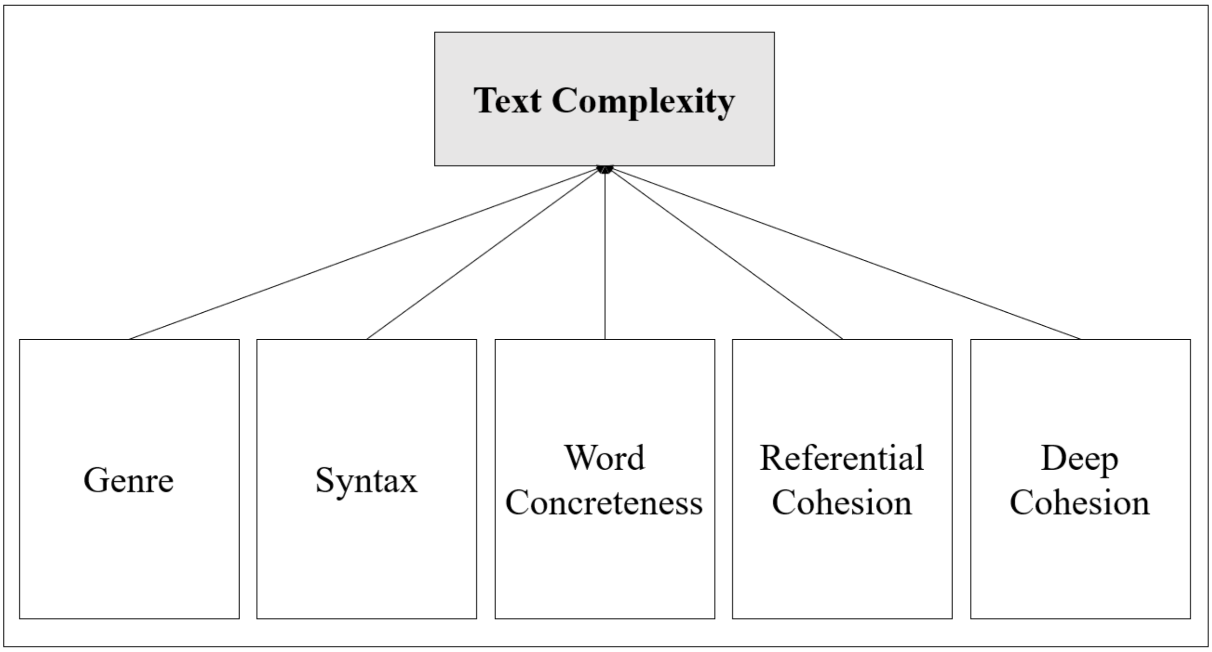

Figure 8. Model of text complexity based on McNamara (2013). Reprinted by permission from SAGE Publications: Discourse Studies, *The epistemic stance between the author and reader: A driving force in the cohesion of text and writing*, McNamara, copyright, 2013.

Syntactic complexity is regarded as the component which "reflects the degree to which the sentences in the text contain […] more words and use complex syntactic structures" (McNamara, 2013, p. 7). For instance, long sentences caused by subordinate clauses but also the use of many nominals increase the complexity of the syntax (McNamara, 2013). The higher syntactic complexity, the more challenging it is to process the provided text; thus, the reader has to increase his or her mental effort to understand the content (Berendes et al., 2018; McNamara, 2013). In contrast, word concreteness helps readers to understand the text more easily. A text is regarded as being word concrete when it contains "words that are concrete, meaningful, and evoke mental images" (McNamara, 2013, p. 7). In reverse, abstract words prevent students to understand the content quickly and forces them to invest more mental effort, which might reduce their cognitive capacity (McNamara, 2013). Lastly, text cohesion strongly determines the complexity of texts. In line with prior research on linguistic text structure (e.g., Cook et al., 1998; Kintsch & van Dijk, 1978), McNamara (2013) distinguished between local cohesion (i.e., referential cohesion) and global cohesion (i.e., deep cohesion). Local cohesion reflects the "linguistic overlap (e.g., word overlap) or conceptual overlap (e.g., argument overlap) between adjacent sentences" (Ozuru et al., 2010, p. 644). Local cohesion helps readers to understand the meaning of adjacent sentences and allows them to process the text easily (McNamara, 2013; Ozuru et al., 2010). Global cohesion, on the other hand, refers to connecting information across the entire text (McNamara, 2013; Ozuru et al., 2010). It supports readers in linking “multiple events, elements, and facts, especially information spread across distant parts of the text” (Ozuru et al., 2010, p. 645). This connection across the entire text helps readers connect the information of the text in a “broader context outside the text,” which deepens their understanding (Ozuru et al., 2010, p. 645).

Prior research on *self-explaining* demonstrated that linguistic text complexity, such as cohesion, strongly influences students’ text processing and affects their learning (e.g., Lachner & Nückles, 2015; McNamara et al., 1996; McNamara & Knitsch, 2009). For instance, Ozuru et al. (2010) conducted a study to investigate whether linguistic text complexity (i.e., local cohesion) moderates the effectiveness of self-explaining. University students (N = 78) participated in the study in which linguistic text complexity was systematically manipulated within two conditions (i.e., low-cohesion condition versus high-cohesion condition). In the first experimental condition, students read an expository passage about reproduction, which was presented in its original form and reflected rather low levels of cohesion and was rather challenging to understand. In the second condition, students were provided with a revised text version, which contained more inferences and overlaps between adjacent sentences. The revised version showed higher levels of cohesion which should be less challenging for students to understand. After reading either the low-cohesive or the high-cohesive text, students of both conditions were asked to self-explain the content. Afterwards, they answered the posttest that contained factual knowledge questions. Interestingly, results showed that students benefited more from reading the low-cohesion text comparted to the high-cohesion text (d = 0.45, medium effect). The authors argued that the contribution of self-explanations to students’ comprehension is higher when they were confronted with a low-cohesion text. Low-cohesion texts contain local conceptual gaps which forces students to build connection by themselves during using additional learning activity, such as explaining (see Ozuru et al., 2010).

This study provides first evidence that text complexity (i.e., cohesion) may moderate students’ learning, at least in the field of self-explaining. However, whether this finding holds true for other formats of explaining, such as explaining in oral or written form to fictitious peers, is still an open question.

6.3 Does Perceived Social Presence Moderate Learning by Explaining?

Perceived social presence of the (fictitious) audience might be a further boundary condition of learning by explaining. This assumption is closely related to the assumption of social presence as the underlying mechanism of explaining (see Chapter 5). However, instead of social presence as a mediator of explaining, in this chapter social presence is discussed as a moderator of explaining. Perceived social presence is defined as the "degree of salience of the other person in the interaction" (Short et al., 1976, p. 65). The interest in perceived social presence during learning particularly rose with technology-mediated learning, such as in online environments (see Weidlich & Bastiaens, 2019). In this context, prior studies mainly analyzed the influence of perceived presence of *real* persons in technology-mediated communications and highlighted that higher levels of perceived social presence positively affected components that are closely related to students' learning, such as satisfaction and perceived learning (see Richardson et al., 2017, for meta-analytic evidence). Interestingly, prior research also analyzed the impact of perceived social presence of virtual others, such as virtual tutors. For instance, Kim (2013) conducted a study to investigate whether social presence of a fictitious peer moderates students' learning regarding their text comprehension (i.e., immediate and delayed posttest). She manipulated social presence within an experiment with three conditions: School students (N = 141) studied the learning content and were presented either with a virtual peer (i.e., high levels of social presence) or with a still image of a robot (i.e., low level of social presence). In a control condition, students learned the content without a simulation (i.e., no social presence). Regarding the immediate posttest, results revealed a main effect of social presence (η_p^2 = .06, medium effect). Post hoc analyses showed that presenting a virtual peer was more beneficial than presenting a still image of a robot or no simulation. Regarding the delayed posttest, analyzes revealed again a main effect of social presence (i.e., η_p^2 = .11, large effect). Post hoc analyses indicated that students who were in the virtual-peer condition outperformed students in the control condition regarding students' text comprehension. In summation, results showed that social presence of virtual others indeed moderated students' learning, as higher levels of induced social presence resulted in higher learning outcomes (see Kim & Baylor, 2007; Ryokai et al., 2003, for similar findings).

Whether social presence also moderates the effect of learning by explaining to fictitious others is still an open question. Hoogerheide et al. (2020) provided first steps forward solving this open research question as the authors manipulated social presence (i.e., to a real versus to a fictitious person) in their study. They conducted two experiments in which university students (experiment 1: N = 88; experiment 2: N = 92) read a text about electrical circuit troubleshooting and explained its content either to a fictitious person by recording their explanation or to a real person, who listened to their explanations. In a control condition, students restudied the provided content. Results of both experiments revealed a significant effect of condition (experiment 1: η_p^2 = .07, medium effect; experiment 2: η_p^2 = .09, medium effect). Contrasting the assumptions, post hoc analyses indicated that students in the control condition outperformed students who explained to a real person regarding their posttest performance. Additionally, explaining to a real person resulted in higher levels of excitement compared to students who explained to a fictitious person or students in the control group.

These results are in line with a growing body of research, indicating no beneficial effects of explaining (see also Fiorella et al., 2017; Fukaya, 2013; Hoogerheide et al., 2016; Lachner et al., 2020; Lachner et al., 2021, for further null findings). However, they contradict prior research on social presence that demonstrated positive effects on students' (perceived) learning (Kim, 2013; Kim & Baylor, 2007; Richardson et al., 2017; Ryokai et al., 2003). An unexpected result of this study was that the control condition outperformed students who explained to a real person but not students who explained to a fictitious person. Since prior research demonstrated that a moderate degree of excitement is most beneficial for learning (Aiello & Douthitt, 2001; Bond & Titus, 1983, see also Chapter 5 for further details on mediating processes of social presence), the authors discussed that explaining to a real person impaired learning because students' level of excitement was too high during explaining.

In summation, the study tested social presence by contrasting explaining to a fictitious versus a real person, assuming that perceived social presence was higher in real settings. However, whether manipulating the level social presence during explaining to fictitious others (i.e., low versus high levels of induced social presence) will affect learning is still an open question. Empirically investigating social presence as a moderator of explaining by systematically manipulating social presence within an experiment is, therefore, essential and should be considered in future research.

6.4 Does Prior Knowledge Moderate Learning by Explaining?

In addition to boundary conditions that rather relate to the implementation of explaining (e.g., explanatory modality, text complexity, social presence), there are also boundary conditions that are related to students' prerequisites which might influence students' learning. According to Ausubel (1968, p. vi) the "most important single factor influencing learning is what the learner already knows". This already existing prior knowledge has been analyzed in several studies over the last decades and plays a crucial role for the success of learning and relatedly of learning strategies (Baadte & Schnotz, 2014; Kalyuga, 2007; McNamara et al., 1996; Richter et al., 2016, 2018). In general, prior knowledge supports students to process new content more easily and reduces cognitive load during learning (Mayer, 2009; Richter et al., 2016, 2018). Therefore, it is not surprising that students have problems with learning new content when they have little prior knowledge in the corresponding topic, as they can only build limited connections between new information and their prior knowledge (Renkl, 2009). This assumption is in line with Castro-Alonso et al. (2021), who proposed an expertise reversal effect of prior knowledge and learning strategy. In their comprehensive review, the authors claimed that students with low prior knowledge especially benefit from additional support and instructional learning strategies, whereas students with high prior knowledge are able to manage their own learning in a more self-regulated manner (Castro-Alonso et al., 2021). In this context, generative learning strategies, such as explaining, can provide such additional support for students with low prior knowledge, which might enable them to make sense out of the new content. However, empirical evidence is rare.

Regarding *self-explaining*, McNamara and Scott (1999) conducted a study and provided evidence for moderation effects of prior knowledge. In their study, university students (N = 43) were randomly assigned to a self-explaining learning activity in which students learned about strategies to generate a self-explanation or to a control condition. In the training phase, students read four science texts. Students in the self-explaining condition additionally listened to a provided self-explanation of each text, detected strategies in these examples, and then actually explained the learned content to themselves. The control group simply read all texts out loud. After each text, all students answered comprehension questions. After this training phase, all students read a new text which was low coherent and rather difficulty to process, and then both conditions generated a self-explanation about the content. Afterwards, students answered a comprehension posttest which contained factual and transfer questions. Results showed that self-explaining was more beneficial than reading out loud, as students answered more questions correctly in the training phase when they generated an explanation (*d* = 0.75, large effect). Regarding students' comprehension in the posttest, results showed that prior knowledge had a significant influence on the effectiveness of the training (*d* = 1.70, large effect). Students with low prior knowledge showed a larger positive effect of self-explaining training on their comprehension compared to students with high prior knowledge (*d* = 0.96, large effect). These results indicated that additional learning strategies, such as explaining, are efficient strategies to especially support students with low prior knowledge as they need additional support to process new content successfully (see McNamara, 2004, 2017, for similar results regarding self-explaining). Nevertheless, students need to have at least some prior knowledge in the corresponding domain to be able to elaborate new information (Brod, 2020; Fiorella & Mayer, 2014; Mayer, 2014; Renkl, 2009).

Whether the moderation effect of prior knowledge regarding self-explaining also holds true for explaining to fictitious others is still unclear as there is a research gap of empirical investigations. One exception is the study by Hoogerheide, Renkl, et al. (2019). Again, university students (*N* = 61) read a text which they either explained to a fictitious peer afterwards or which they restudied. Results showed that students who explained outperformed students who restudied the provided content in both factual knowledge (*d* = 0.54, medium effect) and transfer knowledge (*d* = 0.53, medium effect). Interestingly, additional moderation analyses revealed that this effect mainly occurred due to a moderation effect of prior knowledge: Results showed a significant interaction between students' prior knowledge and conditions (*d* = 0.86, large effect). Additional regression analyses revealed that students' prior knowledge was significantly correlated with students' transfer knowledge in the control condition, but not in the explaining condition. Furthermore, students with high prior knowledge scored comparably among conditions, whereas students with low prior knowledge benefited more from generating an explanation to a fictitious peer. The authors argued that only students with low prior knowledge benefited from explaining compared to restudying the material. This study provides the first evidence that prior knowledge may also influence the effectiveness of explaining to fictitious others. In more detail, explaining seems to be especially beneficial for students with low prior knowledge. However, in this study both conditions (i.e., explaining versus restudying) were not directly compared. The authors only concluded this effect out of the correlation of students' prior knowledge and their transfer knowledge in the restudying condition. Therefore, research is needed which investigates students' prior knowledge as a potential boundary condition of learning by explaining to fictitious peers in more detail.

6.5 Does Academic Self-Concept Moderate Learning by Explaining?

Not only performance-oriented prerequisites, such as prior knowledge, but also belief-oriented prerequisites, such as academic self-concept, might influence the effectiveness of generating explanations. Academic self-concept is regarded as students' self-perceptions of their own competences, which are based on students' prior experiences in that domain (Marsh et al., 2017; Shavelson et al., 1976). Academic self-concept has been demonstrated to be a hierarchic construct, (Marsh et al., 2017; Marsh & Shavelson, 1985; Shavelson et al., 1976; Trautwein & Möller, 2016) which contains different subareas, for instance, academic self-concept in English or Mathematics, which are not necessarily related (see Figure 9). In other words, a student might have a high academic self-concept in English and a low self-concept in Mathematics.

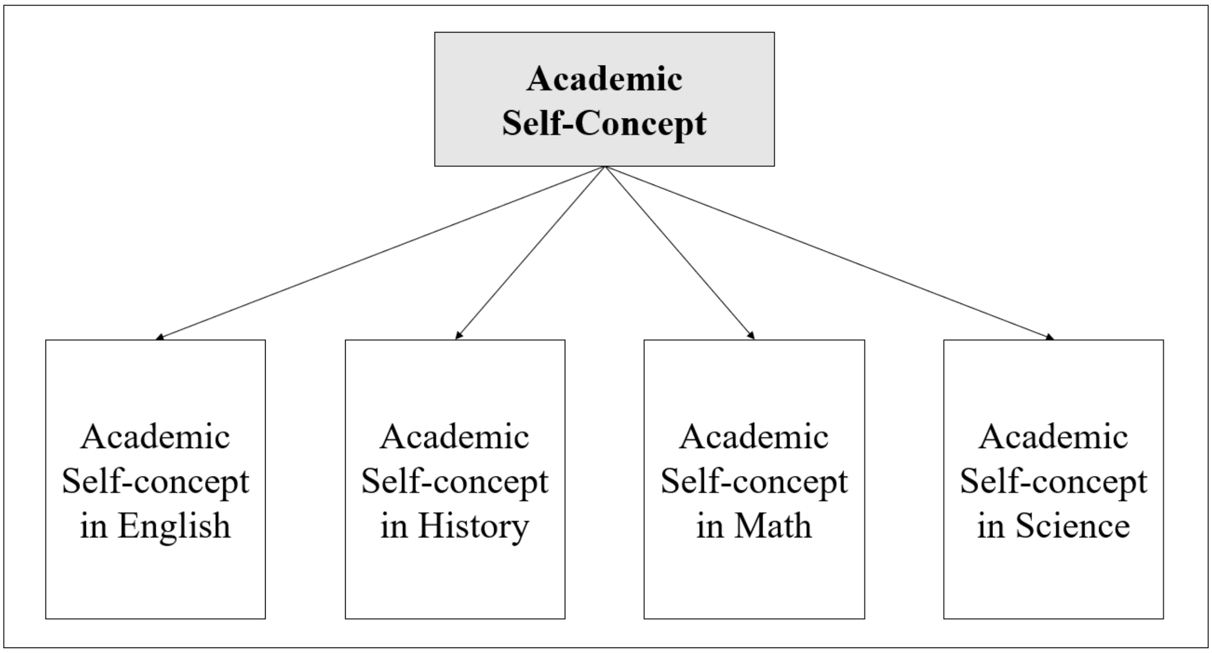

Figure 9. Extract of the hierarchic model of self-concept by Shavelson et al. (1976, p. 413). Reprinted by permission from SAGE Publications: Review of Educational Research, *Self-concept: Validation of construct interpretations*, Shavelson, Hubner, & Stanton, copyright, 1976.

The interest in academic self-concept rose in the last decades as it has been shown to be a crucial factor for learning, as it is positively correlated with students' comprehension (see Huang, 2011; Möller et al., 2009; Möller et al., 2020; Valentine et al., 2004, for meta-analytic evidence; Marsh & Martin, 2011; Marsh & O'Mara, 2008; Trautwein & Möller, 2016). Thus, students who rated their academic self-concept in a specific domain as high also showed high learning outcomes in this domain. Marsh et al. (2017, p. 87) argued that this correlation occurred due to a circle of self-perception and action: The "self-perceptions influence the way one acts, and these acts in turn influence one's self-perceptions".

Even though academic self-concept has been strongly investigated in the context of students' learning performance, there is a lack of investigating the influence of students' academic self-concept on the effectiveness of explaining. There is only one study which considered students' academic self-concept as an potential influence on a related learning strategy: Roelle and Renkl (2020) conducted a study to investigate whether the opportunity to review instructional explanations enhances example-based learning. School students (N = 58) participated in the study and were randomly assigned to one of two conditions (i.e., no-review condition versus review condition). Students first answered a scale about their academic self-concept in chemistry and answered a pre-test. Then, they learned three new chemistry topics and were provided with written instructional explanations which guided them through the new content. Additionally, they received two examples for each topic (i.e., six examples in total) and were asked to self-explain the example. They explained the example with either the opportunity to review the given instructional explanation (i.e., review condition) or without seeing the instructional explanation (i.e., no-review condition). Finally, all students answered the posttest. Results showed no difference among conditions. Interestingly, the authors reported a significant interaction between conditions and academic self-concept (d = 0.63, medium effect). Students with low academic self-concept benefited from the additional support of being able to review the instructional explanation during self-explaining, whereas students with high academic self-concept were hindered by this additional support, and benefited more from self-explaining without the review opportunity.

The interaction between students' academic self-concept and learning strategy might be surprising at first glance. However, these results which indicate that additional support is especially beneficial for students with lower levels of academic self-concept or of prior knowledge are in line with research on adaptive and individualized learning (see Corno, 2008;

Dumont, 2019, for reviews). Research on adaptive learning revealed that students with low self-concept and low prior knowledge have problems with independent and self-regulated learning and, therefore, strongly depend on additional learning support (e.g., Corno, 2008; Hurd, 2006; Nouri, 2016). In contrast, students with high academic self-concept and prior knowledge should be provided with less instructions to motivate them to self-regulate their learning processes (Corno, 2008; Dumont, 2019). Thus, it is not surprisingly that additional learning strategies, such as in the study by Roelle and Renkl (2020), especially helped students with low academic self-concept during learning. With this study, the authors provided the first empirical evidence that students' academic self-concept influences the effectiveness of learning strategies that are related to research on explaining. However, as it is the only experiment in this area, further research is needed that investigates potential influences of academic self-concept on learning by explaining. As prior research highlighted the crucial role of academic self-concept and its strong correlation to students' comprehension, it should be considered as a potential boundary condition of the effectiveness of explaining to fictitious others.

7

OVERVIEW OF STUDIES

7. Overview of Studies

7.1 ISEO Framework Model

Based on prior research, I generated the ISEO framework model (**I**mplementation- and **S**tudent-related boundary conditions of **E**xplaining on students' **O**utcome; see Figure 10). This model aims at systematically presenting different factors regarding the explaining effect (i.e., mechanisms, boundary conditions, outcomes), which resulted from reviewing prior research that revealed several research gaps. First, the underlying mechanism of explaining is still an open question. Following prior research, three possible hypotheses might explain the effectiveness of generating an explanation. The *retrieval practice hypothesis* claims that effects of explaining are a result of retrieving the content during explaining (Koh et al., 2018; Lachner et al., 2021). The *generative learning hypothesis* extended the latter by stating that the process of explaining triggers students to be engaged in elaborative processes that enhance deep learning (Chi et al., 1989; Lachner et al., 2018; Roscoe & Chi, 2008; Webb, 1989; Webb et al., 2009). Finally, the *social presence hypothesis* argues on two levels and claims that explaining to another (fictitious) peer involves higher levels of perceived social presence that might increase students' motivation to explain (i.e., *motivational perspective*; Hoogerheide, Visee, et al., 2019; Lachner et al., 2021) or may trigger students to adapt their knowledge to the audience's needs (i.e., *audience design perspective*; Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021). In the ISEO framework model, a graphical illustration of these hypotheses as underlying mechanisms is presented (see Figure 10). Again, although prior research discussed these three hypotheses as potential underlying mechanisms of the explaining effect, empirical evidence is rare and, therefore, highlights the need for further investigations.

Second, prior research on explaining to fictitious peers reported only mixed findings and a large variance across studies regarding students' comprehension and monitoring accuracy, indicating that further boundary conditions might determine the effectiveness of explaining (Kobayashi, 2019; Lachner et al., 2021). Therefore, I additionally reviewed prior research on learning by explaining or related generative learning strategies to analyze potential boundary conditions in a first step and included them in my framework model in the second step. Based on prior research, I focused on five potential boundary conditions that refer either to the implementation of explaining as a learning strategy or to students' prerequisites (see Figure 10). The first and most discussed potential boundary condition refers to the implementation of the learning task itself: The *explanatory modality* might be a crucial factor as prior research indicated that students' learning strongly depended on whether they explained in oral or in written form (Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021). In this context, prior research demonstrated mixed findings concerning the modality effect of explaining which might be due to differences between the provided learning materials, as they mainly differed in the level *text complexity* (Lachner & Nückles, 2015; McNamara et al., 1996; McNamara & Knitsch, 2009; Ozuru et al., 2010). Therefore, text complexity can be an additional implementation-related boundary condition of explaining as well (Hoogerheide et al., 2016; Lachner et al., 2018). However, in contrast to the explanatory modality, text complexity does not refer to the implementation of the task itself but to the provided learning material and, therefore, should be considered as another subfactor of implementation-related boundary conditions (see left side of ISEO model in Figure 10). Third, besides the explanatory modality and text complexity, reviewing prior research revealed a further potential implementation-related boundary condition: Perceived *social presence* during explaining, meaning that the learning environment in which students explain might be a crucial boundary condition as well. Students who explain in an environment with high levels of social presence may be more motivated to explain to the audience compared to perceiving lower levels of social presence (Hoogerheide, Renkl, et al., 2019; Lachner et al., 2018; Lachner et al., 2021). Thus, I assume that social presence has a curial impact on the effectiveness of learning by explaining and, therefore, included it as a potential boundary condition in the framework model (see Figure 10).

In summation, reviewing prior research revealed three potential implementation-related boundary conditions which might have strong impacts on the effectiveness of explaining. Thus, the explaining effect seems to depend on how the strategy is implemented (i.e., explanatory modality), which material is provided (i.e., text complexity), and in which learning environment the students generate their explanations (i.e., social presence).

Analyzing prior research not only revealed implementation-related boundary conditions, but also student-related boundary conditions, as the effectiveness of explaining seems to depend on students' prerequisites as well. In this context, prior research on learning by explaining and related generative learning strategies pointed out two main prerequisites that should be considered during implementing learning strategies as they showed to highly affect their effectiveness: Prior knowledge and academic self-concept (see right side of ISEO model in Figure 10). First, prior knowledge showed to be a crucial *performance-oriented* prerequisite, which is essential to transfer new learned content into the long-term memory and to enhance meaningful learning, as this learning process is based on connecting new information with existing knowledge (Fiorella & Mayer, 2016; Mayer, 1996, 2009). In other words, students who have only little prior knowledge also have poor connection opportunities, which hinders their learning process. In this context, it is not surprising that learning strategies, such as explaining, especially help students with low prior knowledge, as it served as an additional support to help students concentrate on the relevant information (Hoogerheide, Renkl, et al., 2019; McNamara, 2004, 2017; McNamara & Scott, 1999). As a conclusion, I integrated prior knowledge as a performance-oriented student-related boundary condition in the framework model and assumed that additional learning strategies, such as explaining, should be particularly beneficial for students with low prior knowledge.

Besides prior knowledge, students' academic self-concept is discussed as an important factor for students' learning in prior research (Marsh & Martin, 2011; Marsh & O'Mara, 2008; Trautwein & Möller, 2016). In contrast to prior knowledge, academic self-concept does not refer to students' performance, although they are closely related (e.g., Huang, 2011; Möller et al., 2009; Möller et al., 2020; Valentine et al., 2004). However, academic self-concept rather represents a *belief* that reflects how good students think they are in a specific academic subject. Interestingly, and similar to results on prior knowledge, studies indicated that additional learning support, such as generative learning strategies, is especially beneficial for students who report lower beliefs in themselves regarding a specific topic; thus, who show lower levels of academic self-concept (e.g., Roelle & Renkl, 2020). Therefore, based on prior research, I assumed that students' prior knowledge and academic self-concept might affect the effectiveness of explaining and included both as potential boundary conditions in my framework model. Again, empirical evidence of prior knowledge and academic self-concept on the effectiveness of explaining is rare and, therefore, needs further research.

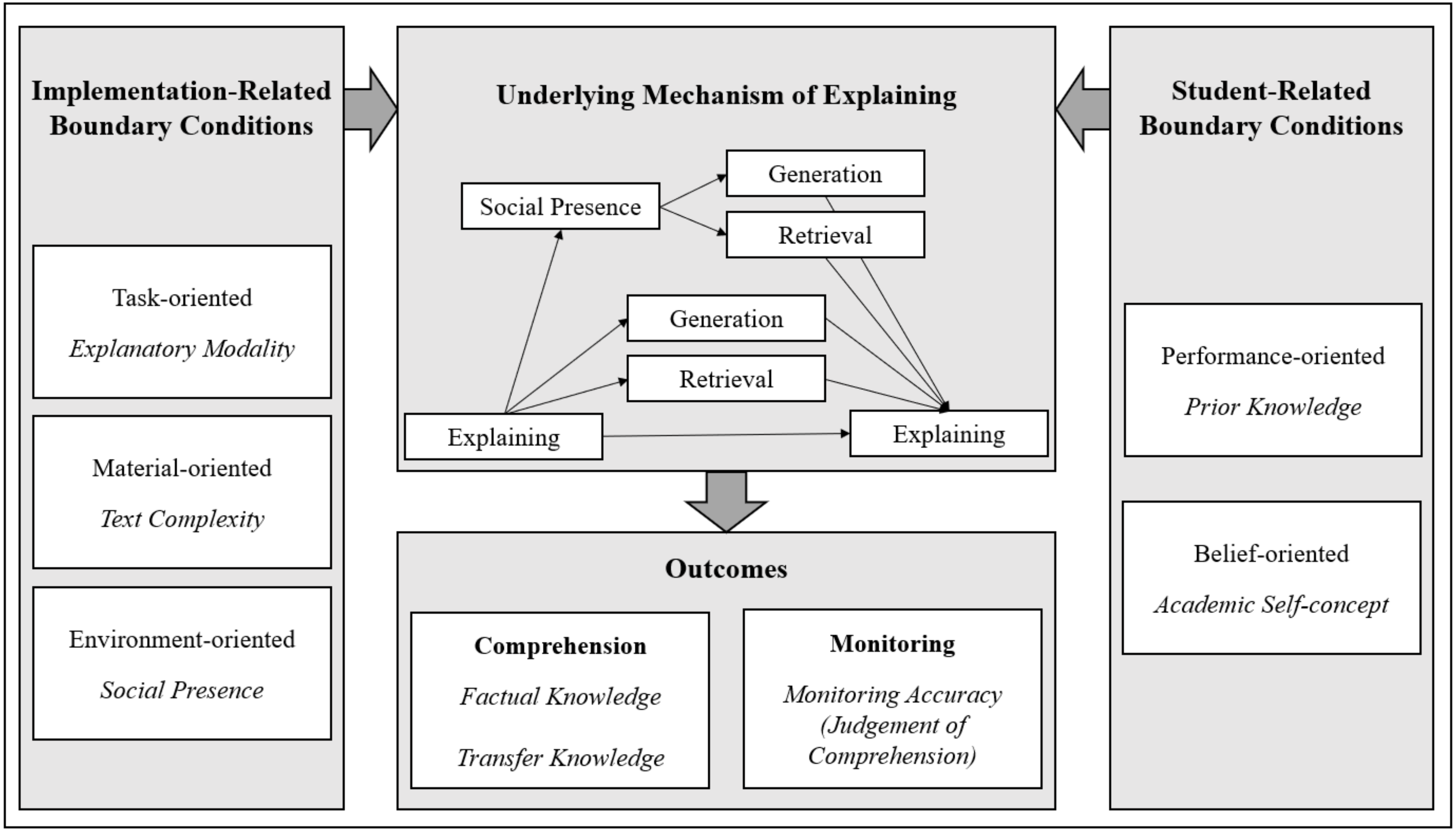

Figure 10. Model of Implementation- and Student-related boundary conditions of Explaining on students' Outcome (ISEO Model).

As an interim conclusion, five potential boundary conditions of learning by explaining can be derived from previous research. These boundary conditions are either implementation-related or student-related and might influence the effectiveness of explaining on students' outcome. Students' outcome, in this context, is twofold as prior research analyzed the effect of explaining on both students' comprehension and their monitoring skills. Regarding to the former, recent research mainly distinguished between two types of knowledge to measure students' comprehension, namely, factual and transfer knowledge. As stated before, generative learning strategies, such as explaining, should trigger students to generate new knowledge due to building connections between new information itself and its content and prior knowledge (Chi, 2009; Chi & Ohlsson, 2005; Chi & Wylie, 2014). Building inferences should enable students to answer transfer questions appropriately. In contrast, factual knowledge represents lower-order knowledge, thus, reflects information and concepts that are directly stated in the text. Retrieving provided content might, therefore, be especially beneficial for answering factual questions. This differentiation is important to consider in empirical studies as explaining versus retrieving might affect students' comprehension regarding factual versus transfer knowledge differently. Thus, factual knowledge and transfer knowledge are both integrated as learning outcomes in the ISEO framework model.

Besides students' comprehension, reviewing prior studies revealed that explaining is also beneficial for students' monitoring accuracy. Accurate monitoring is an essential skill for successful learning as it enables students to detect knowledge gaps for which they can adapt they learning strategies accordingly for achieving their learning goals, for instance, by restudying specific learning content (Baars et al., 2017; Fiorella & Mayer, 2016; Nückles et al., 2009). In this context, prior research indicated that the process of explaining does not only positively affect students' comprehension but also their monitoring accuracy (Fukaya, 2013). As a conclusion, the ISEO framework model combines both outcomes, highlighting that comprehension as well as monitoring skill should be analyzed in studies. Again, empirical evidence is rare which makes further research essential to build evidence for the presented model.

In summation, the ISEO framework model focuses on three aspects of the explaining process. First, it provides three possible underlying mechanisms of explaining. Second, it considers five boundary conditions that are distinguished in implementation and student-related boundary conditions. Lastly, it includes two outcomes, namely students' comprehension and their monitoring skills. In this context, I would like to highlight that the framework model should be regarded as a graphical representation of prior research’s conclusions, meaning that empirical investigation of the model is essential to test its validity.

7.2 Overview of Studies and Research Questions

The aim of this dissertation was to address several open questions revealed from prior research. For this purpose, I generated the ISEO framework model, which I empirically tested with three empirical studies. Empirically testing the framework model allows to draw further conclusions about the explaining effect and to solve open research questions simultaneously. In all three experiments, I investigated the effectiveness of explaining versus retrieving on students' comprehension (i.e., factual knowledge and transfer knowledge) and their monitoring abilities (i.e., monitoring accuracy). This approach extends prior research in two ways: First, prior research mainly contrasted an explaining condition only to a restudying condition (see Koh et al., 2018; Lachner et al., 2020; Lachner et al., 2021, for exceptions). Second, prior research mainly analyzed the explanatory effect on students' comprehension (see Fukaya, 2013; Lachner et al., 2020, for an exceptions). Examining both students' comprehension and their monitoring accuracy as outcomes of explaining, therefore, represents an addition to prior studies.

In this context, I stated the first hypothesis, the *explaining hypothesis*, that students who generate an explanation outperform students who retrieve the content regarding students' comprehension (RQ 1a) and their monitoring accuracy (RQ 1b). Additionally, I aimed at analyzing potential boundary conditions of explaining. As the explaining modality seems to be a crucial factor that determines the explaining effect, I tested the effect of explanatory modality in all three studies. As prior research indicated an interaction between explanatory modality and text difficulty, I stated the *modality hypothesis*, and assumed that oral explaining is more beneficial than writing explanations regarding students' comprehension (RQ 2a) and their monitoring accuracy (RQ 2b), however, only when the learning material is challenging and complex (see Hoogerheide et al., 2016; Lachner et al., 2018).

To test this interaction hypothesis, I conducted a laboratory experiment (Study 1) with 115 university students to investigate whether the text complexity influences the effectiveness of generating oral versus written explanations (see Jacob et al., 2020, for the published study). Students first read either a low or a highly complex text and explained the content afterwards either in oral or in written form. Students in the control condition retrieved the provided content with a free recall test. For low complex learning material, I hypothesized that students who generate an explanation (i.e., in oral or in written form) outperform students who retrieve the content regarding students' comprehension and monitoring accuracy (see RQ1: e*xplaining hypothesis*). Regarding the highly complex learning material, in addition to the explaining effect, I assumed that students who explained orally outperformed students who generated a written explanation regarding students' comprehension and monitoring accuracy (see RQ 2: m*odality hypothesis*).

In Study 2, I conducted another laboratory experiment with 137 university students to investigate whether social presence during the learning task has an influence on the effectiveness of explaining based on prior research (Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021). Again, I assumed that students who explain the learned content outperform students who retrieve the content (see RQ1: *explaining hypothesis*), and that oral explaining results in higher outcomes than writing explanations (see RQ2: *modality hypothesis*). Additionally, I stated the *social presence hypothesis* and assumed that induced social presence resulted in higher outcomes regarding students' comprehension (RQ 3a) and their monitoring accuracy (RQ 3b). For these hypotheses, I replicated Study 1 by using the complex material and conditions. Thus, students randomly explained the content in oral versus written form, or retrieved the content. As a fourth condition, I implemented a further writing condition in which social presence was induced to support students during writing explanations as they showed lower learning outcomes in Study 1. Social presence was induced by providing students with a messenger chat that contained specific social cues, such as the identity of the fictitious peer, which increased perceived social presence in prior research (see Weidlich & Bastiaens, 2019, for similar approach).

In Study 3, two aims were pursued: First, since prior research demonstrated that generative learning strategies, such as explaining, mainly depend on students' development (Brod, 2020), the effectiveness of generating oral versus written was investigated with school students. Second, potential influences of student-related boundary conditions (i.e., prior knowledge and academic self-concept) on students' comprehension and their monitoring accuracy were analyzed, as prior research showed a benefit of additional learning activities for students with lower prerequisites (Hoogerheide, Renkl, et al., 2019; Roelle & Renkl, 2020). In total, 132 seventh grade school students participated in the experiment. Again, I hypothesized that generating either oral or written explanations increases students' comprehension and their monitoring skills compared to retrieving the content (see RQ1: *explaining hypothesis*). Additionally, oral explaining was expected to result in higher learning outcomes than writing explanations (see RQ2: *modality hypothesis*). Additionally, I explored whether students with low prior knowledge benefit more from explaining than students with high prior knowledge regarding their comprehension (RQ 4a) and their monitoring accuracy (RQ 4b), that is the *prior knowledge hypothesis*. I assumed the same pattern for students' academic self-concept and explored whether students with low academic self-concept benefit more from explaining than students with high academic self-concept regarding their comprehension (RQ 5a) and their monitoring accuracy (RQ 5b), that is, the *academic self-concept hypothesis*.

Besides analyzing the effects of potential boundary conditions of explaining, I, additionally, investigated the underlying mechanisms of the explanatory effect. In all three experiments, I addressed this unresolved issue. To investigate the explanatory mechanism, I generated a coding system, and I analyzed all generated explanations in detail. This coding system contained a measure as an indicator for each hypothesis (i.e., *retrieval practice hypothesis*, *generative learning hypothesis*, *social presence hypothesis*). Regarding the *retrieval practice hypothesis*, I investigated the number of mentioned *concepts* within the explanations based on related approaches (see Lachner et al., 2017a). Being able to explain many concepts reflects the active process of retrieving the provided content and, therefore, are a good indicator to measure students' retrieval during explaining. Regarding the *generative learning hypothesis*, I analyzed the number of *elaborations* within students' explanations. Elaborations are an indicator for deep and meaningful learning, resulting in new knowledge structures, as students need to connect new content to their prior knowledge to be able to formulate information that go beyond the provided content (see Chi et al., 1989; Chi et al., 1994; Lachner et al., 2018; Webb, 1989; Webb et al., 2009). Lastly, regarding the *social presence hypothesis*, I analyzed the number of mentioned *personal references* as an indicator for students' perceived social presence during explaining (see Lachner et al., 2018). Personal references (e.g., „I“, „you“), reflect a connection with the audience and, therefore, represents an indicator for perceived social presence (Akinnaso, 1985; Chafe, 1982; Chafe & Tannen, 1987; Einhorn, 1978). To gain more insight into what mechanism might underly the explanatory effect, I evaluated students' generated explanations according to the coding system and analyzed whether concepts, elaborations, and personal references as indicators for the different hypotheses mediated the effectiveness of learning by explaining to fictitious peers.

7.3 Procedure of Studies

All three dissertational studies followed the same approach. First, all studies received approval to conduct the corresponding study from the ethic commission, or, in case of data collection with school students, from the responsible council of governments. Second, an a-priori power analysis was conducted before recruiting students to ensure a power of 80 percent. Third, before data collection, written consent from the participating students were obtained; and, in case of underage students, consent was also obtained from their parents. Additionally, the procedures were rather similar across studies, thus, an overview of the general procedure is provided in this chapter (see Figure 11). The general procedure is based on prior research on explaining (e.g., Fukaya, 2013; Hoogerheide et al., 2014; Hoogerheide et al., 2016; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019; Lachner et al., 2018; Lachner et al., 2020). In the beginning of the study, control variables, such as self-efficacy in explaining, were measured with Likert scales. Then, students answered a prior knowledge test on the related topic. In Study 1 and 2, the same materials were implemented, thus, the prior knowledge test contained five questions with an open-ended answer format. In Study 3 with school students, the prior knowledge test contained five single-choice questions. After completing the test, students were engaged in the study phase. Students were provided with text materials in Study 1 (i.e., texts about reproduction and immunology) and Study 2 (i.e., text about immunology). In Study 3, students attended a lesson on photosynthesis. After the study phase, all students rated their cognitive load (i.e., perceived mental effort and subjective difficulty) related to the lesson and judged their current comprehension (i.e., monitoring judgement) as a control variable for intra-individual differences. Then, the manipulation took place. Students were randomly engaged to the different conditions. In all studies, at least two explaining conditions were implemented to investigate differences between the explanatory modality: students explained either in oral or in written form. In Study 2, as an exception, a third explaining condition was implemented in which social presence was intentional induced during writing an explanation. Additionally, in all three experiments the explaining conditions were contrasted to a control condition in which students were engaged in a free call retrieval practice. Students had 10 minutes to complete the learning activity, except for Study 3 with school students, in which students completed the task in a self-paced manner, as the task should represent a ‘real’ learning environment, such as learning at home. Again, students rated their cognitive load related to the learning activity and judged their current understanding to investigate whether explaining improves students monitoring accuracy. Finally, students answered the posttests which contained factual and transfer knowledge questions. Again, in Study 1 and 2 an open-answered format was used for both factual knowledge questions (i.e., six questions) and transfer questions (i.e., six questions). In Study 3, the factual knowledge posttest comprised nine single-choice questions and the transfer posttest four questions with an open-ended format.

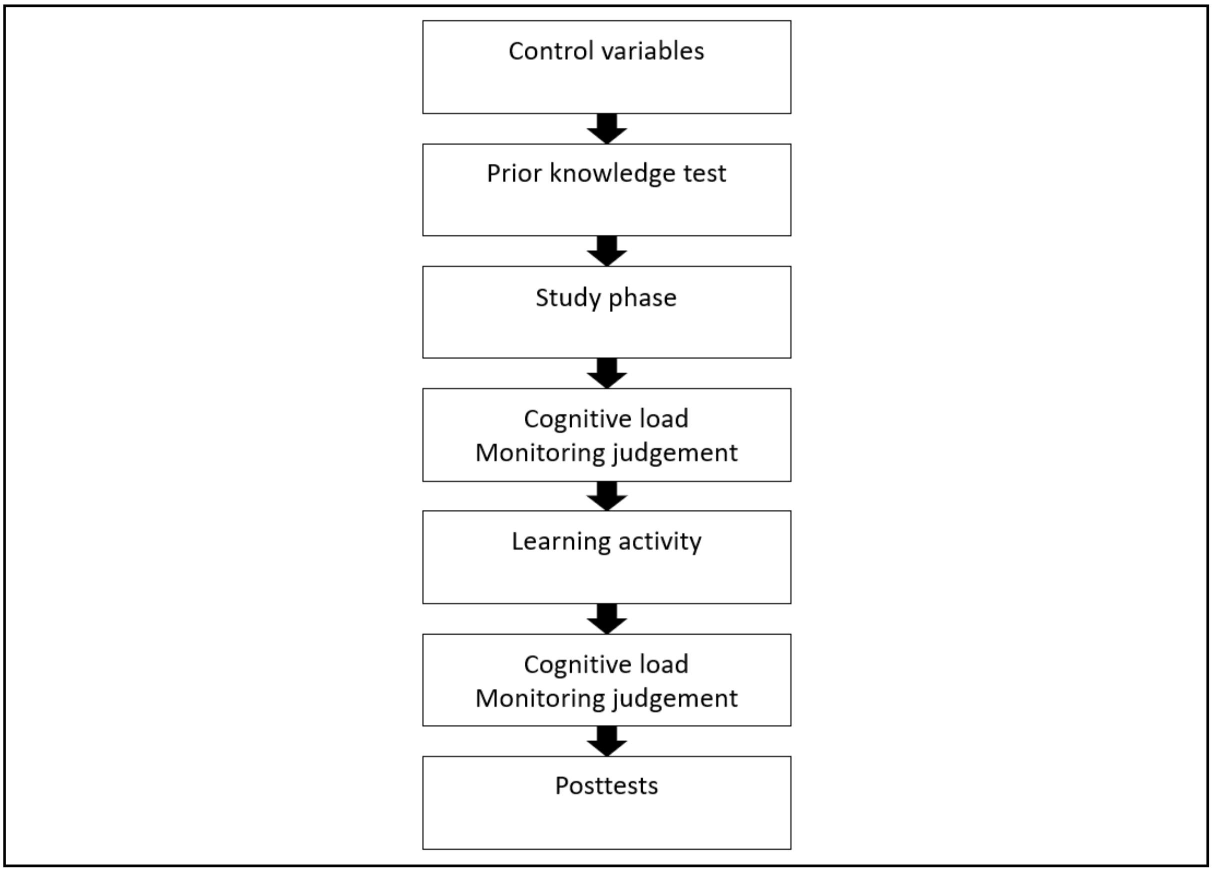

Figure 11. Procedure of the studies.

8

STUDY 1: LEARNING BY EXPLAINING ORALLY OR IN WRITTEN FORM? TEXT COMPLEXITY MATTERS

8. Study 1: Learning by Explaining Orally or in Written Form? Text Complexity Matters

8.1 Abstract

In this experiment, we examined whether linguistic text complexity affects effects of explaining modality on students' learning. Students (*N* = 115) read a high-complex and a low-complex text. Additionally, they generated a written or an oral explanation to a fictious peer. A control group of students retrieved the content. For the low-complex text, we found no significant differences between conditions. For the high-complex text, oral explaining yielded better comprehension than writing explanations. The retrieval condition showed the lowest performance. Mediation analyses revealed that the effect of explaining modality while learning from the high-complex text was mediated by the personal references and the comprehensiveness of the generated explanations. Our findings suggest that the effect of explaining modality emerges when students are required to learn from difficult texts. Furthermore, they show that oral explaining is effective as, likely due to increases of social presence, it triggers distinct generative processes during explaining.

Published as: Jacob, L., Lachner, A., & Scheiter, K. (2020). Learning by explaining orally or in written form? Text complexity matters. *Learning & Instruction.*
https://doi.org/10.1016/j.learninstruc.2020.101344

8.2 Theoretical Background

Learning from texts is a prevalent learning activity at all levels of education. Therefore, students are required to select the relevant information from the text as well as organize and elaborate the selected information to gain a deep understanding (Duran, 2016; Webb et al., 1995; Wharton & Kintsch, 1991). Throughout these cognitive processes, students have to realize distinct metacognitive processes, such as monitoring their understanding to be able to detect and resolve potential comprehension problems (Fiorella & Mayer, 2016; Kiewra, 2005; Nückles et al., 2009). Despite the ubiquitous use of text materials, research indicates that students often face difficulties in realizing these strategies spontaneously which may constrain their general comprehension (O'Reilly et al., 2004; Pressley et al., 1992; Snow, 2002), particularly when students are required to learn from digital texts (Kong et al., 2018; Singer Trakhman et al., 2018). Generating explanations of the learning contents (to fictitious other students) has been shown to be an effective additional learning activity to support students' text comprehension (Duran, 2016; Fiorella & Mayer, 2013, 2014, 2016; Lachner & Neuburg, 2019; Pilegard & Fiorella, 2016; Roscoe, 2014; Roscoe & Chi, 2008; Topping, 2005; Webb et al., 1995; Williams & Lombrozo, 2010). In his meta-analysis, Kobayashi (2019) documented the beneficial effects of explaining. Results indicated a medium effect of explaining to fictitious others on students' comprehension (g = 0.48, based on 28 comparisons). However, the heterogeneity among the included studies was high, indicating that additional factors constrained the effectiveness of explaining. Against this heterogeneity of empirical evidence, the sparse number of empirical studies investigating potential boundary conditions of explaining is surprising (Hoogerheide et al., 2016; Koh et al., 2018; Lachner et al., 2018 for exceptions). Recently, for instance, researchers started to investigate moderation effects of the modality of explaining (Hoogerheide et al., 2016; Lachner et al., 2018). Thus, they analyzed whether students profited more from generating oral or written explanations after reading the learning material regarding students' comprehension. However, results only indicated mixed findings. They either demonstrated that oral explaining was superior to writing explanations or showed that generating oral and written explanations was comparably effective. In this study, we therefore, argue that the modality of explaining is not a boundary condition per se, but rather depends on the complexity of the text material which is used during the study phase. Such effects can be assumed, as previous research on text comprehension highlighted the crucial role of linguistic text complexity features particularly for novice students (O'Reilly & McNamara, 2007b; Ozuru et al., 2010). Therefore, the current study contributes to a better understanding of whether and how the modality of explaining may support novice students' learning, depending on the complexity of the text material of the pre-given study phase.

8.2.1 Learning by Explaining

Explaining is a learning activity which has primarily been used in interactive learning situations in which students either generate explanations to peers (Duran, 2016; Palinscar & Brown, 1984; Plötzner et al., 1999; Roscoe, 2014; Roscoe & Chi, 2008; Topping, 2005; Webb et al., 1995) or to pedagogical agents in computer-supported environments (Chin et al., 2010; Okita & Schwartz, 2013). In line with the generative learning theory (Mayer, 1996; Wittrock, 2010), generating an explanation is regarded to enable learners to build new knowledge by triggering deep-level *cognitive processes* (e.g., organization and integration of information by making examples, see Fiorella & Mayer, 2014, 2016; Pilegard & Fiorella, 2016; Wittrock, 2010). Additionally, explaining may enable to realize distinct *metacognitive processes*, as students are required to externalize their current knowledge which may trigger them to monitor their current level of comprehension (cf., meta-comprehension, see Fukaya, 2013; Lachner et al., 2020). Meta-comprehension is operationalized as the correspondence between students' judgement of their current comprehension and their actual performance on a comprehension test (see Golke & Wittwer, 2018; Maki & McGuire, 2009; Wiley et al., 2005). Accurate meta-comprehension judgments enable students to detect and resolve potential comprehension problems during generative activities, such as explaining (Fiorella & Mayer, 2016; Kiewra, 2005; Nückles et al., 2009).

Interestingly, the beneficial effects of generating explanations have been demonstrated to result also in indirect settings in which students provide an explanation to a non-present, fictitious student (Fiorella & Mayer, 2013, 2014; Hoogerheide, Visee, et al., 2019). For instance, Fiorella and Mayer (2014, Experiment 2) examined the effects of preparing to explain the learning contents versus preparing and explaining the contents to fictitious others on students' learning. First, university students received an expository text about the Doppler effect and were informed that they would answer a test (test expectancy) or that they would explain the learning contents to a fictitious student (explaining expectancy) after the study phase. Afterwards, half of the students either restudied the text (no teaching condition) or provided an explanation on video (explaining condition). The authors found that explaining expectancy only improved students' short-term retention, (teaching expectancy effect: d = 0.55) but that the actual act of explaining was particularly beneficial for long-term retention (explaining effect: d = 0.56). Hoogerheide et al. (2014, Experiment 1) asked students to learn from an instructional text on syllogistic reasoning. Two groups were required either to prepare for a test or to prepare to explain the learning materials. A third group of students prepared to explain, and additionally explained the learning contents on video. As in the study by Fiorella and Mayer (2014), students who provided an explanation outperformed students who only prepared to explain (d = 0.75). In a related study, Fukaya (2013) demonstrated that explaining not only contributed to better comprehension, but also to more accurate monitoring during explaining, as a crucial metacognitive outcome (cf., meta-comprehension accuracy). In his study, students who generated an explanation to a fictitious student showed higher levels of meta-comprehension accuracy than students who only expected to explain or a control condition, which only produced keywords of the learning material. These findings demonstrate that the act of generating explanations can contribute to students' comprehension and meta-comprehension.

8.2.2 The Modality of Explaining Affects Students' Learning

Recently, researchers aimed at replicating the effects of learning by explaining in written form, as writing explanations can be more easily implemented in educational practice and may be a more feasible instructional approach during teaching. For instance, Hoogerheide et al. (2016, Experiment 1) used the identical learning materials from their previous study (Hoogerheide et al., 2014), and investigated potential effects of the intention to provide an explanation or to actually write an explanation. Contrary to their previous findings, neither the intention to provide an explanation nor writing an explanation affected students' learning (see Fukaya, 2013, for related findings). In their second experiment, the authors directly compared written versus oral explaining, and additionally included a control condition in which students simply restudied the learning material. Results indicated that explaining on video (d = 0.43) but not writing an explanation (d = 0.19) was more effective than restudying. However, there were no significant differences between the written and oral condition (d = 0.24). The authors supposed that generating oral but not written explanations triggered students' awareness of social presence, which inclined them in more generative processes, such as reorganizing and elaborating the content toward the need of potential recipients (Chafe, 1982; Chafe & Tannen, 1987; Sindoni, 2013). Additional explorative content analyses showed that students in the oral explaining condition used more personal references such as "me", and "you" than students who wrote an explanation (d = 1.54), as potential behavioral proxies for the perceived social presence during explaining (see also Akinnaso, 1985; Chafe, 1982; Chafe & Tannen, 1987;

Einhorn, 1978). However, the increased social presence did not result in a better comprehension.

In their replication study, Lachner et al. (2018) provided students with a Wikipedia entry on combustion engines as learning material. Afterwards, the students either generated an oral or written explanation. Contrary to Hoogerheide et al. (2016), the authors found that students who explained orally outperformed students who wrote an explanation (d = 0.67). Additional analyses showed that students who generated an oral explanation, again used more personal references (d = 0.98). At the same time, they found that the higher test performance of students in the oral explaining condition was explained by higher levels of generative processes, as students provided more elaborations in their explanations (d = 1.53). The authors suggested that the differences between their study and the study by Hoogerheide et al. (2016, Experiment 2) can be ascribed to the different learning materials. Whereas Hoogerheide et al. (2016) used well-designed instructional material which likely can be characterized by low levels of linguistic text complexity for novice students, Lachner et al. (2018) used a naturalistic expository text from Wikipedia with likely higher levels of linguistic text complexity for novices, which may account for the contradictory findings. Following these suggestions, although oral explaining may generally trigger higher levels of social presence regardless of the level of linguistic text complexity, it may be superior to writing explanations only when students are required to learn from high-complex text material (as in the study by Lachner et al., 2018). When learning from low-complex text material (as in the study by Hoogerheide et al., 2016), the different explaining modalities may be comparably effective, as students may have ample cognitive resources available to process the information of the learning material (Ozuru et al., 2009). When learning from high-complex texts, subsequent learning activities such as oral explaining may be effective to trigger additional generative process (e.g., elaborations) that aim at building a coherent situation model of the contents, particularly, when students have low prior knowledge regarding the topic.

8.2.3 Linguistic Text Complexity Affects Students' Learning

Following McNamara (2013), linguistic text complexity can be described as a multi-dimensional construct comprising different text features, that is syntactic complexity, concreteness, and cohesion of a text, that constrain the comprehensibility of texts, particularly for novice students.

Syntactic complexity refers to the difficulty of the phrase structure and the associated dependencies within a text (Berendes et al., 2018). For instance, complex syntactical structures comprise the use of nominals, or the inclusion of subordinate clauses. Complex syntactical structures often unnecessarily increases the sentence length, and as such increases readers' cognitive efforts to process the text (Berendes et al., 2018). Contrarily, concreteness refers to the use of content words that are concrete and meaningful which make texts easier to process and comprehend (McNamara, 2013). Finally, cohesion refers to explicit indicators that mark distinct relations between sentences (e.g., connectives, argument overlap) to process adjacent sentences as meaningful units (Ozuru et al., 2010). Thus, during processing expository texts, cohesion mainly helps students establish a coherent initial understanding because they highlight interrelations among adjacent sentences. Research indicated that linguistic text complexity features, such as cohesion or concreteness, play a distinct role in comprehending texts (Hinds et al., 2001; Lachner & Nückles, 2015; McNamara et al., 1996; McNamara, 2013; McNamara & Knitsch, 2009; Nickerson, 1999). In this study, we used learning material which differed regarding cohesion and concreteness as an indicator for linguistic text complexity (see Materials for details).

8.2.4 The Present Study: Effects of Explaining Modality and Linguistic Text Complexity on Students' Learning

Based on the previous considerations, we conducted an experiment in which we investigated effects of explaining modality (i.e., generating written versus oral explanations) and linguistic text complexity (high-complex versus low-complex text), and potential interaction effects on students' learning. For this purpose, we provided university students with two different texts from a previous study Golke and Wittwer (2018) which systematically differed regarding linguistic text complexity (i.e., level of cohesion and word concreteness, see McNamara, 2013), and showed to affect students' text comprehension and meta-comprehension accuracy. Additionally, as a subsequent learning activity, the students explained the contents of the learning materials either in written or oral form. A control condition retrieved the contents of the learning material. Retrieving material is an effective and robust strategy to enhance learning (c.f., testing effect, Roediger & Karpicke, 2006b), and therefore, can be regarded as a strong control condition in comparison to explaining. We stated the following hypotheses (which were preregistered on As.Predicted, https://aspredicted.org/sv3zv.pdf).

8.2.4.1 Learning-Outcome Hypotheses

Regarding effects of linguistic text complexity, we hypothesized that students should gain higher test scores when learning from the low-complex text than from the high-complex text. Regarding effects of explaining modality, based on the contradictory findings by Hoogerheide et al. (2016) and Lachner et al. (2018), we hypothesized an interaction effect with linguistic text complexity: For the *low-complex text*, we assumed that students who provide an explanation (i.e., in written or oral form) should outperform students in the retrieval condition. The effect of the explaining modality should be less pronounced in the low-complex text condition, as students would have ample cognitive resources available (i.e., mental effort) to build an adequate understanding of the text which should result in comparable learning gains of the two explaining conditions. However, for the *high-complex text*, generating an oral explanation should be more beneficial than generating a written explanation (Lachner et al., 2018), as oral explaining may occupy fewer cognitive resources than writing during explaining. The lower cognitive load may result in more generative processing and contribute to students' comprehension. Students who only retrieved the contents of the learning material (i.e., retrieval practice) should show the lowest performance on the knowledge test.

8.2.4.2 Additional Explorative Analyses

Regarding the underlying generative processes which account for the differences within students' learning performance in the high-complex text condition, we analyzed specific characteristics of the generated oral and written explanations. Based on previous research (Chafe & Tannen, 1987; Hoogerheide et al., 2016; Lachner et al., 2018), we assumed that oral explaining would increase the awareness of social presence, as indicated by more personal references in the explanations. The increased awareness of social presence would incline students to provide more comprehensive explanations (as measured by the number of explanatory concepts, Boshuizen & Schmidt, 1992) and more elaborations (Chafe & Tannen, 1987; Lachner et al., 2018) to make the content tangible to a fictitious novice student. These generative processes should finally contribute to students' comprehension.

As further dependent variables, we collected subjective ratings of students' mental effort and perceived difficulty as potential indicators of their perceived cognitive load. Additionally, since generating explanations has demonstrated to result in higher levels of students' monitoring accuracy (Fukaya, 2013; Lachner et al., 2020), we further collected students' meta-comprehension ratings, as proxies for their meta-comprehension accuracy.

8.3 Method

8.3.1 Participants and Design

To obtain a novice student sample, we recruited university students (*N* = 119) from study programs which were not related to the study topic (i.e., biology). We excluded students who did not complete the study (*n* = 3) or reached very low performance (i.e., zero points) in the reading skill test (*n* = 1).

The mean age of the students was 25 years (*SD* = 7.76) and 77 % of them were female. The students were advanced students, on average in their 5th semester (*SD* = 3.72), had low prior topic knowledge ($M_{\text{low-complex text}} = 2.75$, $SD_{\text{low-complex text}} = 1.19$; $M_{\text{high-complex text}} = 2.17$, $SD_{\text{high-complex text}} = 1.06$), and had excellent reading skills (*M* = 22.89, *SD* = 8.17). Most students were enrolled in humanity programs (*n* = 89).

We conducted a 2×3 mixed factorial design with linguistic text complexity as a within-students factor (low-complex text vs. high-complex text) and learning activity as a between-students factor (generating oral vs. written explanations vs. retrieval practice). Students were randomly assigned to one of the three learning activities (oral explanation group: *n* = 40; written explanation group: *n* = 39; retrieval group: *n* = 36). The dependent variable was students' text comprehension, measured by two knowledge subtests comprising text-based questions and inference-based questions.

To control for potential differences regarding students' prerequisites, we controlled for prior knowledge and reading skills (Ozuru et al., 2009). Additionally, we collected data on students' perceived mental effort, subjective difficulty, and their meta-comprehension judgments.

The size of the recruited sample exceeded the required sample size of 102, as determined by an a-priori power analysis for contrast analyses (GPower, Version 3.1.9.2). Power was set to .80, α-error to .05, and the assumed effect size to $\eta^2_p = .05$, as recent studies documented medium effects of explaining modality (Hoogerheide et al., 2016; Lachner et al., 2018).

8.3.2 Materials

The entire experiment was presented in the Qualtrics online survey tool (https://www.qualtrics.com).

8.3.2.1 Study Texts

We used two validated German texts from the domain of biology from the study by Golke and Wittwer (2018). The first text was about reproduction (low-complex text) and dealt with types of reproduction (i.e., sexual versus asexual reproduction). The second text was about immunology (high-complex text) and dealt with immune research based on laboratory mice. The authors aimed at systematically manipulating text complexity particularly regarding cohesion (see Golke & Wittwer, 2018). In their study, Golke and Wittwer (2018) could show that the low-complex text contributed more to students' comprehension than the high-complex text, as an indicator of the prognostic validity of the text. As further treatment checks, that means that differences regarding students' (meta-) comprehension can be ascribed to textual differences of the learning materials, we realized two further safeguards. First, we conducted an automated linguistic complexity analysis to disentangle whether the low-complex and the high-complex texts descriptively differ on the central dimensions of linguistic text complexity by means of an established computer-linguistic analysis tool for German language (Berendes et al., 2018; Hancke et al., 2012). The findings indicated that the high-complex text was less cohesive and less concrete than the low-complex text (see Appendix A, Table A). However, both texts were comparable regarding syntactic complexity. Thus, we can conclude that the texts systematically differed regarding their overall text-complexity, at least on two of three dimensions of linguistic text complexity.

Second, we conducted an empirical validation study (N = 175 students) to investigate whether differences of the texts also resulted in perceived differences of cognitive load (i.e., perceived difficulty, mental effort). Additional goal was to test whether the findings were not confounded by other subjective text features such as perceived interestingness or simplicity. In line with our assumptions, we found that students reported to have invested lower mental effort and perceived lower levels of difficulty while reading the low-complex text than the high-complex text. Interestingness and perceived simplicity were rated to be comparable across the texts (see Appendix A for the complete pre-study).

8.3.2.2 Prior Knowledge Tests

We used the prior knowledge tests from Golke and Wittwer (2018) which contained one prior knowledge test per text. Each test included five open-ended questions (see Appendix B for examples) which measured different subtopics of reproduction and immunology, as they represent multidimensional constructs (McDonald's $\omega_{t\ \text{low-complex text}} = .63$; McDonald's $\omega_{t\ \text{high-complex text}} = .49$). For each question, students could receive one point, yielding a maximum score of five points per test. Two independent raters coded 20 % of the tests. As interrater reliability was excellent ($ICC_{\text{low-complex text}} = .97$; $ICC_{\text{high-complex text}} = .91$, see Koo & Li, 2016), one rater coded the remaining answers.

8.3.2.3 Knowledge Posttests

For assessing students' learning outcome, we also used the knowledge tests from Golke and Wittwer (2018). The knowledge posttests comprised six open text-based questions per text, which aimed at measuring students' basic knowledge of the text (see Appendix B for examples; McDonald's $\omega_{t\ \text{low-complex text}} = .67$; McDonald's $\omega_{t\ \text{high-complex text}} = .62$). Furthermore, six open inference-based questions per text measured students' advanced understanding (i.e., situational model, see Appendix B

; McDonald's $\omega_{t\ \text{low-complex text}} = .96$; McDonald's $\omega_{t\ \text{high-complex text}} = .63$), as the students had to combine different aspects and information across the text, and relate them to their prior knowledge. Again, students could receive one point per answer, yielding a maximum score of six points for each knowledge test. The posttests consisted of different questions than the prior knowledge test questions. Two independent raters rated 20 % of the tests. As interrater reliabilities for the text-based questions ($ICC_{\text{low-complex text}} = .93$; $ICC_{\text{high-complex text}} = .94$) and for the inference-based questions ($ICC_{\text{low-complex text}} = .98$; $ICC_{\text{high-complex text}} = .90$) were good to excellent (see Koo & Li, 2016), one rater coded the remaining answers.

8.3.2.4 Perceived Mental Effort and Subjective Difficulty

We asked students after the study phase and after the learning activity to rate their invested mental effort ("How much effort did you invest in explaining the material?", Paas, 1992), and subjective difficulty ("How easy was it for you to explain the material?", see DeLeeuw & Mayer, 2008) on a scale from one to nine, as subjective proxies to perceived cognitive load.

8.3.2.5 Meta-comprehension Accuracy

To investigate students' meta-comprehension accuracy, we asked them to make prospective judgments about their expected performance on the posttests ("How confident are you that you can answer questions to the text correctly?", Baars et al., 2017; Prinz et al., 2018) both after the study phase and after the learning activity on a scale from 0 % (no confidence) to 100 % (absolute confidence; see Schleinschok et al., 2017). We operationalized students' meta-comprehension accuracy in terms of bias (identically to Golke & Wittwer, 2018; Griffin et al., 2009; Lachner et al., 2020; Schraw, 2009; Wiley et al., 2016). We calculated the difference between the proportion of students' meta-comprehension ratings and the proportion of students' actual overall performance on the entire posttest (i.e., $X_{Judgment} – X_{Performance}$), separately for both texts. This approach allowed for measuring students' over- and underestimations of their judged test performance per text condition. Positive values indicate an overestimation and negative values indicate an underestimation of the judged performance. A value of zero indicates an accurate judgment. Thus, we calculated two different meta-comprehension accuracy scores (i.e., after the study phase, after the learning activity) per text to investigate whether students' meta-comprehension accuracy increased after the explaining task (see Golke & Wittwer, 2018).

8.3.2.6 Additional control measures.

Reading skills. As students' reading skills could impact their text comprehension (Golke & Wittwer, 2018; Ozuru et al., 2009), we assessed their reading skills by a German reading and speed comprehension test (LGVT 6-12; Schneider et al., 2007). The LGVT is a speed test that requires students to read a text (text length: 1700 words), as quickly and accurately as possible (duration 4 minutes). The text comprises 25 gaps in which students are required to choose which of three different alternative words had to be filled in (for more details see Schneider et al., 2007). Cronbach's alpha was good: $\alpha = .81$.

Topic interest. As a further control variable, we measured students' interest in biology with three items (i.e., "I regard biological topics as important", "I am excited about biological topics", "Biological topics are fascinating", based on Baumert et al., 2008) on a four-point Likert scale from 1 ("I completely disagree") to 4 ("I completely agree"). Cronbach's alpha was excellent: $\alpha = .92$.

Self-efficacy in explaining. As a further control variable, we assessed students' self-efficacy in explaining. We used three adapted items (e.g., "I always find ways to explain even difficult contents") from Jerusalem and Schwarzer (1999), and asked students to rate their explaining skills on a four-point Likert scale from 1 ("I completely disagree") to 4 ("I completely agree"). Cronbach's alpha was good: $\alpha = .79$.

8.3.3 Procedure

The instructor informed the students about the study scope. After providing written consent, the students were randomly assigned to the learning activity conditions (i.e., oral explanation, written explanation, retrieval practice). Afterwards, all students were seated individually in front of a laptop in multi-group sessions (n = 4) in two separate rooms. To avoid potential disturbances by other participating students, they wore sound-proof ear protectors during the experiment. The entire study was self-paced. First, the students answered the prior knowledge test, and rated their explaining skills. In the study phase, students read the first text in a self-paced manner, assessed their invested mental effort and perceived subjective difficulty during reading, and provided a meta-comprehension judgment. In the explaining phase, dependent on learning activity, they randomly provided an oral or a written explanation or retrieved the information of the text. The oral explanation group recorded their explanation via the built-in microphone function of the Windows system. The written explanation group wrote an explanation in the Qualtrics environment. Both explaining groups received the following instruction:

Imagine the following scenario: One person could not participate in the study. However, she is highly interested in the topic. She has not read anything about the topic yet. Therefore, she asks you to explain the central contents of the text. Please provide her a clear and detailed explanation, so that she can understand the content without additional information.

The control group was engaged in a retrieval practice:

Please retrieve the information of the text.

For the learning activity, we followed a more conservative approach and kept time on task constant (i.e., 10 minutes) across conditions, as recent research indicated that particularly for generative activities, such as explaining, their effectiveness may be confounded by differences of students' time investment (see Hoogerheide, Staal, et al., 2019). Additionally, during the learning activity, all the students could take notes on a separate sheet (see also Fiorella & Kuhlmann, 2019; Lachner et al., 2018, for related approaches). They were informed that the notes were only available in the learning activity phase but not in the testing phase. Then, the students again provided judgments of their mental effort and subjective difficulty, as well as a meta-comprehension judgment. Afterwards, the students answered the knowledge posttest. Given that students processed two different tests, the procedure was repeated analogously for the second text. The order of the texts was randomized across students. At the end of the study, the students completed the reading skill test and answered the demographic questionnaire (i.e., age, gender, mother language, number of enrolled semesters, interest in topic). The average study time was 90 minutes. After completing the study, the students were debriefed and rewarded with 16 Euros.

8.3.4 Analysis and Coding of the Explanations

8.3.4.1 Number of Personal References

We counted the number of personal references per explanation (i.e., “I”, “you”, etc.) as a behavioral proxy for the perceived social presence (see Chafe & Tannen, 1987; Hoogerheide et al., 2016; Lachner et al., 2018, for related approaches). This procedure was further motivated, as recent studies demonstrated distinct relations between the number of personal references and subjective ratings of perceived social presence in explaining contexts (see Jacob, Lachner, & Scheiter, 2020). Two independent raters counted the number of personal references for 20 % of all explanations. Interrater reliability was excellent, *ICC* = 1.00. Therefore, one rater coded the remaining explanations.

8.3.4.2 Comprehensiveness

As an indicator of the level of comprehensiveness, we counted the number of concepts per explanation (Boshuizen & Schmidt, 1992). To automatically detect the concepts per explanation, we used *CohViz* by Burkhart, Lachner, and Nückles (2017b) which automatically detects and counts the number of concepts within an explanation by means of natural language processing technologies. For instance, the sentence “After both *groups* got a *vaccination* against *fever*, the *group* with previous *infections* showed a different *reaction*” contains five concepts (marked in italics). Redundant concepts (e.g., “group”) were ignored (see also Boshuizen & Schmidt, 1992, for related approaches). CohViz showed a high precision in comparison to expert raters (*ICC* = .93, see Lachner et al., 2017b).

8.3.4.3 Number of Elaborations

As an indicator of the level of elaboration of the explanations, we counted the number of elaborations per explanation. An elaboration was operationalized as an idea unit which was not covered in the explanation, such as examples, analogies, and own experiences (see also Fiorella & Kuhlmann, 2020; Lachner et al., 2018). For instance, the sentence "Mice from the pet shop have a completely different immune profile than laboratory mice, similar to the profile of human adults who have already had many diseases, *and therefore, have already formed more immune cells*" contains one elaboration (marked in italics) since the section was not covered in the study text. Again, two independent raters counted the number of elaborations for 20 % of the explanations. Interrater reliability was excellent (*ICC* > .95). Thus, one rater coded the remaining explanations.

8.4 Results

We used partial η_p^2 and Cohens' *d* as effect size measures, qualifying values of η_p^2 = .01, .06, .14, and *d* = 0.20, 0.50, 0.80 as small, medium, and large effects (Cohen, 1977). Additionally, we used an alpha level of α = .05. In cases of directed hypotheses, we used one-tailed tests (Cho & Abe, 2013; Furr & Rosenthal, 2003).

8.4.1 Preliminary Analyses

Analyses showed no significant differences between the learning activity groups concerning age, $F(2, 112) = 1.54$, $p = .220$, $\eta_p^2 = .027$, and their number of enrolled semesters,[2] $F(2, 106) = 0.71$, $p = .492$, $\eta_p^2 = .013$.[3] Students' cognitive and motivational prerequisites were comparable among groups, as the learning activity groups did not differ in their reading skills, $F(2, 112) = 0.27$, $p = .767$, $\eta_p^2 = .005$, in their prior knowledge of the low-complex text, $F(2, 112) = 0.71$, $p = .496$, $\eta_p^2 = .012$, or their prior knowledge of the high-complex text, $F(2, 112) = 1.26$, $p = .288$, $\eta_p^2 = .022$. Similarly, there were no differences regarding students' general topic interest, $F(2, 112) = 1.09$, $p = .340$, $\eta_p^2 = .019$, or their self-efficacy in explaining, $F(2, 112) = 1.19$, $p = .310$, $\eta_p^2 = .021$. Additionally, the reading times of the texts were comparable across conditions (*F*s < 1), indicating that all learning activity groups invested the same amount of time to process the texts. Regarding the subjective ratings

[2] Eleven students did not state their number of enrolled semesters.
[3] Two students did not report their number of semesters.

of cognitive load[4] of the study phase (i.e., mental effort, subjective difficulty), there was no difference among learning activity groups for the mental effort ratings, $F(2, 110) = 2.87$, $p = .061$, $\eta_p^2 = .039$, or for the subjective difficulty, $F(2, 110) = 1.45$ $p = .240$, $\eta_p^2 = .018$. Similar to the pre-study, students stated to invest more mental effort in reading the high-complex text ($M = 6.33$, $SD = 1.82$) compared to reading the low-complex text ($M = 5.96$, $SD = 1.92$), $t(112) = -2.78$, $p = .025$, $d = 0.20$. However, this was not directly reflected in students' perceived difficulty, as students rated both texts as comparably difficult ($M_{\text{low-complex text}} = 4.13$, $SD_{\text{low-complex text}} = 2.22$; $M_{\text{high-complex text}} = 4.37$, $SD_{\text{high-complex text}} = 1.96$), $t(112) = -1.09$, $p = .279$, $d = 0.11$.

As a further control, we tested whether our data was confounded by potential outliers. Graphical boxplot analyses of the dependent variables revealed that there were three outliers for the text-based questions (see Figure C.1 and C.3 in Appendix C). The outliers were included in the main analysis, as removing them from the sample did not change the findings. Correlations between prior knowledge, the dependent variables, students' subjective ratings of cognitive load, their judgement of learning, and their reading skills are shown in Appendix D.

8.4.2 Learning-Outcome Hypotheses

We performed two separate mixed factorial ANCOVAs for our two knowledge tests as dependent variables (i.e., text-based questions; inference-based questions). Learning activity (i.e., generating oral vs. written explanations vs. retrieval practice) was the between-students factor and linguistic text complexity (i.e., low-complex text vs. high-complex text) was the within-students factor. Additionally, we controlled for students' prior knowledge as covariate.

Regarding the text-based questions, as expected, we found a main effect of linguistic text complexity, $F(1, 111) = 56.51$, $p < .001$, $\eta_p^2 = .128$, indicating that students performed better in the low-complex text condition than the high-complex text condition (see Table 2). In line with our hypotheses, there was no main effect of learning activity, $F(1, 111) = 0.78$, $p = .459$, $\eta_p^2 = .007$, and no interaction effect between linguistic text complexity and learning activity, $F(2, 111) = 0.58$, $p = .563$, $\eta_p^2 = .003$, indicating that the learning activity groups performed comparable on the text-based questions. The same results occurred after controlling for text order, $F(2, 111) = 0.58$, $p = .563$, $\eta_p^2 = .003$, note taking, $F(2, 110) = 0.57$, $p = .566$, $\eta_p^2 = .003$, and reading time, $F(2, 110) = 0.63$, $p = .537$, $\eta_p^2 = .003$.

[4] Two students did not report their cognitive load.

Regarding the inference-based questions, we found a significant effect of linguistic text complexity, $F(1, 111) = 15.85$, $p < .001$, $\eta_p^2 = .039$, indicating that students scored higher in the low-complex text condition than in the high-complex text condition (see Table 2). Again, the main effect of learning activity was not significant, $F(1, 111) = 1.34$, $p = .261$, $\eta_p^2 = .013$. In line with our hypothesis, we found a significant interaction between learning activity and linguistic text complexity, $F(2, 111) = 4.67$, $p = .011$, $\eta_p^2 = .023$. The same results occurred after controlling for text order, $F(2, 110) = 4.67$, $p = .011$, $\eta_p^2 = .023$, note taking, $F(2, 110) = 4.52$, $p = .013$, $\eta_p^2 = .022$, and reading time, $F(2, 110) = 4.55$, $p = .013$, $\eta_p^2 = .023$. There was no interaction of prior knowledge, reading skills, and text condition, $F(1, 97) = 3.43$, $p = .067$, $\eta_p^2 = .009$.

To break up the significant interaction between linguistic text complexity and learning activity, we performed planned contrasts to specifically test our hypotheses. With the first contrast, we tested whether explaining was more beneficial than retrieval practice (i.e., retrieval practice: -2; written explanation: 1; oral explanation: 1). With the second contrast, we tested whether generating oral explanations was more effective than written explanations (i.e., retrieval practice: 0; written explanation: -1; oral explanation: 1). For the high difficult text condition, the contrast analysis revealed that students who generated explanations (oral and written form) outperformed students who simply retrieved the learning material, $t(111) = 1.90$, $p = .030$, $d = 0.36$ (see Table 2). More importantly, students who explained orally outperformed students who explained in written form, $t(111) = 1.81$, $p = .037$, $d = 0.32$. For the low-complex text condition, none of the contrasts approached significance (explaining contrast: $t(111) = -1.25$, $p = .785$, $d = 0.33$; modality contrast: $t(111) = 0.80$, $p = .576$, $d = 0.22$). Together, in line with our hypotheses, we found that the modality of explaining only had an effect for high-complex text material but not for low-complex text material. This effect, however, was only significant for higher-order test questions (i.e., inference-based questions), but not for lower-order test questions (i.e., text-based questions).

Table 2

Means and Standard Deviations for all Measurements

	Low-complex text			High-complex text		
	Retrieval practice	Written explaining	Oral explaining	Retrieval practice	Written explaining	Oral explaining
Prior knowledge (0 – 5)	2.92 (1.12)	2.59 (1.25)	2.76 (1.20)	2.25 (0.96)	2.32 (1.22)	1.96 (0.98)
Learning outcomes (0 – 6)						
Text-based questions	4.21 (1.28)	4.01 (1.22)	4.34 (1.14)	3.39 (1.10)	3.15 (1.13)	3.30 (1.28)
Inference-based questions	3.28 (1.03)	2.79 (1.12)	3.05 (1.21)	2.28 (1.08)	2.50 (1.17)	2.89 (1.25)
Perceived cognitive load during the experimental task (1 – 9)						
Perceived mental effort	5.42 (1.98)	6.08 (2.01)	6.55 (1.52)	5.64 (2.17)	6.38 (1.76)	6.83 (1.24)
Perceived difficulty	3.72 (2.05)	4.54 (2.08)	5.28 (1.96)	4.06 (2.03)	4.51 (2.01)	5.12 (1.73)
Meta-compr. accuracy	0.13 (1.50)	0.24 (1.38)	-0.46 (1.33)	0.54 (1.42)	0.90 (1.34)	0.60 (1.50)
Reading time (in minutes)	2.95 (2.75)	3.07 (2.07)	3.19 (1.35)	3.94 (2.21)	3.65 (2.07)	3.73 (1.58)
Characteristics of explanations						
Personal references		1.54 (1.89)	2.21 (3.33)		0.10 (0.45)	0.64 (1.46)
Concepts		21.31 (7.57)	23.08 (6.15)		18.95 (7.63)	25.90 (9.12)
Elaborations		2.46 (1.68)	2.41 (1.53)		1.00 (0.86)	1.39 (1.46)

Note. Meta-comprehension accuracy: A score of 0 indicates an exact estimation of learning; negative numbers indicate an underestimation of learning; positive numbers indicate an overestimation of learning. Characteristics of explanations: Personal references and elaborations were counted by two independent raters. Amounts of concepts were automatically analyzed by tool CohViz (see Lachner et al., 2017b).

8.4.3 Additional Explorative Analyses

8.4.3.1 Effects of Explanatory Features

To explore the processes underlying the significant modality effect in the high-complex text condition, we analyzed potential differences regarding the number of personal references, the number of concepts, and the number of elaborations by applying independent *t*-tests. Oral explanations contained more personal references (e.g., “I”, “you”, etc.) than written explanations, $t(77) = 2.20$, $p = .033$, $d = 0.50$, indicating that the social presence was higher in the oral condition than in the written condition (see Table 2). Regarding the generative processes, oral explanations contained more concepts, as an indicator of the level of comprehensiveness, than written explanations, $t(77) = 3.65$, $p < .001$, $d = 0.83$. However, we did not find a significant difference between written explanations and oral explanations regarding the number of elaborations, $t(77) = 1.44$, $p = .155$, $d = 0.33$. The explorative analysis of the differences of the explanations suggested that the effect of generating oral versus written explanations on students’ performance on the inference-based questions mainly resulted, as oral explaining triggered higher amounts of social presence (indicated by the number of personal references) which resulted in more comprehensive explanations (indicated by the number of concepts). To explore this mediation assumption, we conducted a serial mediation analysis. The number of personal references and the number of concepts were the serial mediators, learning activity was the dummy-coded predictor (1 = written explanation, 2 = oral explanation), and students’ performance on the inference-based questions (in the high-complex text) was the dependent variable. We applied a bootstrapping methodology based on Hayes (2013) with 10,000 simulations. Results indicated a significant indirect effect (via number of personal references → number of concepts) of $a_1 \times d_{21} \times b_2 = .09$, 95% CI [.010, .313], as zero was not included in the confidence interval (see Figure 12, for the full mediation model). There were no direct effects via personal references or comprehensiveness. The significant serial mediation effect suggests that the superiority of generating oral versus written explanations resulted, as students had higher levels of social presence during oral explaining (as indicated by the number of personal references), which inclined them to provide more comprehensive explanations and include more concepts in their explanations.

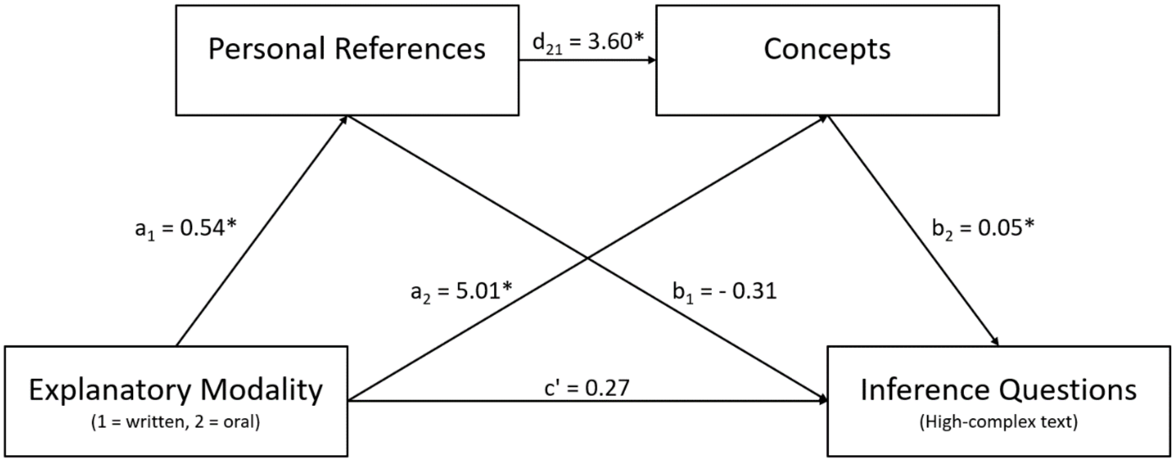

Figure 12. Serial mediation analysis regarding concepts in explanations based on the high-complex text. **p* < .05.

8.4.3.2 Effects on Students' Mental Effort

To investigate students' mental effort during the learning activity (i.e., explaining or retrieving), we conducted a repeated measures ANOVA[5] with students' mental effort ratings as a dependent variable, learning activity (i.e., generating oral vs. written explanations vs. retrieval practice) as a between-students factor, and linguistic text complexity (i.e., low-complex text vs. high-complex text) as a within-students factor. Results showed no main effect of linguistic text complexity, $F(1, 110) = 2.21$, $p = .140$, $\eta_p^2 = .004$, indicating that processing the texts after administering the subsequent learning activities (i.e., generating oral vs. written explanations vs. retrieval practice) was perceived as comparably effortful across conditions of linguistic text complexity. However, we found a main effect of learning activity, $F(2, 110) = 5.15$, $p = .007$, $\eta_p^2 = .069$. The interaction between learning activity and linguistic text complexity was not significant, $F(2, 110) = .080$, $p = .923$, $\eta_p^2 < .001$. Post-hoc comparisons (Scheffe) indicated that students reported more effort during oral explaining than retrieving the contents ($p < .001$, see

[5] Two students did not report their cognitive load.

Table 2). There were no significant differences between generating oral or written explanations (p = .251) or between writing explanations and retrieving the contents (p = .057).

8.4.3.3 Effects on Students' Subjective Difficulty

Similar to previous analyses, a repeated measures ANOVA[6] revealed no significant main effect of linguistic text complexity, $F(1, 110) = 0.12$, $p = .733$, $\eta_p^2 < .001$, but a main effect of the learning activity, $F(2, 110) = 7.00$, $p = .001$, $\eta_p^2 = .080$. Again, the interaction effect was not significant, $F(2, 110) = 0.38$, $p = .684$, $\eta_p^2 < .001$.

Similar to the mental effort ratings, post-hoc tests revealed that students perceived oral explaining as more difficult than retrieving the contents ($p < .001$, see Table 2). Again, there were no differences between generating oral versus written explanations (p = .050) or between writing explanations and retrieving the contents (p = .140).

8.4.3.4 Meta-comprehension accuracy

To investigate whether students' meta-comprehension accuracy increased after the explaining task, we conducted an ANCOVA with students' meta-comprehension accuracy after the learning activity as a dependent variable, the low- and the high-complex text as a within-students factor, and type of learning activity (i.e., generating oral vs. written explanation vs. retrieval practice) as a between-students factor. Additionally, we controlled for students' meta-comprehension accuracy during the study phase to control for potential intra-individual differences (see Hertzog et al., 2013; Lachner et al., 2020). Results demonstrated a main effect of learning activity, $F(2, 111) = 9.05$, $p < .001$, $\eta_p^2 = .032$, and a main effect of linguistic text complexity, $F(1, 111) = 41.11$, $p < .001$, $\eta_p^2 = .039$, indicating that students rated their comprehension more precisely in the low-complex text condition than in the high-complex text condition (see Table 2). There was no interaction between learning activity and linguistic text complexity, $F(2, 111) = 0.04$, $p = .963$, $\eta_p^2 = .000$. Post-hoc tests showed a significant difference between students who explained orally and students who wrote an explanation ($p < .001$). Students who explained orally estimated their comprehension more accurately (see Table 2). Additionally, there was a significant difference between students who generated an oral explanation and students who retrieved the material (p = .026). Again, explaining orally resulted

[6] Two students did not report their cognitive load.

in better meta-comprehension accuracy. However, results indicated no effect between students who wrote an explanation and students who retrieved the material (p = .446).

8.5 Discussion

In the current study we investigated whether the effect of explaining modality depends on the level of linguistic text complexity of the studying material. As expected, we showed that explaining modality mainly affected students' performance when they learned from high-complex learning material, as students who generated oral explanations outperformed students who generated written explanations (on the inference-based questions). Students who simply did retrieval practice showed the lowest performance. For the low-complex learning material, the three learning activities were comparably effective since there were no significant differences among learning activities. Thus, linguistic text complexity moderated the effect of explaining modality on students' learning. Whereas low-complex learning material may suffice to support students in constructing a coherent representation of the text (Berendes et al., 2018; McNamara, 2013), which may make subsequent learning activities obsolete, for the high-complex materials, the addition of a generative learning activity was a necessary precondition to help students establish a coherent understanding of the text (see also Roelle & Nückles, 2019, for related findings). This effect was more pronounced for explaining orally than for writing explanations. Therefore, our findings highlight that the need for additional learning activities depends on the complexity of the pre-given study materials.

Second, we showed that the superior effects of generating oral versus written explanations for complex text materials occurred, as oral explaining likely triggered higher levels of social presence during explaining, as indicated by behavioral proxies of the number of personal references. The increased social presence resulted in more generative processes, as indicated by more comprehensive explanations in the oral explaining condition. Therefore, our findings corroborate the social presence hypothesis (Hoogerheide et al., 2016; Lachner et al., 2018) that effects of explaining are not only due to generative processes but also due to increases of social presence. Our mediation analysis helped get a better understanding of the cognitive mechanisms of increases of social presence. For instance, Hoogerheide et al. (2019) suggested both, a direct effect of social presence, due to increases of arousal, and an indirect effect of social presence by affecting the generative processes that take place during explaining. Our findings rather support the indirect social presence hypothesis, as we only obtained a significant serial mediation effect of the explaining modality on students' learning via social presence (i.e., personal references) and comprehensiveness of explanations (i.e., number of concepts) but not a direct effect solely via social presence.

Our explorative analysis regarding students' meta-comprehension accuracy showed a further interesting detail. Oral explaining generally resulted in more accurate judgements compared to retrieval practice. These results are in line with previous research which demonstrated that oral explaining supports students' meta-comprehension accuracy (Fukaya, 2013; Lachner et al., 2020). Additionally, explaining orally resulted in more accurate judgements compared to writing explanations. This result adds to previous research, as it shows that improved meta-comprehension accuracy depends on the modality of explaining.

An unexpected finding was that only the comprehensiveness of the explanation but not the elaboration of the explanations mediated the effect of explaining modality on students' learning outcomes in the high-complex text condition. This finding is surprising, as recent studies documented the central role of elaborations during explaining (Fiorella & Kuhlmann, 2020; Lachner et al., 2018). We attribute this finding to the fact that students processed relatively abstract biological contents, which made it difficult to generate elaborations such as examples during explaining. This was also reflected in the medium number of elaborations in the explaining conditions (see Table 2).

8.5.1 Study Limitations and Future Research

One potential caveat refers to the fact that we only measured characteristics of explanations as a proxy for the underlying processes of the modality effect (see Fiorella & Kuhlmann, 2020; Lachner et al., 2018). First, based on prior research, we measured three characteristics of students' explanations (i.e., personal reference, comprehensiveness, and elaborations). However, further research is needed to investigate additional characteristics which could result in different learning outcomes. Second, assessing students' products of explaining (i.e., comprehensiveness and elaborations) – more importantly for the written explanation condition — does not provide an "online" measure of the underlying processes of the explaining modality effect. For instance, it could be the case that, particularly in the written explanation condition, students could have elicited more elaborations than they actually realized in their explanations. Such an effect can be expected, as writing typically involves more conscious decisions about inserting or withholding information than providing oral explanations (Lachner et al., 2018; Nückles et al., 2009). Thus, future studies should include online-measures, such as think-aloud protocols (Ericsson & Simon, 1999) or log-file-data to more directly assess the underlying cognitive mechanisms of explaining.

Another caveat refers to the relatively low homogeneity indices, particularly for the prior knowledge tests. However, we want to note that we measured a broad prior understanding by asking different and independent questions with a restricted set of 5 items per topic (e.g., “How does offspring arise from asexual reproduction?”, “How do children’s genes differ from their parents’ genes?”, “How do mammals select a partner for reproduction?”). This decision likely resulted in a trade-off regarding the internal consistency of the prior knowledge measures, as they likely covered diverse sub-components of students’ prior knowledge (see Dunn, Baguley, & Brunsden, 2014). Further research with advanced statistical procedures is needed which explicitly takes the multidimensionality into account, for instance, by using bifactor-(S-1) measurement models, which however, require larger sample sizes (Eid, Geiser, Koch, & Heene, 2017).

Furthermore, we have to note that we conducted our study in a laboratory setting with university students. Thus, although this study indicates that the exploratory modality effect depends on the complexity of the learning material, the external validity is limited. Therefore, more research is needed in applied contexts and different study populations, for instance in school settings, to test the robustness of our findings (see Hoogerheide, Visee, et al., 2019, for examples), and to provide guidelines for educational practice.

Additionally, we would like to point out that our findings are restricted to reading online-texts only, as we provided our students with text materials on the laptop. Replicating our findings, particularly regarding potential interaction effects of linguistic text complexity and explaining modality would be interesting as current research showed potential advantages of paper-based versus online-based text materials regarding students’ comprehension materials (Kong et al., 2018; Singer Trakhman et al., 2018). Therefore, it is up to further research to investigate whether our findings would replicate in learning settings in which students learn from paper-based text materials.

Finally, we want to draw attention to the fact that we only implemented one learning strategy (i.e., explaining to fictitious other students) to support students' (meta-) cognitive processing. Given that recent research has started to use combinations of cognitive and metacognitive instructional support strategies (Fiorella et al., 2017; Fiorella & Kuhlmann, 2020; Rogiers et al., 2019; Weinstein et al., 2018), it would be interesting to investigate potential synergistic but also moderating effects in the context of learning by explaining.

8.5.2 Conclusion

Our findings demonstrate that explaining to fictitious others is a beneficial strategy to support students' learning, particularly when they are required to learn from difficult learning materials. Furthermore, our findings suggest that the effectiveness of explaining depends on the modality of explaining. Due to differences in perceived social presence between written and oral explaining, providing written and oral explanations might be both appropriate if the primary learning goal is the acquisition of basic knowledge. However, if the main goal is to enable students to construct coherent higher-order knowledge, generating oral explanations may best enhance students' learning.

8.6 Appendix

8.6.1 Appendix A

We empirically investigated the linguistic differences among text-conditions in a pre-study as a further manipulation check. We invited university students (who did not participate in the main study) via the university mailing list to judge the linguistic text complexity of the two texts by means of an online questionnaire. The findings are based on a sample of 175 students (42 students were excluded, as they did not complete the pre-study or German was not their mother language). The students were comparable to the main study (mean age: 22.65 years, *SD* = 3.66, 79% female). Students randomly read the high- or the low-complex text, and judged the subjective difficult, and their invested mental effort (see Material section of main study) on a scale from one to nine, as subjective proxies to perceived cognitive load. Additionally, the students rated general characteristics of the text by means of a questionnaire of Jucks (2001) comprising the following dimensions: a) simplicity (5 items, for instance, "The text contains short and easy sentences”, Cronbach’s $\alpha_{\text{low-complex text}}$ = .56, Cronbach’s $\alpha_{\text{high-complex text}}$ = .67); b) cohesion (6 items, for instance, “The text has a consistent story line”, Cronbach’s $\alpha_{\text{low-complex text}}$ = .84, Cronbach’s $\alpha_{\text{high-complex text}}$ = .87); c) conciseness (6 items, for instance, “The text is limited to the essential information”, Cronbach’s $\alpha_{\text{low-complex text}}$ = .46, Cronbach’s $\alpha_{\text{high-complex text}}$ = .61); d) interestingness (4 items, for instance, “The text is interesting”, Cronbach’s $\alpha_{\text{low-complex text}}$ = .84, Cronbach’s $\alpha_{\text{high-complex text}}$ = .78). Then, the students read the second text and answered identical questions.

In line with the linguistic analysis, students invested more mental effort while reading the high-complex text, $t(174) = 2.60$, $p = .010$, $d = 0.15$. Additionally, they stated higher difficulties when reading the high-complex text compared to the low-complex text, $t(174) = 3.17$, $p = .002$, $d = 0.31$. However, we did not find significant differences between the low-complex and the high-complex text for perceived simplicity, $t(174) = 0.14$, $p = .890$, $d = 0.01$, perceived text cohesion, $t(174) = -1.08$, $p = .281$, $d = -0.11$, and interestingness of the texts, $t(174) = -.07$, $p = .287$, $d = 0.11$, indicating that apparently students were not directly aware of the linguistic differences between both texts (see descriptive statistics in Table A). Interestingly, students rated the high-complex text as less concise than the low-complex text, $t(174) = -3.36$, $p = .001$, $d = -0.29$, although both texts had a comparable text length (words $_{\text{low-complex text}}$: 380, words $_{\text{high-complex text}}$: 397). These findings may be indicative of the fact that students had to invest more mental effort during reading the high-complex text which resulted in different perceptions of text length (i.e., conciseness).

Table A

Features of Both Learning Texts

Features	Low-complex text	High-complex text
Topic	Reproduction	Immunology
Genre	Informative factual text	Informative factual text
Linguistic features		
Number of words	380	397
Number of sentences	25	26
Words per sentence	15.2	15.3
Syntactical structure	9.12	9.79
Global cohesion	0.16	0.08

Word concreteness	15.5	9.5
Flesh-Index	46	41
Perception of the text (Pre-Study)		
Mental effort	4.72 (*SD* = 1.91)	5.01 (*SD* = 1.94)
Subjective difficulty	2.92 (*SD* = 1.78)	3.48 (*SD* = 1.84)
Simplicity	3.56 (*SD* = 0.61)	3.57 (*SD* = 0.70)
Cohesion	3.70 (*SD* = 0.72)	3.61 (*SD* = 0.80)
Conciseness	2.97 (*SD* = 0.53)	2.81 (*SD* = 0.55)
Interestingness	3.41 (*SD* = 0.88)	3.50 (*SD* = 0.83)

Note. Linguistic features were calculated according to McNamara (2013), Berendes et al. (2018), and Lachner and Nückles (2015). Perception of the text was rated by 175 students (see Materials section).

8.6.2 Appendix B

Item Examples of Knowledge Tests

	Low-complex text	High-complex text
Prior knowledge questions	How does offspring arise from asexual reproduction?	Why can viruses be dangerous for humans?
Text-based questions	What is the disadvantage of sexual reproduction?	What is the main concern against research with laboratory mice?
Inference-based questions	How would the health of animals be affected if they produced offspring with asexual reproduction?	How can the results help to clarify the basic problems of immune research with mice?

8.6.3 Appendix C

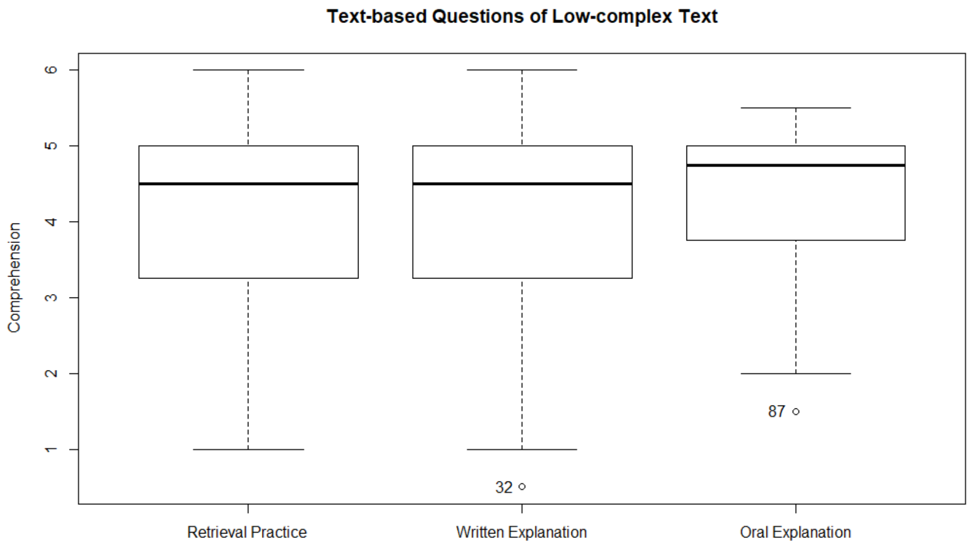

Figure D.1. Boxplots of Posttest Performance per Condition.

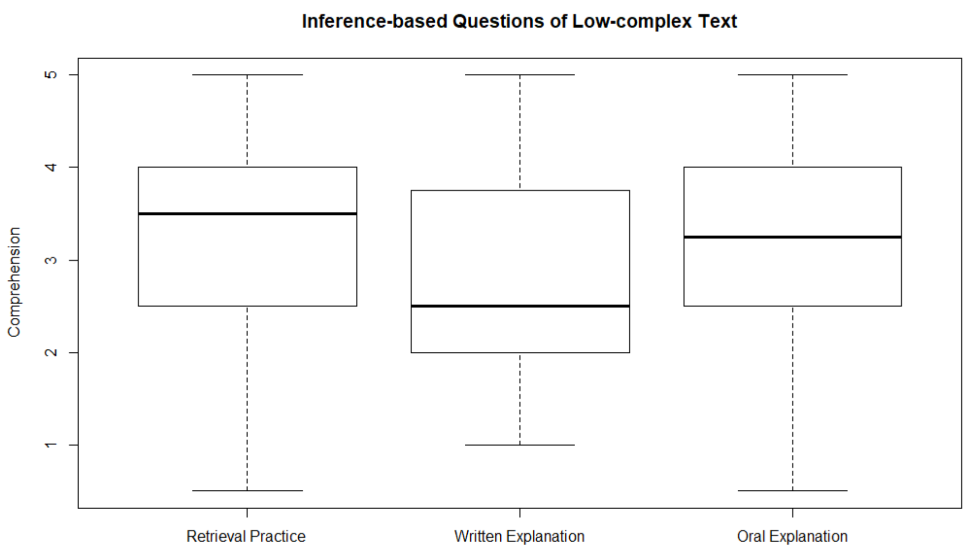

Figure D.2. Boxplots of Posttest Performance per Condition.

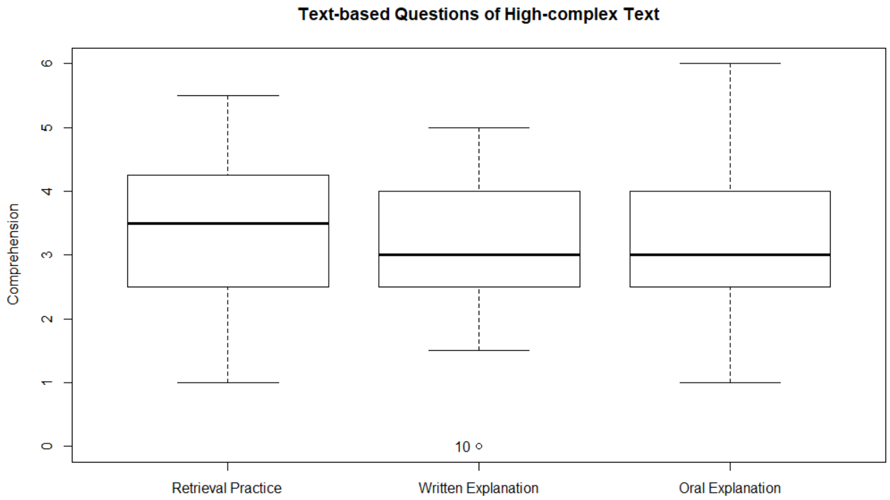

Figure D.3. Boxplots of Posttest Performance per Condition.

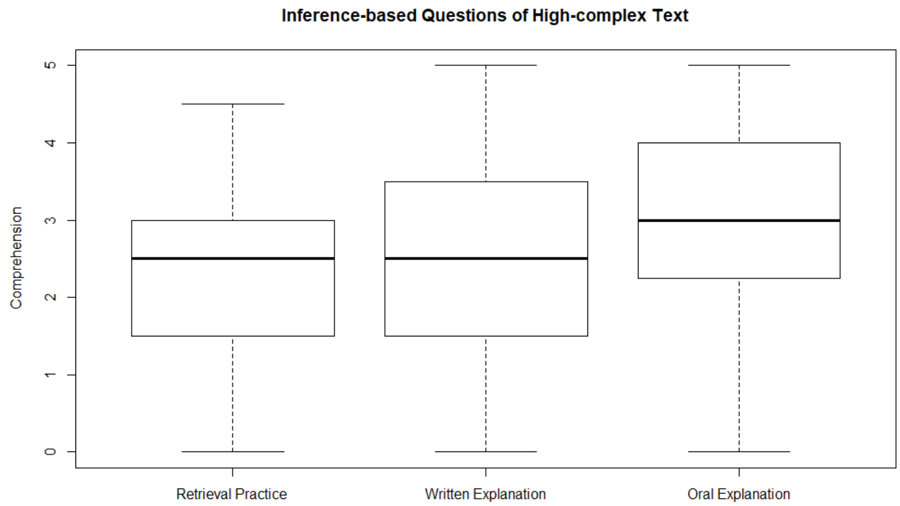

Figure D.4. Boxplots of Posttest Performance per Condition.

8.6.4 Appendix D

Correlations Separated by Experimental Condition and Text Condition

Low-complex text	1	2	3	4	5	6	7
1 Prior knowledge	-						
2 Text-based qu.	(.24|.27|.39*)	-					
3 Inference-based qu.	(.40*|.28|.60*)	(.41*|.29|.43*)	-				
4 Perceived mental effort	(-.16|.14|.09)	(-.13|.05|.05)	(.03 |.12|.12)	-			
5 Perceived difficulty	(-.02|-.29|-.25)	(-.19|-.08|-.30)	(-.27|-.07|-.37*)	(.33*|.28|-.07)	-		
6 Meta-compr. accuracy	(-.10|.03|.02)	(-.49*|-.33*|-.23)	(-.53*|-.32|-.08)	(-.13|-.15|.35*)	(-.09|-.50*|-.17)	-	
7 Reading skills	(.22|-.03|.16)	(-.21|.30|.17)	(.06|.05|.02)	(.04|-.10|.04)	(-.21|.17|.12)	(.17|-.14|.11)	-
High-complex text	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1 Prior knowledge	-						
2 Text-based qu.	(.19|-.04|.36*)	-					
3 Inference-based qu.	(.28|.14|.26)	(.18|.46*|.62*)	-				
4 Perceived mental effort	(-.29|.05|.17)	(-.03|-.01|.03)	(.02|.29|.06)	-			
5 Perceived difficulty	(-.24|-.00|-.04)	(-.10|-.38*|.07)	(-.00|-.18|.13)	(.40*|.47*|-.13)	-		
6 Meta-compr. accuracy	(-.12|.08|-.11)	(-.48*|-.47*|-.67*)	(-.53*|-.54*|-.61*)	(-.12|-.06|.03)	(-.27|-.08|-.30)	-	
7 Reading skills	(-.05|-.02|-.11)	(-.09|.15|-.17)	(.32|.30|-.15)	(.13|.11|.14)	(.10|-.08|.09)	(-.19|-.18|.18)	-

Note. Numbers in brackets represent the correlations separated for experimental conditions: Left = retrieval condition; middle = writing condition; right = oral condition. * indicates $p < .05$.

9

STUDY 2: DOES INCREASING SOCIAL PRESENCE ENHANCE THE EFFECTIVENESS OF WRITING EXPLANATIONS?

9. Study 2: Does Increasing Social Presence Enhance the Effectiveness of Writing Explanations?

9.1 Abstract

Writing explanations has demonstrated to be less effective than providing oral explanations, as writing triggers less amounts of perceived social presence during explaining. In this study, we investigated whether increasing social presence during writing explanations would aid learning. University students (N = 137) read an instructional text about immunology; their subsequent task depended on experimental condition. Students either explained the contents to a fictitious peer orally, wrote their explanations in a text editor, or wrote them in a messenger chat, which was assumed to induce higher levels of social presence. A control group retrieved the material. Surprisingly, we did not obtain any differences in learning outcomes between experimental conditions. Interestingly, explaining was more effortful, enjoyable, and interesting than retrieving. This study shows that solely inducing social presence does not improve learning from writing explanations. More importantly, the findings underscore the importance of cognitive and motivational conditions during learning activities.

Published as: Jacob L., Lachner A., & Scheiter K. (2021). Does increasing social presence enhance the effectiveness of writing explanations? PLOS ONE, 16(4): e0250406. https://doi.org/10.1371/journal.pone.0250406

9.2 Theoretical Background

Generating explanations is regarded as a successful strategy to enhance students' understanding, as it triggers generative processes associated with deep learning (Fiorella & Kuhlmann, 2020; Fiorella & Mayer, 2014; Fukaya, 2013; Lachner et al., 2020; Palinscar & Brown, 1984; Plötzner et al., 1999; Roscoe, 2014; Roscoe & Chi, 2008). Seminal studies on learning by explaining started to investigate the role of explaining in interactive settings, such as during collaborative learning or tutoring, in which the explainer received feedback from the recipient, for instance, in form of direct questions (Plötzner et al., 1999; Roscoe & Chi, 2008). Results showed that students who explained learned material were engaged in deeper learning processes and showed higher learning outcomes compared to restudying (Palinscar & Brown, 1984; Plötzner et al., 1999; Roscoe, 2014; Roscoe & Chi, 2008). Recent research replicated the beneficial of explaining in non-interactive settings, in which no recipient was present; thus, students explained learned material to a *fictitious* peer which also resulted in higher learning outcomes (Fiorella & Mayer, 2014; Hoogerheide, Renkl, et al., 2019; Kobayashi, 2019). Moreover, learning by explaining was more often shown to be effective when students were required to generate oral explanations instead of written ones (Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018). A possible reason for the benefit of oral explaining might be the difference between perceived *social presence* during explaining. In this context, Jacob, Lachner, and Scheiter (2020) provided first evidence that writing explanations induces lower levels of social presence during explaining than providing oral explanations, and reduces the quality of the generated explanations (Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018). Social presence is a central concept in discourse theory, and is commonly defined as the extent to which a person feels that a communication partner is present during a mediated conversation, such as in online learning environments (Gunawardena & Zittle, 1997; Oh et al., 2018; Short et al., 1976; Weidlich & Bastiaens, 2017, 2019). In this context, social presence not only holds true for real persons, but also for virtual or even fictitious communication partners (Atkinson, 2002; Kim, 2013; Wang & Antonenko, 2017). Given that writing explanations is a learning activity that can easily be implemented in learning contexts, the question arises whether and how writing explanations can be made more effective. As previous studies highlighted the role of social presence, we investigated whether inducing social presence would increase the effectiveness of writing explanations regarding students' comprehension. Additionally, as recent studies showed that learning by explaining affected (meta-)cognitive and motivational factors (Hoogerheide et al., 2014;

Hoogerheide, Visee, et al., 2019), we explored whether perceived mental effort and subjective difficulty as central facets of perceived cognitive load (DeLeeuw & Mayer, 2008; Paas, 1992), and students' monitoring accuracy as a metacognitive factor (Fukaya, 2013) differed across experimental conditions. Further, we investigated whether students' enjoyment and interest as crucial facets of students' valence-related motivational orientations (Eccles & Wigfield, 2002) varied among conditions.

9.2.1 Learning by Explaining to Fictitious Peers

Learning by explaining is a generative learning activity which aims at enhancing students' meaningful learning (Fiorella & Mayer, 2016). In line with the generative learning theory, the process of explaining as a generative act may elicit cognitive (e.g., mental effort) and metacognitive (e.g., monitoring) processes, which should contribute to students' comprehension: First, and in line with Mayer's SOI model (1996), when explaining students need to *select* the most relevant information if the provided materials and to *organize* the information in a coherent way. Then, they need to connect the new contents with their already existing knowledge to *integrate* them into their long-term memory (Brod, 2020; Fiorella & Mayer, 2016; Mayer, 2009). Through this connection, students are able to provide explanations that include further details and information that go beyond the giving materials, which results in new knowledge and meaningful learning (Brod, 2020; Chi, 2009; Chi & Wylie, 2014; Fiorella & Mayer, 2016; Mayer, 1996, 2009; Watson, 1989; Wittrock, 2010). This process additionally triggers students to monitor whether they understood all relevant contents correctly or whether they need to restudy specific information. As a consequence, students' metacognitive monitoring may become more accurate when explaining, which has also been observed for other generative learning activities such as keyword generation or gap filling (Bruin et al., 2017; Fukaya, 2013; Lachner et al., 2020; van Loon et al., 2014). Learning by explaining is commonly implemented in interactive learning settings in which students explain learned contents to present and interactive peers; this setting allows students to exchange ideas and thought, which additionally enhances their understanding (Chi et al., 1989; Chi et al., 1994; Chi, 2009; Chi & VanLehn Kurt, 1991; Dillenbourg, 1999; Duran, 2016; McNamara, 2004, 2017; McNamara & Scott, 1999; Plötzner et al., 1999; Roscoe, 2014; Topping, 2005; Webb et al., 1995; Webb et al., 2009). Interestingly, recent research started to investigate the effectiveness of explaining to a *fictitious* peer and reported promising results (Fiorella & Kuhlmann, 2020; Fiorella & Mayer, 2014; Hoogerheide et al., 2016; Hoogerheide, Visee, et al., 2019; Jacob et al., 2020; Lachner et al., 2020). For instance, Fiorella and Mayer

(2014) conducted an experiment in which university students first studied a text with the expectation to either answer a test (test expectancy) or to explain the content to a fictitious peer (explaining expectancy) after a learning phase. Then, they were randomly assigned to the experimental conditions: They either restudied the material (no explaining condition) or explained the contents to a fictitious peer by generating a video (explaining condition). Results demonstrated that explaining expectancy had an effect only on students' short-term retention (d = 0.55, medium effect), whereas the actual act of explaining learned material resulted in better performance regarding their long-term retention (d = 0.56, medium effect, see also Hoogerheide et al., 2014, for replications).

In another study, Hoogerheide, Visee, Lachner, and van Gog (2019) demonstrated that learning by explaining did not only result in higher learning outcomes but also affected students' motivation during learning. The authors conducted a study with three conditions: Students either explained learned material to a fictitious peer (explaining condition) or wrote a summary about the contents (summarizing condition). A control group restudied the materials (restudy condition). Results showed that students in the explaining condition outperformed students in the restudy condition. Interestingly, students' enjoyment mediated the explaining effect, as students enjoyed explaining more than summarizing, which, in turn, yielded better comprehension. Additionally, students who explained the materials reported higher levels of invested mental effort during the learning task compared to restudying, which, however, was not linked to higher learning outcomes (Hoogerheide, Visee, et al., 2019).

Furthermore, several studies indicated that generating explanations additionally supports students' monitoring accuracy, which is a crucial metacognitive facet of successful learning (Fukaya, 2013; Lachner et al., 2020). For instance, during reorganizing and connecting learned contents for generating an explanation, students might detect unsolved problems, misunderstandings, or missing information that prevent them from understanding the contents deeply (Fiorella & Mayer, 2016). This detection helps students to judge their current understanding more accurately and supports them in restudying information which they did not understand clearly to reach their learning aims (Thiede et al., 2003b). In a study by Fukaya (2013), for instance, students first read five different texts with the intention to either explain the contents or to write keywords about the texts. A control group only had the intention to explain the content but did not actually generate an explanation. Results showed a difference among conditions: Students who actually generated an explanation judged their comprehension more accurately than students who only had the intention to explain or who wrote keywords (*d* = 0.91, large effect). Thus, the actual act of explaining seems to increase students' monitoring accuracy.

Even though serval studies indicated beneficial effects of learning by explaining to fictitious others on students' comprehension and monitoring skills, little is yet known about the underlying mechanism of why explaining is effective. Recently, researchers emphasized the role of social presence during explaining (Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018; Lachner et al., 2021), as higher levels of social presence (which also may arise from virtual of even fictitious characters) are linked to central components of learning, such as cognition or motivation (Atkinson, 2002; Kim, 2013; Richardson et al., 2017; Russo & Benson, 2005; Wang & Antonenko, 2017; Weidlich & Bastiaens, 2017, 2019). On the one hand, the social presence of a fictitious communication partner may engage students to adapt their knowledge to the audience's needs (Clark & Brennan, 1991; Nickerson, 1999), for instance, by providing further details and elaborations that go beyond the contents of the learning material. Such audience-adjustments may result in deeper elaborative processes and contribute to meaningful learning (Wittwer et al., 2010). On the other hand, from a motivational perspective (Hoogerheide, Visee, et al., 2019), the social presence of a fictitious person may also increase the feeling of relatedness (Deci & Ryan, 2000). Higher levels of relatedness may yield higher levels of enjoyment and investments during providing an explanation, and, in turn, contribute to comprehension (Deci & Ryan, 2000; Hoogerheide, Visee, et al., 2019; Lachner et al., in prep.).

9.2.2 Learning by Writing Explanations

Although prior research documented beneficial effects of explaining compared to retrieving, recent studies demonstrated that generating oral explanations is more beneficial than writing an explanation. In an experiment by Hoogerheide, Deijkers, Loyens, Heijltjes, and van Gog (2016) university students first read an instructional text about syllogistic reasoning, and then either wrote an explanation to a fictitious peer or restudied the learning material. In contrast to prior research on oral explaining, results indicated that writing an explanation did not result in higher learning outcomes than restudying. Therefore, the authors directly compared the influence of the explanatory modality in a second experiment. Results showed that only students who explained orally (*d* = 0.43, medium effect), but not in written form (*d* = 0.19, small effect) outperformed students who restudied the learning material. However, there was no significant difference between oral and written explanations regarding students' learning outcome. Additionally, similar to Hoogerheide, Visee, Lachner, and van Gog (2019), results indicated that students who generated an explanation in oral (d = 1.96, large effect) or in written form (d = 0.93, large effect) invested higher levels of mental effort during the learning task compared to students who restudied the material. Relatedly, Lachner, Ly, and Nückles (2018) provided students with a Wikipedia article about combustion engines; after studying it, students were asked to explain the contents to a fictitious peer in either oral or written form. The findings revealed that students who generated an oral explanation reached higher scores in the comprehension posttest compared to students who wrote an explanation (d = 0.67, medium effect), which could be explained by more elaborated explanations given in the oral condition. The authors attributed the findings to the fact that they used more difficult learning materials than Hoogerheide, Deijkers, Loyens, Heijltjes, and van Gog (2016). Against this background, Jacob, Lachner, and Scheiter (2020) conducted a further experiment to resolve the conflicting findings regarding the explanatory modality. Results revealed an interaction effect between learning activity and text difficulty (d = 0.31). The effect of explaining was only significant in the high-difficult condition, but not in the low-difficult condition. Thus, the explaining effect only held true when students learned from difficult but not from less difficult material. More interestingly, students who explained the contents to a fictitious peer orally outperformed students who wrote an explanation. Again, this effect was only significant when the learning material was difficult, but not when it was less complex. Additionally, perceived social presence and the richness of explanations mediated the explanatory effect: Students who explained orally perceived a stronger presence of the fictitious peer (measured by the number of personal references) compared to students who wrote an explanation. Higher levels of social presence, in turn, was associated with richer explanations (measured by the number of mentioned concepts) which resulted in better learning outcomes, at least for the difficult text. Apparently, the social presence during explaining accounted for the superiority of oral explaining. Furthermore, students who explained orally judged their current understanding more precisely than students who wrote an explanation and invested more mental effort than students who retrieved the materials (Hoogerheide, Visee, et al., 2019; Jacob et al., 2020).

These findings are in line with literature in applied linguistics, in which it is generally argued that writing is a rather solitary process, since the writer is normally separated from the audience regarding time and place (Chafe, 1982; Sindoni, 2013). Due to the lack of a social presence during written activities, the writer in contrast tends to adopt a knowledge-telling perspective, and as such only retrieves the content without adapting the content to a particular audience's needs (Scardamalia & Bereiter, 1987). These detrimental effects may increase, as writing often induces higher levels of cognitive load than speaking, which can be considered as an automated and less demanding process compared to writing. Overall, the lower levels of audience adjustments may be less conducive to learning (Clark & Brennan, 1991), as, for instance, students provide lower amounts of examples to elaborate the content.

9.2.3 The Present Study: Inducing Social Presence to Enhance Writing Explanations

We conducted an experiment to investigate whether learning by writing explanations could be supported by inducing social presence during explaining. We used validated experimental materials (Golke & Wittwer, 2018; 2020), and provided university students with a text about immunology during the study phase. Afterwards, they were randomly assigned to one of four conditions: They explained the contents in oral form (oral condition), wrote their explanations in a text editor (standard written condition), or wrote them in a chat messenger program (chat condition, see Figure 13). In this messenger program, students could see a profile picture and a message from the fictitious peer, which we assumed would induce higher levels of social presence (Weidlich & Bastiaens, 2019). Students in the control condition were asked to retrieve the contents (retrieval condition). We stated the following hypotheses, which were preregistered on AsPredicted.org (https://aspredicted.org/3nu5m.pdf).

In line with previous evidence, we hypothesized that students who generate an explanation (oral or written form) outperform students who retrieve the material regarding students' comprehension posttests (*Hypothesis 1*). Additionally, based on previous findings, we hypothesized that students who explain orally outperform students who write an explanation (standard written condition) to a fictitious peer (*Hypothesis 2*). Furthermore, we hypothesized that students who write an explanation with increased social presence (chat condition) outperform students in the standard written condition (*Hypothesis 3*). Additionally, we explored potential differences between the oral condition and the chat condition and assumed comparable outcomes regarding students' comprehension since social presence was induced in the chat condition aiming at reaching a similar level of perceived social presence as explaining orally.

Based on previous research, we additionally investigated potential differences regarding (meta-)cognitive (i.e., monitoring accuracy, mental effort, and subjective difficulty; Fukaya, 2013; Jacob et al., 2020) and motivational factors, such as enjoyment and interest (Hoogerheide et al., 2014; Hoogerheide, Visee, et al., 2019). To account for the quality of the generated explanations, we measured three characteristics of the generated explanations (i.e., personal references, concepts, elaborations), which are commonly measured in research on learning by explaining (Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018).

9.3 Method and Methods

9.3.1 Participants and Design

The current study was approved in written form by the ethics committee of the Leibniz-Institut für Wissensmedien in Tübingen (approval number: LEK2019/009). We recruited university students (N = 137) from study programs that were not related to the study topic (i.e., biology). This sample reached the required sample size of 126 participants, as determined by an *a priori* power analysis. Power was set to .80, α-error to .05, and the assumed effect size to η^2 = .08, as recent studies documented differences of medium to large effect size of explaining (Hoogerheide et al., 2016; Lachner et al., 2018) and social presence in learning settings (Weidlich & Bastiaens, 2017).

The mean age of the students was 23.23 years (SD = 2.65) and 73 % of them were female. The students either stated to be German native speakers (85 %), or that they grew up bilingually with German (15 %). The students were advanced students, on average in their 5th semester (SD = 3.19) of their current study program and were mostly enrolled in humanities programs (71 %). On average, they had taken biology classes in school for 8 years (SD = 2.13), had 10 points (on a scale from 0 to 15, corresponding to a B-) in their last report card in biology (SD = 3.25), and showed low to medium prior knowledge skills in the prior knowledge test (M = 2.15; SD = 1.13; on a scale from 0 to 5).

Students were randomly assigned to one of four experimental conditions (i.e., retrieval condition: n = 34; standard written condition: n = 34; chat condition: n = 35; oral condition: n = 34), which was the independent variable. The dependent variable was students' text comprehension, measured by two knowledge subtests comprising text-based questions and inference questions. To investigate potential (meta-)cognitive differences among conditions, we additionally collected data regarding students' perceived mental effort, subjective difficulty, and monitoring accuracy during the study phase (i.e., reading the learning material) as additional control variables, and during the learning activity (i.e., explaining vs. retrieving) as potential mediators. Moreover, to investigate potential differences regarding students' motivation during the learning activity, we asked them to rate their enjoyment and interest during the learning activity. As potential underlying processes, we additionally analyzed three characteristics of students' generated explanations: Personal references as an indicator for social presence, number of elaborations, and number of concepts to investigate further underlying learning processes. As control variables, we measured students' prior knowledge, self-efficacy in explaining, their biological interest at the beginning of the study, and their perceived social presence during explaining.

9.3.2 Study Text

We used a validated text from the domain of biology from Golke and Wittwer (2018). The text was about immunology and dealt with immune research based on laboratory mice. The text was previously used in a study on learning by explaining (Jacob et al., 2020). Overall, the text had a length of 397 words and constituted a relatively complex text regarding common measures of text difficulty (Jacob et al., 2020).

9.3.3 Prior Knowledge Test as Control Variable

We used the prior knowledge test from Golke and Wittwer (2018), which contained five questions with an open-ended answer format and represented a multidimensional construct, measuring different subcomponents of immunology (e.g., "Why can viruses be dangerous for humans?"; McDonald's ω_t = .51). For each right answer students could receive one point, yielding a maximum score of five points. Two independent raters coded 20 % of the tests. As interrater reliability was excellent ($ICC_{2,1}$ = .91), one rater coded the remaining answers (Koo & Li, 2016).

9.3.4 Knowledge Posttests as Dependent Variables

We measured students' text-based knowledge and inference knowledge with two different posttests from Golke and Wittwer (2018). The tests consisted of different questions than the prior knowledge test and represented multidimensional constructs.

9.3.5 Text-based Questions

The text-based questions comprised six open-ended questions, which aimed at measuring students' basic knowledge of the text (e.g., "What is the main concern against research with laboratory mice?"; McDonald's ω_t = .44). For each correct answer, students could reach one point, resulting in a maximum score of six points. Two independent raters coded 20 % of the tests. As interrater reliability was excellent ($ICC_{2,1}$ = .95), one rater coded the remaining answers.

9.3.6 Inference Questions

The inference questions measured students' advanced understanding (e.g., "How can the results in the text help to clarify the basic problems of immune research with mice?"; McDonald's ω_t = .45), as students had to combine different aspects and information across the text, and to draw conclusions. Again, students could receive one point per answer, yielding a maximum score of six points. Two independent raters coded 20 % of the tests. As interrater reliability was excellent ($ICC_{2,1}$ = .92), one rater coded the remaining answers.

9.3.7 (Meta-)Cognitive Processes During the Learning Activity

9.3.7.1 Monitoring Accuracy

To measure students' monitoring accuracy, we asked them to make prospective judgments about their expected performance on the posttest (i.e., text-based and inference questions combined) to investigate their monitoring accuracy by rating the following item: "How confident are you that you can answer questions to the text correctly?" (Baars et al., 2017; Prinz et al., 2018). Monitoring accuracy is commonly measured by the correspondence between students' judgements of their own current understanding and their actual performance on a comprehension test (Golke & Wittwer, 2018; Maki & McGuire, 2009; Wiley et al., 2005). Therefore, students rated their comprehension after the study phase and after the learning activity on a scale from 0 % (no confidence) to 100 % (absolute confidence) (Jacob et al., 2020; Schleinschok et al., 2017). We operationalized students' monitoring accuracy in terms of bias (Golke & Wittwer, 2018; Griffin et al., 2009; Jacob et al., 2020; Lachner et al., 2020; Schraw, 2009; Wiley et al., 2016). Bias refers to the signed difference between students' estimated performance and the actual performance (i.e., $X_{Judgment} - X_{Performance}$). Hence, this estimation indicated whether students over- or underestimate their own performance. Positive values indicate an overestimation and negative values indicate an underestimation of their judged performances. A value of zero indicates an accurate judgment.

9.3.7.2 Cognitive Load

We asked students to rate their perceived mental effort (i.e., “How much effort did you invest in explaining the material?”), as subjective proxies to perceived cognitive load, after the study phase and after the learning activity on a 9-point Likert scale from 1 “very little” to 9 “very much” (Paas, 1992).

Additionally, students rated their subjective difficulty (i.e., “How easy was it for you to explain the material?”), as subjective proxies to perceived cognitive load after the study phase and after the learning activity on a 9-point Likert scale from 1 “very easy” to 9 “very difficult” (DeLeeuw & Mayer, 2008).

9.3.8 Motivation During the Learning Activity

9.3.8.1 Enjoyment During the Learning Activity

We measured students’ enjoyment during the learning activity by using two self-generated items (e.g., “I enjoyed doing the task”). The students rated their enjoyment on a 4-point Likert scale from 1 “not at all” to 4 “absolutely”. Reliability was good (McDonald’s ω_t = .96).

9.3.8.2 Interest in the Learning Activity

We measured students’ enjoyment during the learning activity by using two self-generated items (e.g., “I enjoyed doing the task”). The students rated their interest on a 4-point Likert scale from 1 “not at all” to 4 “absolutely”. Again, reliability was good (McDonald’s ω_t = .81).

9.3.9 Characteristics of Explanations

Based on prior research, we analyzed three characteristics of the generated explanations as indictors for underlying processes during explaining: Personal references, concepts, and elaborations (Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018).

9.3.9.1 Personal References

Based on discourse literature, we used personal references (i.e., “I”, “you”, etc.) as an indicator for the perceived social presence (Akinnaso, 1985; Chafe, 1982; Chafe & Tannen, 1987; Sindoni, 2013) which is frequently also used in learning by explaining literature

(Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018). We automatically counted the number of personal references with RStudio (R Core Team, 2018).

9.3.9.2 Concepts

As an indicator of the level of comprehensiveness, we counted the number of concepts per explanation (Boshuizen & Schmidt, 1992). Concepts are the mentioned constructs within an explanation. For instance, the sentence "*Laboratory mice* grow up in sterile *environments* and, therefore, the *laboratory mice* do not have any *contact* with *diseases*" contains four concepts (marked in italics). Redundant concepts (e.g., "laboratory mice") were ignored (Boshuizen & Schmidt, 1992). Two independent raters coded 20 % of the explanations. As interrater reliability was excellent ($ICC_{2,1}$ = .95), one rater coded the remaining explanations.

9.3.9.3 Elaborations

As a third characteristic, we counted the number of elaborations within the explanations. An elaboration was operationalized as an idea unit, such as examples, analogies, and own experiences (Fiorella & Kuhlmann, 2020; Lachner et al., 2018). For instance, the sentences "Laboratory mice are too clean *because they have been cultivated over generations*" is an elaboration, as it was not mentioned in the study text. The student, therefore, combined the information of the text with her or his prior knowledge to generate the example. Two independent raters counted the number of elaborations for 20 % of all explanations. Interrater reliability was excellent, $ICC_{2,1}$ = .95. Therefore, one rater coded the remaining explanations.

9.3.10 Additional Control Measures

9.3.10.1 Perceived Social Presence

We asked students to rate their feelings of a social presence form the fictitious peer during explaining. We asked students to rate three self-generated items (i.e., "How strongly did you imagine that Lisa was real?"; "How important was it for you that Lisa understands the contents?"; "How strongly did you perceived being in a communicative situation"; McDonald's ω_t = .84). Based on related approaches on subjective assessments of task characteristics, we used a 9-point Likert scale from 1 "not at all" to 9 "completely" (DeLeeuw & Mayer, 2008; Paas, 1992). Since students in the retrieval condition were not asked to explain to a fictitious peer, only students in explaining conditions (i.e., standard written condition, chat condition, oral condition) rated their perceived social presence.

9.3.10.2 Self-efficacy in Explaining

We assessed students' self-efficacy in explaining as a further control variable. We used three adapted items (e.g., "I always find ways to explain even difficult contents"; McDonald's ω_t = .74) from Jerusalem and Schwarzer (1999). Students rated their explaining skills on a 4-point Likert scale from 1 "I completely disagree" to 4 "I completely agree".

9.3.10.3 Biological Interest

We measured students' interest in biology as an additional control variable by using three items (e.g., "I am fascinated by biological topics"; McDonald's ω_t = .91), based on Kunter et al. (Kunter et al.). Students rated the items on a 4-point Likert scale from 1 "I completely disagree" to 4 "I completely agree".

9.3.11 Procedure

The instructor welcomed the students and informed them about the study procedure. After providing written consent, all students were seated individually in front of a laptop in noise-cancelling cubicles so that they neither could see or hear each other. A maximum of 6 students could participate in one session. The entire study was self-paced and all instructions were provided in an online learning environment created by Klemke (2017).

First, students rated their efficacy in explaining and their interest in biology. Then, they answered the prior knowledge test. Afterwards, they read the study text in a self-paced manner without the intent to explain or retrieve the material afterwards. After this study phase, they rated their mental effort and difficulty during reading, and provided a monitoring judgment. Then, the students were randomly assigned to one of four experimental conditions (i.e., retrieval condition, standard written condition, chat condition, oral condition). All students had 10 minutes to accomplish the learning activity. During the learning activity, they could take notes on a separate sheet but were not able to see the learning materials anymore. Students in the oral and standard written condition were given the following instruction:

Imagine the following scenario: One person could not participate in the study. However, she is highly interested in the topic of the study. She has not read anything about the topic yet. Therefore, she asks you to explain the central contents of the text. Please provide her a clear and detailed explanation, so that she can understand the content without additional information. You may take notes on a separate sheet.

Students in the oral explanations recorded their explanation with a microphone which was connected to the laptop. Students in the standard written condition wrote their explanation into a text editor box. In contrast, students in the chat condition received a more personalized instruction:

Here you can see a chat with Lisa. Lisa is highly interested in the topic of the study. She has not read anything about the topic yet. Therefore, she asks you to explain the central contents of the text. Please provide her a clear and detailed explanation, so that she can understand the content without additional information. You may take notes on a separate sheet.

Students in the chat condition provided their explanation in a messenger mockup (Figure 13), which was an imitation of WhatsApp Messenger (freeware created by WhatsApp Inc. in 2009 and acquired by Facebook in 2014). We provided different social cues to induce higher levels of social presence (see Weidlich & Bastiaens for a similar approach; Weidlich & Bastiaens, 2019). First, students could see a profile picture from *Lisa*. Second, they received a short message from Lisa who directly asked the students for an explanation (Figure 13). We automatically adapted the time of the received message to increase the synchrony of the communication. Students could send Lisa a message to share their explanation within the chat. Students in the retrieval condition only retrieved the material via an open recall task (Jacob et al., 2020; Lachner et al., 2020) with the following instruction:

Please retrieve the information of the text. You may take notes on a separate sheet.

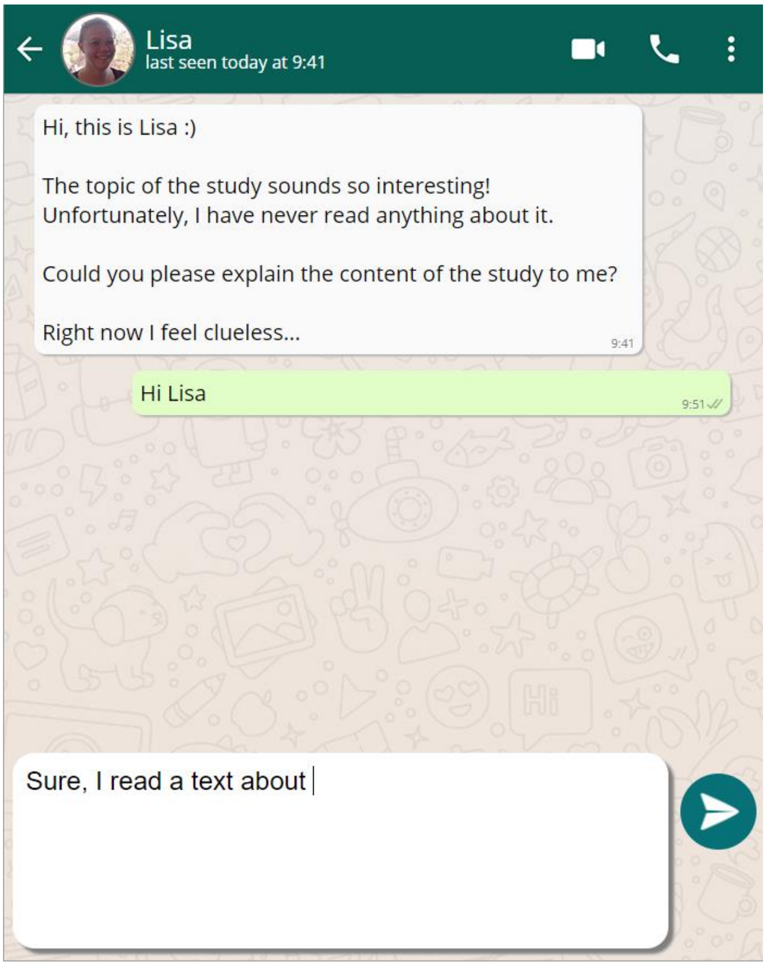

Figure 13. Simulated mockup messenger chat for chat condition with induced social presence. Mockup messenger chat with a profile picture and message from the fictitious student Lisa. Students in the chat condition could send text messages which appeared in the chat afterwards. We received written approval to publish the picture of the corresponding individual who completed the consent form for publication in a PLOS journal.

After this explaining task, students again rated their effort and difficulty during explaining or retrieving and provided a monitoring judgment. Additionally, they rated their enjoyment and interest in the learning activity and students who generated an explanation further stated their perceived social presence during explaining. Finally, all students answered the posttests. The study took approximately 1 hour and was rewarded with 8 €.

9.4 Results

We used partial η_p^2 and Cohens' d as effect size measures, qualifying values of η_p^2 = .01, .06, .14 and d = .20, .50, .80 as small, medium, and large effects (Cohen, 1977). Additionally, we applied an alpha level of α = .05.

9.4.1 Preliminary Analyses

Preliminary analyses showed no significant differences between gender among conditions, χ^2 (6, 137) = 5.50, p = .538, η_p^2 = .07. Results of a MANOVA showed no significant differences between age, students' number of semesters, grade in their last report card in biology, prior knowledge, self-efficacy in explaining, and biological interest across conditions, F(3, 133) = 1.08, p = .370. Additionally, students rated the social presence as comparably across the three explaining conditions, F(2, 100) = 2.23, p = .113, η_p^2 = .04.

9.4.2 Learning Outcome

The descriptive statistics (see Table 3) suggested that students showed comparable learning outcomes across conditions. Therefore, although our preregistered analysis plan was to conduct two separate contrast analyses for each dependent variable (which would have resulted in four tests in total), we decided to use MANCOVAs instead to reduce the number of statistical tests and to use the most economic statistical approach (Ranganathan et al., 2016; Streiner, 2015). The text-based and the inference questions were the dependent variables, experimental condition the independent variable, and prior knowledge the control variable. Results showed a main effect of prior knowledge, F(1, 132) = 14.59, p < .001, η_p^2 = .10, but no differences between conditions, F(3, 132) = 2.03, p = .062, η_p^2 = .04 (text-based questions: F(3, 132) = 2.56, p = .058, η_p^2 = .05; inference questions: F(3, 132) = 0.69, p = .562, η_p^2 = .01). As recommended by Biel and Friedrich (2018), to investigate whether the non-significant finding can be attributed to the fact that that the null-hypothesis was true (i.e., no differences among conditions), we additionally computed Bayesian factors with the *bain* package (Hoijtink et al., 2019). Values between 0 and 1 can be interpreted as evidence in favor of the alternative hypothesis. Values greater than 1 are suggested as evidence in favor of the null hypothesis. A value of 1 represents no preference for either hypothesis. Results showed a Bayes factor higher than 1 regarding both posttests (text-based questions: BF_{01} = 9.43; inference questions: BF_{01} = 128.04), suggesting that we can assume that the null-hypothesis is true and that all conditions resulted in comparable learning outcomes.

Table 3

Summary of Means and Standard Deviations (in Parentheses) for All Measurement

	Retrieval Condition	Written Condition	Chat Condition	Oral Condition
Learning outcomes				
Text-based questions (0 – 6)	3.29 (1.05)	2.94 (1.12)	2.77 (1.15)	3.35 (0.97)
Inference questions (0 – 6)	2.79 (1.23)	3.09 (1.09)	2.89 (1.18)	2.75 (0.98)
(Meta-)Cognitive processes				
Monitoring accuracy	0.11 (0.19)	0.14 (0.22)	0.17 (0.20)	0.09 (0.18)
Perceived mental effort (1 – 9)	5.24 (2.09)	6.71 (1.36)	6.66 (1.47)	6.09 (1.73)
Subjective difficulty (1 – 9)	4.24 (2.20)	4.65 (1.84)	4.94 (2.07)	5.18 (2.22)
Motivation during learning activity				
Enjoyment (1 – 4)	2.29 (0.75)	2.91 (0.67)	3.14 (0.73)	2.91 (0.87)
Interest (1 – 4)	2.12 (0.71)	2.74 (0.67)	3.17 (0.69)	2.84 (0.79)
Characteristics of explanations				
Personal references	-	0.12 (0.48)	1.89 (1.97)	0.71 (1.66)
Concepts	-	24.47 (5.70)	21.54 (6.18)	28.53 (11.39)
Elaborations	-	1.09 (1.03)	0.94 (0.87)	1.68 (1.63)
Control variables				
Prior-knowledge (0 – 5)	2.06 (1.17)	2.26 (1.03)	2.03 (1.18)	2.25 (1.15)
Interest in biology (1 – 4)	2.69 (0.90)	2.54 (0.87)	2.57 (0.79)	2.97 (0.74)
Self-efficacy explaining (1 – 4)	2.83 (0.57)	2.59 (0.62)	2.73 (0.65)	2.80 (0.59)
Perceived social presence (1 – 9)	-	5.78 (1.80)	5.73 (2.04)	4.88 (2.10)

9.4.3 Influence of Learning Activity on (Meta-)Cognitive Processes

9.4.3.1 Monitoring Accuracy

In a first step, we conducted one-sample t-tests to investigate whether students significantly over- or underestimated their comprehension after the learning activity (i.e., bias). We contrasted students' actual judgements of their current understanding relative to their performance with an accurate judgment, which was represented by a value of zero. Results revealed that students in all conditions significantly overestimated their current understanding (retrieval condition: $t(33) = 3.43$, $p = .002$, $d = .58$; written condition: $t(33) = 3.74$, $p < .001$, $d = .64$; chat condition: $t(33) = 5.09$, $p < .001$, $d = .85$; oral condition: $t(33) = 2.81$, $p = .008$, $d = .50$).

In a second step, we analyzed potential differences among conditions. We conducted an ANCOVA with experimental condition as independent variable, students' judgements of their current understanding after the learning activity as dependent variable, and students' judgements of their understanding after the study phase as covariate to control for potential intra-individual differences (Hertzog et al., 2013; Lachner et al., 2020). Results showed a main effect of students' judgements after the study phase, $F(1, 132) = 136.47$, $p < .001$, $\eta_p^2 = .50$, but not of experimental condition, $F(3, 132) = 2.41$, $p = .070$, $\eta_p^2 = .03$. Again, we conducted a Bayesian analysis. Results showed a Bayes factor of $BF_{01} = 51.68$, indicating that all conditions resulted in comparable (biased) judgements.

9.4.3.2 Cognitive Load

First, to analyze whether experimental conditions differed regarding students' mental effort during the learning activity, we performed an ANOVA with mental effort as dependent variable and experimental condition as independent variable. Results showed a significant difference among conditions, $F(3, 133) = 5.63$, $p = .001$, $\eta_p^2 = .11$. Post-hoc comparisons (Scheffé) revealed that students who wrote an explanation investigated more mental effort than students who retrieved the material (standard written condition: $p = .006$; chat condition: $p = .008$). The remaining group comparisons were not significant. See Table 4 for correlations with students' learning outcomes.

Second, to investigate potential differences among conditions regarding students' subjective difficulty, we conducted an ANOVA with subjective difficulty as dependent variable and experimental condition as independent variable. Results indicated no differences among conditions, $F(3, 133) = 1.29$, $p = .282$, $\eta_p^2 = .03$. Bayesian analyses showed a Bayes factor of $BF_{01} = 44.79$, indicating that there were no differences between conditions regarding students' subjective difficulty.

Table 4

Correlations with Confidence Intervals Regarding All Variables

		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	**Prior knowledge**											
2	**Text-based questions**	**.31**										
		[.15, .45]										
3	**Inference questions**	**.39**	**.36**									
		[.24, .52]	[.20, .50]									
4	**Monitoring accuracy**	-.14	**-.27**	**-.32**								
		[-.30, .03]	[-.42, -.11]	[-.46, -.16]								
5	**Perceived mental effort**	.05	.09	**.19**	.16							
		[-.12, .22]	[-.08, .25]	[.02, .35]	[-.01, .32]							
6	**Subjective difficulty**	-.04	-.12	**-.18**	**-.36**	.07						
		[-.21, .13]	[-.28, .05]	[-.34, -.01]	[-.50, -.21]	[-.10, .24]						
7	**Enjoyment**	.05	.09	**.17**	**.21**	**.34**	**-.25**					
		[-.12, .21]	[-.08, .25]	[.00, .33]	[.05, .37]	[.18, .48]	[-.40, -.08]					
8	**Interest**	.02	-.02	.08	**.25**	**.28**	**-.18**	**.81**				
		[-.15, .19]	[-.19, .15]	[-.09, .24]	[.08, .40]	[.11, .42]	[-.33, -.01]	[.75, .86]				
9	**Social presence**	.08	-.02	.04	**.21**	**.28**	**-.28**	**.47**	**.40**			
		[-.11, .27]	[-.21, .17]	[-.15, .24]	[.01, .38]	[.09, .45]	[-.45, -.09]	[.30, .61]	[.23, .55]			
10	**Personal references**	.05	.02	.06	.02	**.17**	.10	**.17**	.16	.18		
		[-.12, .21]	[-.15, .19]	[-.11, .23]	[-.15, .19]	[.01, .33]	[-.07, .26]	[.00, .33]	[-.01, .32]	[-.01, .36]		
11	**Concepts**	**.34**	**.45**	**.34**	-.03	**.20**	**-.24**	.10	.05	.07	.07	
		[.15, .50]	[.28, .59]	[.15, .50]	[-.22, .16]	[.01, .38]	[-.42, -.05]	[-.10, .29]	[-.15, .24]	[-.12, .26]	[-.12, .26]	
12	**Elaborations**	**.28**	**.30**	**.20**	.02	.16	**-.26**	.18	.13	**.21**	.05	**.70**
		[.09, .45]	[.11, .47]	[.00, .38]	[-.17, .21]	[-.04, .34]	[-.43, -.07]	[-.01, .36]	[-.07, .31]	[.02, .39]	[-.15, .24]	[.59, .79]

Note. Significant correlations are highlighted in bold font; $p < .05$.

9.4.4 Motivation During the Learning Activity

9.4.4.1 Enjoyment

We explored differences regarding students' enjoyment during the learning activity by conducting an ANOVA with experimental condition as independent variable. Results revealed significant differences among conditions, $F(3, 133) = 7.91$, $p < .001$, $\eta_p^2 = .15$. Post-hoc comparisons showed that students in all explaining conditions (i.e., standard written condition: $p = .012$; chat condition: $p < .001$; oral condition: $p = .012$) enjoyed the learning activity more than students in the retrieval condition.

9.4.4.2 Interest

Similarly, we investigated potential differences regarding the interestingness of the learning activity among conditions. Results of an ANOVA with experimental condition as independent variable showed a significant effect of condition, $F(3, 133) = 13.11$, $p < .001$, $\eta_p^2 = .23$. Post-hoc comparisons revealed that students rated all three explaining conditions (i.e., standard written condition: $p = .006$; chat condition: $p < .001$; oral condition: $p < .001$) as more interesting than the retrieval condition.

9.4.5 Characteristics of the Explanations

As expected, within the explaining conditions, the generated explanations differed among conditions: Personal references, $F(2, 100) = 12.19$, $p < .001$, $\eta_p^2 = .20$; concepts, $F(2, 100) = 6.38$, $p = .002$, $\eta_p^2 = .11$, and elaborations, $F(2, 100) = 10.35$, $p = .034$, $\eta_p^2 = .07$. Post-hoc comparisons (Scheffé) showed that students mentioned more personal references in the chat condition compared to the standard written condition, $p < .001$, and the oral condition, $p < .001$, which can be regarded as an indicator that the chat condition indeed induced higher levels of social presence. Additionally, students who explained orally mentioned more concepts, $p = .003$, and provided more elaborations, $p = .048$, than students in the chat condition. None of the other comparisons were significant.

9.5 Discussion

The aim of the current study was to investigate whether increasing social presence during writing explanations aids learning. Contrarily to our preregistered hypotheses, we did not obtain a significant effect of induced social presence during the learning activity compared to a less social learning environment. Apparently, only raising the social presence did not contribute to learning. Additionally, we did not find significant differences between the explaining and the retrieval conditions. Thus, we did not replicate prior findings on the effectiveness of explaining (Fiorella & Mayer, 2013, 2014; Hoogerheide et al., 2014; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019). Relatedly, we were also not able to replicate beneficial effects of oral explaining compared to writing explanations (Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018). The null findings regarding students' comprehension might be attributed to low levels of prior knowledge. Little prior knowledge limits students in learning new contents adequately (Brod, 2020; Renkl, 2009). However, prior research revealed that generative learning activities, such as explaining, are in particular beneficial for low prior knowledge students (Hoogerheide, Renkl, et al., 2019). Additionally, students showed comparable monitoring accuracy ratings across conditions. Interestingly, all students overestimated their current understanding. This, however, is in line with prior research that highlighted that students generally tend to overestimate their current comprehension (Eitel, 2016; Mayer et al., 2007; Prinz et al., 2020). Nevertheless, we need to reject our hypotheses regarding effects on students' comprehension and their monitoring accuracy. We want to note, however, that the obtained null-findings are in line with a growing body of empirical research that was also not able to document an effect of explaining (Fukaya, 2013; Lachner et al., 2020; Rhoads et al., 2019; Roscoe & Chi, 2008).

In this context, our study is of particular interest, as we used previously tested materials and knowledge tests in the context of explaining (Jacob et al., 2020), and had sufficient test power to test our hypotheses. As additional safe-guard we also computed Bayes factors, which similarly suggested that our findings are rather in favor of the null-hypothesis. Thus, we are confident that this study represents reliable evidence to reject the explaining and modality hypothesis, at least in the current study context.

Furthermore, the analyses of the characteristics of the explanations demonstrated that the quality of the generated explanations was rather low, suggesting that the students were less capable to generate effective explanations. This finding may explain why we did not find any effects on learning. Our explorative analysis regarding the (meta-)cognitive and motivational conditions of explaining revealed distinct differences between explaining and retrieval, as students perceived more effort and more motivation (i.e., enjoyment, interest) during the learning activity, which is in line with previous research (Hoogerheide, Visee, et al., 2019; Jacob et al., 2020). Cognitive conditions, such as mental effort, are strongly linked to students' learning (van Gog et al., 2020), and, therefore, should be investigated in more detail in combination with additional learning activities, such as learning by explaining, in future research. Moreover, learning enjoyment and relatedly interest are regarded as important value-related facets of intrinsic motivation (Deci & Ryan, 2000), and intrinsic motivation is particularly key to whether and how students would persist to use explaining as a learning strategy beyond the experimental context (Yi & Hwang, 2003). To explore the differences regarding students' enjoyment and mental effort in more detail, we additionally conducted explorative mediation analyses by applying a bootstrapping approach with 1,000 simulations based on Hayes (Hayes, 2013). Results revealed that enjoyment but not mental effort mediated learning by explaining on students' text-based comprehension as the indirect effect was significant (ACME = 0.24, 95 % CI (0.03, 0.50), p = .024). These findings are in line with the results by Hoogerheide and colleges who demonstrated that explaining led to higher levels of enjoyments, which, in turn, enhanced students' learning (Hoogerheide, Visee, et al., 2019).

Nevertheless, against the background of the relatively low quality of students' explanations, as a theoretical consequence of our study, we suggest that future research should focus on how students could be supported to generate high-quality explanations. Given that students are generally not familiar with explaining challenging contents (Rozenblit & Keil, 2002), they might depend on further support to be able to apply this learning activity successfully. As a first attempt, Lachner and Neuburg (2019) investigated the role of formative feedback during explaining (in written form). They found that formative feedback helped students generate more cohesive explanations, which finally contributed to their comprehension. Additionally, the self-explaining literature focusses on two main instructional approaches (Fonseca & Chi, 2011), supporting self-explaining directly by means of pre-trainings (McNamara, 2017) or indirectly by the use of prompts (Chi et al., 1994; Schworm & Renkl, 2007), which should work as strategy activators to enact deep-level explaining strategies. Whether and under which conditions such direct and indirect support procedures also hold true for generating explanations to fictitious others has to be investigated by further research.

9.5.1 Study Limitations and Future Research

With our study, we provide the first empirical approach to systematically investigate the influence of perceived social presence on the effectiveness of explaining to fictitious others. In this context, however, we would like to point out some limitations of our study. First, an unexpected finding was that we indeed found an effect of our intervention on the number of personal references, but not on the social presence ratings. This finding might be contradictory at first glance. However, we want to note that such contradictory findings frequently occur in the context of inducing discourse situations (Savary et al., 2015; Weinhuber et al., 2019). As such, these findings are often interpreted as suggesting that inducing discourse situation may not affect subsequent communication processes directly, but rather subconsciously. This may explain the differences between the number of personal references and the intentional presence ratings.

Another surprising finding was that students in the explaining conditions did not outperform students in the retrieval practice condition. This not only contradicts our assumptions but also prior studies that showed beneficial effects of explaining in contrast to restudying or retrieving the learned material (Fiorella & Kuhlmann, 2020; Fiorella & Mayer, 2014; Hoogerheide et al., 2016; Hoogerheide, Visee, et al., 2019; Jacob et al., 2020; Lachner et al., 2020). In this context, we would like to point out that we used the same learning materials as in a previous study by Jacob, Lachner, and Scheiter who obtained effects of explaining with this difficult learning material (Jacob et al., 2020). In contrast to Jacob et al., students in our study showed slightly lower learning outcomes. Additionally, results indicated a higher correlation between prior knowledge and students' learning outcome, suggesting that the provided learning material was too difficult, which may explain the null findings. Seemingly, students have problems in applying this learning strategy successfully when the learning material is overly difficult (Brod, 2020), which highlights the need for additional support during learning, such as pre-trainings or prompts (Chi et al., 1994; McNamara, 2004; Schworm & Renkl, 2007).

Another caveat refers to the measurement of the generated explanations. Based on prior research, we measured the number of personal references, concepts, and elaborations in students' generated explanations (Akinnaso, 1985; Chafe, 1982; Chafe & Tannen, 1987; Einhorn, 1978; Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018). We used these characteristics as a coarse proxy for the underlying processes during the learning activity (Fiorella & Kuhlmann, 2020; Jacob et al., 2020; Lachner et al., 2018). Further research is needed to explore the generated explanations in more detail which could provide a deeper insight into the explaining effect. This indeed could disentangle which detrimental effects occurred during explaining. Further research should, therefore, implement additional online-measures, for instance, think-aloud protocols or log-file-data (Ericsson & Simon, 1999).

Another limitation refers to the relatively low homogeneity indices. However, we want to note that we measured a broad understanding of the topic by asking different and independent questions with a restricted set of 5 to 6 items per test. This decision likely resulted in a trade-off regarding the internal consistency of the knowledge measures, as they likely covered diverse sub-components of students' comprehension (Dunn et al., 2014). Further research with advanced statistical procedures is needed which explicitly takes the multidimensionality into account, for instance, by using bifactor-(S-1) measurement models which, however, require larger sample sizes (Eid et al., 2017).

Finally, we would like to point out restrictions regarding the measurement of enjoyment, as we only measured enjoyment quantitatively by means of two items. Our results revealed, in line with the study by Hoogerheide et al. (2019), that explaining resulted in higher levels of enjoyment compared to retrieving the materials, which is in part indicative for the prognostic validity of our instrument. However, as we only used a quantitative measure, we are not able to conclude whether and how enjoyment affected learning outcomes in a qualitative manner. Therefore, further research is needed that investigates the impact of enjoyment on the explaining effect in more detail by implementing qualitative methods, such as interviews (Csikszentmihalyi, 2000; Wienke & Jekauc, 2016).

9.5.2 Conclusion

The findings of our study show that explaining does not necessarily foster learning, although it enhances investments of mental effort and increases students' motivation. Furthermore, our findings indicate that simply inducing social presence does not increase the effectiveness of explaining. It is, therefore, up for further research to find ways to trigger deep-learning processes during explaining. As such, explaining can contribute to students' mental effort and motivation, which are also linked to higher learning outcomes.

10

STUDY 3: WHAT'S APP? SCHOOL STUDENTS' SELF-CONCEPT MODERATES THE EFFECTIVENESS OF GENERATING TECHNOLOGY-MEDIATED EXPLANATIONS

10. Study 3: What’s App? School Students’ Self-Concept Moderates the Effectiveness of Generating Technology-Mediated Explanations

10.1 Abstract

Asking students to generate explanations to fictitious others by means of different technologies (e.g., instant messenger, video) gains popularity at schools. Although the potential of these technology-mediated explaining activities is appealing, less is known whether generating these explanations is effective for school students. Additionally, empirical evidence regarding general effects of learning-by-explaining proposes that generating explanations is constrained by different boundary conditions. In this experimental field study, we investigated the effectiveness of technology-mediated explaining with seventh grade students (N = 129). More importantly, we contrasted different implementation modalities of explaining (i.e., written versus oral explaining), and examined the moderating role of cognitive and motivational prerequisites (i.e., prior knowledge, academic self-concept). After assessing prior knowledge and academic self-concept, students attended a lesson on photosynthesis. Then, students randomly explained the contents in written versus oral form in a mock-up messenger or retrieved the contents (control condition). We neither obtained an effect of explaining, nor did the explanatory modality account for students’ learning. However, academic self-concept, but not prior knowledge moderated the explaining effect on comprehension, as only students with low self-concept profited from explaining. The findings highlight that technology-mediated explaining is not necessarily an effective strategy for school students, but strongly depends on students’ motivational prerequisites, that is, their academic self-concept.

STUDY 3: WHAT’S APP? SCHOOL STUDENTS´ SELF-CONCEPT MODERATES THE EFFECTIVENESS OF GENERATING TECHNOLOGY-MEDIATED EXPLANATIONS

Publication process: Major revision: Computers & Education

10.2 Theoretical Background

In informal contexts, it is common practice that students use technologies, such as videos or messenger services to provide explanations about distinct contents to a remote or even fictitious audience (Alon & Herath, 2014; Orús et al., 2016; Pereira et al., 2014). Following generative learning theory (Fiorella & Mayer, 2016; Wittrock, 2010), these technology-mediated explaining activities can be beneficial for learning, as they trigger students’ active sense-making of the previously learned contents (Fiorella & Kuhlmann, 2020; Fukaya, 2013; Lachner et al., 2020). To this end, also in formal contexts (i.e., school learning) explaining activities have gained continuous interest and are frequently applied as consolidation activities to strengthen knowledge after a lesson. It is an open question, however, whether these explaining activities are beneficial for school students. On the one hand, the empirical basis of generating explanations is indeed steadily increasing. For instance, results of a recent meta-analysis demonstrated a small but significant explaining effect regarding comprehension (g = 0.22, k = 18 comparisons; Lachner et al., 2021). The meta-analytic findings, however, also documented large variability among studies, as there were positive, negative, and null findings. Part of this variance could be explained by the type of explanatory implementation, such as their modality, as the authors found stronger effects of oral explaining than for writing explanations (Hoogerheide et al., 2016; Jacob et al., 2020). Relatedly, the effectiveness of explaining may not only depend on differences in implementation, but also on different student characteristics. Given that a growing body of empirical literature documented the moderating role of student characteristics regarding the effectiveness of instructional interventions (Hoogerheide, Renkl, et al., 2019; Jacob et al., 2020; Lachner et al., 2021; Roelle & Renkl, 2020), the variability of previous findings in the context of explaining may also depend on interindividual student differences, emphasizing the need to consider these characteristics in future studies.

It is difficult to decide whether explaining is beneficial, particularly for school students, as most of the research on explaining is almost exclusively restricted to adult students (university students). The latter students are likely to possess higher levels of explaining skills, as well as cognitive and metacognitive prerequisites (Brod, 2020). Therefore, it is unclear whether explaining is effective for different student populations, particularly for younger students, such as school students (see Hoogerheide, Visee, et al., 2019, for an exception), especially as recent researchers suggest that the benefits of generative activities depend on developmental differences (Brod, 2020).

In this paper, we provide an empirical attempt to resolve the previously mentioned issues. First, we investigated the effectiveness of technology-mediated explaining in a secondary school teaching context in which students attended a scripted biology lesson on photosynthesis. Second, we investigated the moderating role of explanatory implementation and students' characteristics: Considering potential effects of implementation (explaining modality), we implemented two versions of explaining activities in which students provided an explanation in written or oral form to a fictitious student via a mock-up instant messaging service. We followed a non-interactive explaining approach (Lachner et al., 20xxb; Fiorella and Mayer, 2013; Hoogerheide et al., 2014a; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019) to deliberately trace back potential effects of our interventions to the act of explaining, and not to subsequent learning activities (Chi and Wylie, 2014). Additionally, we investigated the moderating role of student' prerequisites. We relied on two common prerequisites, namely, prior knowledge and academic self-concept, which have frequently shown to be highly predictive for students' achievement (Hoogerheide, Renkl, et al., 2019; Huang, 2011; Kalyuga, 2007; McNamara et al., 1996; McNamara & Scott, 1999; Möller et al., 2020; Richter et al., 2016, 2018; Roelle & Renkl, 2020).

10.2.1 Generating Explanations to Fictitious Others

Generating an explanation to a fictitious person is a generative learning strategy that enhances students' learning (Fiorella & Mayer, 2013, 2014, 2016; Plötzner et al., 1999; Roscoe & Chi, 2008). In line with generative learning theory, explaining triggers students to select and organize relevant information of the provided materials (Mayer, 1996; Wittrock, 2010). To explain the previously learned contents, students need to elaborate the contents by connecting new information with their existing knowledge which ideally results in new knowledge structures (Fiorella & Mayer, 2014, 2016; Pilegard & Fiorella, 2016; Wittrock, 2010). Besides these cognitive strategies, students need to use metacognitive strategies as well, as they have to monitor the accuracy of their current comprehension (cf. monitoring accuracy, see Fukaya, 2013). Accurate monitoring allows students to find knowledge gaps and to choose learning strategies accordingly, for instance, by restudying specific information (Fiorella & Mayer, 2016; Kiewra, 2005).

In contrast to these theoretical assumptions, prior studies have demonstrated only mixed findings concerning the effectiveness of generating explanations to fictitious others on students’ understanding and monitoring accuracy. On the one hand, previous research has indicated positive effects of explaining. As an example, Fiorella and Mayer (2014, Experiment 2) conducted a study with a 2 × 2 design with expectation (i.e., to be tested versus to explain) and activity (i.e., explaining versus no explaining) as within-subject factors. University students studied a text (topic: doppler effect), expecting to either be tested or to generate an explanation afterwards. Then, they either explained the contents to a fictitious peer or restudied the contents. Results revealed that students who explained outperformed students who restudied the contents regarding comprehension (d = 0.56). On the other hand, even though positive effects of explaining were demonstrated in several studies (Fiorella & Mayer, 2013; Hoogerheide et al., 2014; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019), there is a growing countercurrent of research revealing null or even negative findings of explaining. For instance, Lachner et al. (2021) conducted two experiments in which university students explained provided materials either to themselves or to a fictitious peer. In a third condition, students retrieved the contents from memory (control condition). Results of both experiments revealed no differences between explaining to a fictitious peer and retrieving regarding comprehension (see Fiorella et al., 2017; Fukaya, 2013; Hoogerheide et al., 2016, for similar findings).

The same holds true for students’ monitoring. Jacob et al. (2020), for instance, investigated potential effects of explaining on monitoring accuracy (correspondence between judgement and actual performance) by conducting a study in which university students studied a text and either provided an explanation in written versus oral form, or retrieved the information. The findings showed that oral explaining contributed more to monitoring accuracy than retrieval or writing explanations (η_p^2 = .03; see Fukaya, 2013, for similar effects). In contrast, Lachner et al. (2020) conducted a similar study in which university students also studied a text (topic: combustion engines) and then either generated an explanation (between versus after paragraphs) or retrieved the materials. Contrary to their hypothesis that explaining would improve monitoring accuracy, the authors did not find a difference among conditions (see Lachner et al., 2021, for similar findings).

10.2.2 Boundary Conditions of Technology-mediated Explaining

The mixed findings of previous studies regarding comprehension and monitoring suggest that explaining is not necessarily an effective strategy; rather, its effectiveness depends on different boundary conditions. Recent research started to discuss both, implementation-related and student-related characteristics as potential boundary conditions of explaining (Hoogerheide, Renkl, et al., 2019; Jacob et al., 2020; Lachner et al., 2021; Roelle & Renkl, 2020). We are aware that there may be an infinite set of potential boundary conditions (see Chi et al., 1994; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019, for further boundary conditions). We, therefore, restricted the literature review to three central boundary conditions which are prominent examples in educational research.

10.2.2.1 Modality as Implementation-Related Boundary Condition

Due to the advent of recent technologies, there are many possibilities to generate explanations. A central common principle of technology-mediated explaining is that students make active use of predominantly *verbal* media. For instance, students explain via audio messages or videos. Prior research indicated no differences among these verbal forms of explaining (video-based versus audio-based explaining; Waldeyer et al., 2020; Wassenburg et al., sub.). Interestingly, recent research revealed that most of the differences of explaining depend on the explanatory *modality* (i.e., oral versus written explanations) as an implementation-related boundary condition. For instance, Hoogerheide et al. (2016) investigated whether writing an explanation was more beneficial than restudying. In their experiment, university students studied a text on syllogistic reasoning, and afterwards either restudied the material or wrote an explanation to a fictitious peer. In contrast to prior research, results showed no differences between the learning tasks (restudying versus explaining), neither regarding the text-based test (d = 0.07) nor the inference test (d = 0.14). To investigate whether the null effects resulted from the modality of writing, the authors conducted a second study to directly compare the impact of generating oral versus written explanations. Results demonstrated that explaining orally (d = 0.43), but not explaining in written form (d = 0.19) was more beneficial than restudying the material. The difference between both modalities was not significant (d = 0.24). Relatedly, Lachner et al. (2018) provided university students with a text on combustion engines. Afterwards, they randomly explained the contents orally or in written form to a fictitious student. Interestingly, and in contrast to Hoogerheide et al. (2016), analyses showed that explaining in oral form resulted in higher learning outcomes than writing an explanation (*d* = 0.67). Jacob et al. (2020) resolved these seemingly contradictory findings, as they showed that the benefit of oral explaining predominantly holds true only for complex learning materials. Thus, the students were provided with a low-complex or a high-complex text and were asked to explain the contents in oral or written form. Again, a control group retrieved the materials. The authors showed that students benefited more from explaining orally than from writing explanations when studying high-complex text (*d* = 0.32), but not when studying low-complex text (*d* = 0.22); students who retrieved the materials showed the lowest performance (see Lachner et al., 2021, for meta-analytic evidence).

10.2.2.2 Student-Related Boundary Conditions

Besides implementation-related boundary conditions, students' prerequisites might also determine the effectiveness of explaining. In instructional literature, *prior knowledge* is discussed as a critical cognitive student characteristic (Kalyuga, 2007; McNamara et al., 1996; Richter et al., 2016, 2018). Prior knowledge commonly enhances processing of new information, and likewise reduces students' cognitive demands to regulate their learning (Mayer, 2009; Richter et al., 2016, 2018). In this context, prior research demonstrated that additional instructions, such as asking to provide an explanation, are particularly beneficial for students with low prior knowledge (e.g., McNamara & Scott, 1999). It can be assumed that instructional interventions (e.g., explaining) may act as support especially for low prior knowledge students, as it may guide their learning processes and help them realize a higher level of (meta-)cognitive processes during explaining. Learners with high prior knowledge contrarily may have ample cognitive resources available to realize adequate learning activities on their own (Kalyuga, 2007). Empirical evidence on the effect of prior knowledge in the context of explaining to fictitious others, however, is scarce. The only empirical example is the study by Hoogerheide, Renkl, et al. (2019) who contrasted an explaining condition to a restudy condition in the domain of electrical troubleshooting tasks. Results showed that students who explained outperformed students who restudied regarding comprehension (η_p^2 = .07). An additional analysis revealed that the explanatory effect was qualified by a significant moderation effect of students' prior knowledge, as in line with the aforementioned reasoning only students who had lower prior knowledge profited more from explaining than students who had higher prior knowledge.

As a second student-related boundary condition, prior research discussed students' *academic self-concept* as a crucial motivational prerequisite that determines learning.

Academic self-concept is regarded as one’s own self-perception of the own competence in a specific academic domain (Marsh et al., 2017; Shavelson et al., 1976). Several studies indicated that students’ academic self-concept is positively correlated with learning outcomes and, therefore, is crucial to consider during learning activities (see Huang, 2011; Möller et al., 2020, for meta-analytic evidence). An empirical example for the moderating role of self-concept in the context of instructional interventions is the recent study by Roelle and Renkl (2020). The authors examined effects of guidance with high school students by providing constrained or permanent access to instructional explanations while learning from direct instruction. Results showed that the effectiveness of guidance depended on students’ academic self-concept, but not on their prior knowledge, as only students with low self-concept profited from the guidance function. Whether these moderating effects replicate in the context of explaining to fictitious others, however, is still an open question.

10.2.3 Developmental Perspective on Learning-by-Explaining

Judging whether explaining is beneficial or not, and which boundary conditions may constrain its effectiveness, requires considering developmental differences of the participating students. First, from a cognitive perspective, school students have fewer cognitive and metacognitive resources available, which constrains the effectiveness of generative activities (Breitwieser & Brod, 2021; Brod, 2020). For instance, students’ working memory and inhibitory capacities are still in development (Best & Miller, 2010; Schneider, 2015), which are integral part of shifting between mental states and tasks (Best & Miller, 2010; Brod, 2020). The lower levels of task switching may limit the potential reasoning abilities (e.g., elaboration), and as such reduce the successful application of generative activities, such as explanations (Bascandziev et al., 2018; Brod, 2020; Brod & Breitwieser, 2019; Zaitchik et al., 2014).

Second, from an expertise development perspective, successful explaining requires professional experience and deliberate practice (Ericsson et al., 2006). Thus, it cannot be assumed that students possess the required skills to generate high-quality explanations, which may in turn reduce the potential benefits for explaining. These negative effects may be even more pronounced for written explanations, as writing commonly requires students to invest more cognitive effort than producing oral language since students first need to structure their thoughts to then externalize them into the writing system; in contrast, speaking is a rather automatic and less organized process (Bourdin & Fayol, 1994, 2002; Grabowski, 2010;

Liberman, 1994). These findings highlight the need to replicate previous findings of explaining with school students.

10.2.4 The Present Study

The aims of the current experiment were two-fold. First, we aimed at replicating whether technology-mediated explaining (explaining in a messenger chat) is an effective generative activity for school students, as previous research most exclusively relied on university samples. Second, we investigated implementation-related (modality) and student-related (prior knowledge, academic self-concept) boundary conditions of technology-mediated explaining, as previous research suggested considerable heterogeneity among studies.

We conducted a field experiment with students of secondary biology classes. Before the students attended a scripted lesson on photosynthesis, we assessed their prerequisites (prior knowledge, academic self-concept). After the lesson, we realized three different consolidation activities. In two explaining conditions, students generated a technology-mediated explanation, as they were asked to generate an explanation about the contents to the fictitious peer Lisa in a mock-up instant messenger in either orally or in written form (Figure 14). Students in the control condition simply retrieved the contents. The study was conducted in the TüDiLab (Tübingen Digital Teaching Lab, see https://www.tuedilab-tuebingen.de/), a classroom, which is equipped with technological infrastructure (e.g., internet, notebooks, tablets). Given that the technological infrastructure largely varies among German schools (Eickelmann et al., 2019), using the TüDiLab allowed us to conduct the study independently of the infrastructure of the particular school, and to keep the experimental sessions as comparable as possible. We examined the following research questions:

10.2.4.1 RQ 1: Is Explaining Effective for School Students?

Given that previous evidence regarding the benefits of explaining in school contexts is rare (Hoogerheide, Visee, et al., 2019), and recent researchers plead for replications in different study samples (Breitwieser & Brod, 2021; Brod, 2020), we investigated whether explaining is effective for school students. Based on previous findings, we hypothesized that students who explain (in oral or in written form) would outperform students who retrieve the material a) on the comprehension tests (Fiorella & Mayer, 2013, 2014; Hoogerheide, Visee, et al., 2019), and b) demonstrate more accurate monitoring judgments (Fukaya, 2013; Jacob et al., 2020).

10.2.4.2 RQ 2: Does the Implementation of Explaining Activities Constrain the Effectiveness of Explaining?

Technologies such as instant messengers allow generating both written and oral explanations. Given that previous evidence particularly suggested differences between oral and written modality (Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018), we tested whether students who generate oral explanations in the instant messenger would outperform students who write an explanation a) on the comprehension tests (Hoogerheide et al., 2016; Lachner et al., 2018), and b) on their monitoring accuracy (Jacob et al., 2020). These analyses would allow to examine whether modality of explaining is a critical boundary condition of explaining also for school students.

10.2.4.3 RQ 3: Do Students’ Prerequisites Affect the Effectiveness of Learning-by-Explaining?

Based on previous evidence (Hoogerheide, Renkl, et al., 2019), we argue that the act of generating explanations might be a particularly adequate instructional support for students with lower cognitive and motivational prerequisites, as the act of explaining would trigger students to enact generative activities, which they would not realize spontaneously. Therefore, we explored whether prior knowledge and self-concept moderate the effect of explaining.

10.2.4.4 Further Explorative Analyses

Based on previous studies, we explored three characteristics of students’ generated explanations to explore the underlying mechanism of explaining, (Chafe & Tannen, 1987; Hoogerheide et al., 2016; Lachner et al., 2018). First, we examined whether oral explaining would trigger students to perceive social presence more strongly compared to students who write an explanation. We used personal references within students’ explanations which is considered as an indicator for perceived social presence. We further counted mentioned concepts as a proxy for the comprehensiveness of the explanations (Boshuizen & Schmidt, 1992; Jacob et al., 2020), and the number of elaborations for generative processes (Chafe & Tannen, 1987; Lachner et al., 2018). Both processes should increase students’ comprehension. Furthermore, we analyzed students’ self-reported mental effort and their perceived difficulty as subjective measures of cognitive load during the learning activity.

10.3 Method

10.3.1 Participants and Design

In total, seven 7th grade classes (Gymnasium) of a district in southwestern Germany participated in the study (N = 132). The first class was a pilot study (n = 22). After this pilot study, we shortened the knowledge tests and the lesson. Then, we collected data from six further classes. However, since the pilot class showed comparable characteristics and learning outcomes as the other classes, we integrated these students in our analyses, nevertheless. Three students completed the lesson but did not participate in the learning activity, and therefore were excluded from our analyses. In sum, we included data from 129 students in the final analysis. We conducted a power analysis before recruiting students (GPower, Version 3.1.9.2). We analyzed the required sample size for contrast analyses and assumed an effect size of η_p^2 = .05, as the only study with school students revealed medium effects of generating explanations (Hoogerheide, Visee, et al., 2019). We set power to .80, and the α-error to .05. Thus, our sample reached the required size of 128 students.

On average the students were 12 years old (SD = 0.54), and 57 % were female. The majority stated to be a German native speaker or grew up bilingually (85 %), and few students had another native language than German (15 %). Their average biology grade was 2.00 (SD = 0.67, on a scale from 1 "very good" to 6 "failed") in their last report cards. Additionally, students had medium prior knowledge regarding photosynthesis (M = 2.40, SD = 1.23; on a scale from 0 to 5 points) and possessed an average academic self-concept in biology (M = 2.86; SD = 0.46; on a 4-level Likert scale from 1 "very low" to 4 "very high").

Within each class, students were randomly engaged in one out of three learning activities (oral explaining condition: n = 41; written explaining condition: n = 45; retrieval condition: n = 43).

10.3.2 Measures

Apart from the learning activity, data collection was paper-based. For completing the learning activity, students used a laptop (see Procedure for details). Since the study was integrated in a larger research project, we refrained from reporting the entire set of variables, but only those of interest for the current study. A list of the entire variables of the study is provided in Appendix A.

10.3.2.1 Academic Self-Concept

As a potential moderator variable, we measured academic self-concept in biology by using 11 items (e.g., "I learn new biological topics quickly", McDonald's $\omega_t = .88$), which addressed competence and affect components (based on OECD, 2016). Students rated their self-concept on a Likert scale from 1 "completely disagree" to 4 "completely agree".

10.3.2.2 Prior Knowledge Test

We developed a prior knowledge test with 5 single-choice questions (e.g., "Why are plants green?"). For each correctly answered questions, students got 1 point resulting in a maximum score of 5 points. The items comprised the entire process of photosynthesis (McDonald's $\omega_t = .43$). The prior knowledge test substantially predicted students' learning outcomes (conceptual knowledge: $r(127) = .34$, $p < .001$; transfer knowledge: $r(127) = .42$, $p < .001$), suggesting high prognostic validity of our prior knowledge test measure.

10.3.2.3 Comprehension Posttest

The posttest comprised 2 subtests which measured conceptual and transfer knowledge. The conceptual test combined 9 single-choice questions (e.g., "Which variable can be used to measure photosynthesis?"). Ror each correctly answered question, students got 1 point resulting in a total score of 9 points (McDonald's $\omega_t = .74$). The transfer test combined 4 questions with an open-ended answer format (McDonald's $\omega_t = .63$). Students could get 2 points per question, resulting in a maximum score of 8 points. In total, 20 % of the open-ended answers were assessed by two independent raters. The interrater reliability was very good ($ICC_{2,1} = .90$), therefore, one rater rated the remaining answers.

10.3.2.4 Monitoring Accuracy

Students prospectively judge their assumed performance on the conceptual knowledge test to investigate their monitoring accuracy ("On the next page, we will ask you to answer 9 questions about photosynthesis. You will get 1 point for each correct answer. How many points do you think you will get?"; see Baars et al., 2017; Prinz et al., 2018) after both, the lesson and the learning activity (explaining or retrieving) on a scale from 0 – 9. In this context, we would like to highlight that the first participating class answered more questions in the comprehension test, which we excluded for the further classes due to time constraints. We adjusted their monitoring ratings to the new maximum score of the posttest. Students' monitoring accuracy was operationalized in terms of bias (Baars et al., 2017; Prinz et al., 2018). Bias reflects the difference between students' estimated performance and their actual performance ($X_{Judgment} – X_{Performance}$). Hence, this estimation indicated whether students over- or underestimate their own performance. An overestimation of performance is represented by positive values and an underestimation is reflected by negative values. An accurate judgement is represented by a value of 0.

10.3.2.5 Mental Effort and Subjective Difficulty

Students were instructed to state their perceived mental effort (i.e., "How much effort did you invest in explaining the material?"; see Paas, 1992), and the subjective difficulty (i.e., "How easy was it for you to explain the material?"; see DeLeeuw & Mayer, 2008) after the lesson and after the activity on a 9-point Likert scale.

10.3.2.6 Control measures

Self-efficacy in explaining. We assessed students' self-efficacy in explaining as a control variable. To measure their explaining skills, students were asked to rate 3 adapted items (e.g., "I always succeed in explaining complex contents when I make an effort", McDonald's ω_t = .55) from Jerusalem and Schwarzer (1999) on a 4-point Likert scale from 1 "completely disagree" to 4 "completely agree".

Engagement and situational interest during learning activity. As further control variables, we used 4 items to ask students to rate their engagement (e.g., "While explaining, I strongly focused on the task"; we excluded one item in terms of content and reliability misfit; McDonald's ω_t = .81), based on Frank and Kluge (2014), Seidel, Stürmer, Blomberg, Kobarg, and Schwindt (2011), and Vorderer et al. (2004), and 6 items to ask them to rate their situational interest (e.g., "I liked to explain the materials"; McDonald's ω_t = .92), based on Knogler, Harackiewicz, Gegenfurtner, and Lewalter (2015), during the learning activity on a 4-point Likert scale.

10.3.2.7 Characteristics of The Generated Explanations

We analyzed three characteristics of students' generated explanations based on previous research: Personal references, concepts, and elaborations (Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018).

Personal references. The number of personal references within students' explanations ("I", "you", etc.) were automatically counted with a script implemented in RStudio (R Core

Team, 2018, version 3.5.1). For instance, the sentence "[…] now *you* might think 'that is very difficult' – but it's not" contains one personal references (you). Previous research used personal references as an indicator for perceived social presence (see Chafe & Tannen, 1987; Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018, for related approaches).

Concepts. We counted mentioned concepts within the explanations as an indicator of their comprehensiveness (Boshuizen & Schmidt, 1992). Concepts are the mentioned constructs within an explanation. For instance, the sentences "*Photosynthesis* is when *plants* transform *carbon dioxide* into *oxygen*" contains four concepts (marked in italics). We used *CohViz* by Lachner et al. (2017b) to automatically analyze the concepts, which is a tool that detects concepts by means of natural language processing technologies.

Elaborations. The number of elaborations within the explanations were measured as a third characteristic. Elaboration were defined as further information, such as examples, that go beyond the provided materials (see also Fiorella & Kuhlmann, 2020; Lachner et al., 2018). For instance, the sentences "*Glass fogs up when you breath on it, because you exhale hydrogen*" is one elaboration as it was not mentioned during the lesson, and the student combined new information with prior knowledge to express this example. Elaborations for 20 % of all explanations were coded by two independent raters. Interrater reliability revealed to be excellent, $ICC_{2,1}$ = 0.94, thus, one rater assessed the remaining explanations.

10.3.3 Procedure

The experiment was incorporated in a larger research context in which the implementation of innovative instructional concepts for technology integration into teaching were examined. First, we received approval from the council of governments to conduct our study and obtained consent from students and their parents. The study was conducted in the *Tübingen Digital Teaching Lab* (www.tuedilab-tuebingen.de), a classroom with sufficient infrastructure for realizing technology-enhanced teaching. Per experimental session, one class participated. After welcoming the class, students were asked to answer the first questionnaire which included demographical questions (Appendix A), academic self-concept in biology, self-efficacy in explaining, and the prior knowledge test. Afterwards, an instructed teacher started a scripted lesson about photosynthesis which we developed in cooperation with advanced biology pre-service teachers. The same teacher taught all seven classes. After receiving instruction on the components of photosynthesis, students conducted a virtual experiment in which they investigated the relation between the central components (e.g., light intensity). The lesson took about 45 minutes. Afterwards, students stated their mental effort and difficulty during the lesson, provided a monitoring rating, and answered the conceptual knowledge posttest. After the lesson, students were randomly assigned to one of three conditions (oral explanation vs. written explanation vs. retrieval practice) and guided to three different rooms. Students worked individually with a laptop. To ensure data protection, all data were stored locally on the laptops. Students who explained the contents were linked to a mockup messenger chat (Figure 14) generated by Klemke (2017), which was an imitation of WhatsApp Messenger (created by WhatsApp Inc. in 2009 and acquired by Facebook in 2014). In this chat, students could see a profile picture and the message from a fictitious peer called Lisa who asked them for an explanation (Figure 14, see also Hoogerheide, Visee, et al., 2019; Jacob et al., 2020, for related approaches). They received the following instruction:

Here you can see a chat with Lisa. Lisa goes to school as well and is about your age. She also wanted to take part in this study but unfortunately, this was not possible. However, she would love to learn the process of photosynthesis and asks you to explain it to her. She has not read anything about the topic yet and does not know what photosynthesis is. Please explain Lisa the process of photosynthesis.

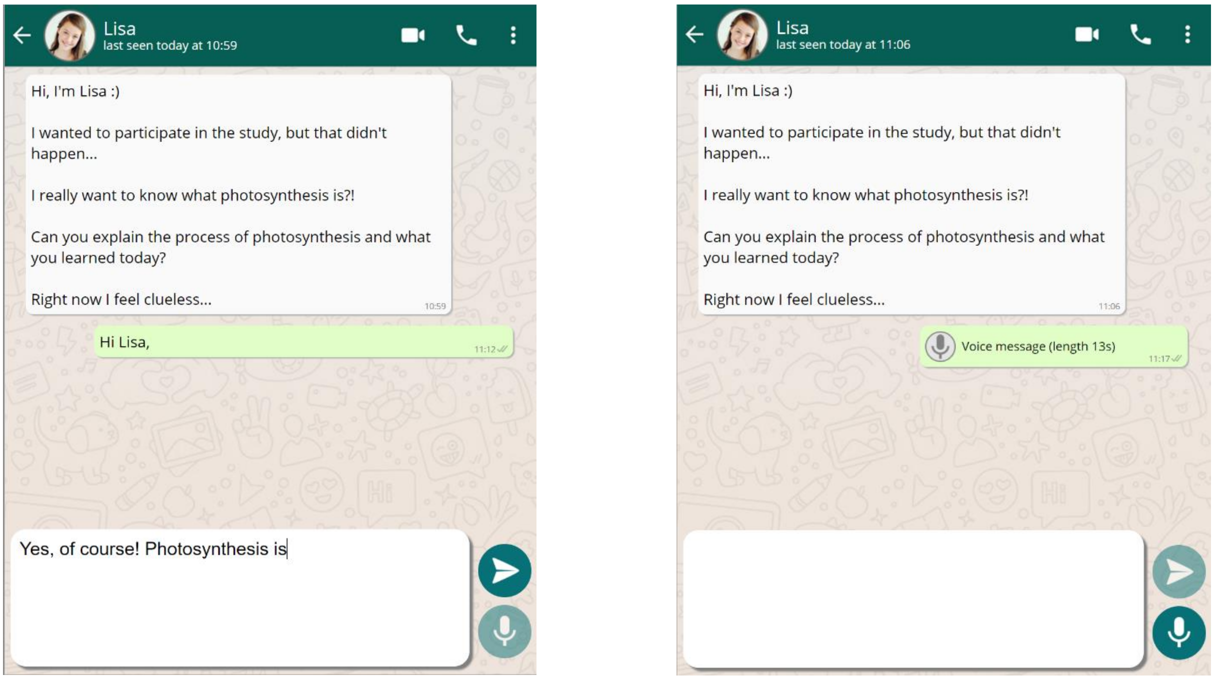

Figure 14. Simulated mockup messenger chat for explaining condition with a profile picture

and message from the fictitious peer Lisa. Students in the writing condition (left) could send text messages and students in the oral condition (right) could send voice messages.

Depending on experimental condition, students could either send a voice message (oral condition) or a text message (written condition) in which they explained the contents. Their messages appeared in the mockup chat after sending them (Figure 14). Students could send as many messages as they wanted to. As in previous studies, we followed a non-interactive scenario, in which the recipient Lisa could not respond to the explanation, to trace back potential effects of technology-mediated explaining to the act of explaining exclusively. The control group was engaged in retrieval practice and could write notes about the lesson on the laptops in a text box. Students received the following instruction:

Please retrieve the contents of the lesson. You can take notes.

Students completed the task in a self-paced manner. Afterwards, students again stated their mental effort and their subjective difficulty and judged their current comprehension. Then, they rated their engagement and situational interest during explaining or retrieving the material. Afterwards, they completed the knowledge posttest (20 minutes). Each class received 100 Euros for participating.

10.3.4 Data Analyses

In cases of direct hypotheses, contrast coding was applied to analyze our data. To test whether there is a main effect of explaining (RQ 1), we contrasted the explaining conditions (oral and written explanations) against the retrieval condition by using the following contrast weights: retrieval: -2/3; writing explanations: 1/3; oral explaining: 1/3. To test for differences between the two explaining modalities (RQ 2), we used the following contrast weights: retrieval: 0; writing explanations: -1/2; oral explaining: 1/2.

Regarding the moderation hypotheses (RQ 3), separate moderation analyses were conducted with the PROCESS macro in SPSS (Hayes, 2018). Experimental condition was the independent variable and learning outcome (conceptual knowledge; transfer knowledge) and monitoring accuracy the dependent variable. Prior knowledge and self-concept were separately included as moderator variables. As in our previous analysis, experimental condition was implemented as a contrast variable (explaining contrast: retrieval: -2/3; writing explanations: 1/3; oral explaining: 1/3; modality contrast: retrieval: 0; writing explanations: -

1/2; oral explaining: 1/2; by means of Helmert coding) to directly calculate the interactions for each contrast (see Hayes & Montoya, 2017, for more details). Moderation occurred when the interaction effect between the contrast coded predictor (explaining contrast; modality contrast) and one of the potential moderator variables (prior knowledge; self-concept) was significant. In cases of a significant interaction, the Johnson-Neyman-Technique (Hayes & Montoya, 2017) was used to identify the regions of the moderator in which the effect of the experimental condition on the conceptual questions was significant.

Since all test materials were paper-based, missing data naturally occurred. The maximum score of missing values per scale was 6 % percent and no pattern of missingness was observed. We, therefore, expected that missing values were missing at random (MAR) which allowed us to use multiple imputations (Allison, 2009; Enders, 2010; Graham, 2009; Rubin, 1976). This methodology is an alternative to maximum likelihood methodology and is a recommended technique to handle missing data (Enders, 2010). Instead of replacing missing data with a mean or median, this approach estimates missing data based on the distribution of the observed variables. We impute the missing values of the affected scale by using the R-package *mice* (van Buuren & Groothuis-Oudshoorn, 2011; version 3.8.0) to.

10.4 Results

Cohens' *d* and partial η^2_p were used to measure effect sizes (small effects: η^2_p = .01, *d* = 0.20; medium effects: η^2_p = .06, *d* = 0.50; large effects: η^2_p = .14, *d* = 0.80). Alpha level was α = .05.

10.4.1 Preliminary Analyses

We performed MANOVAs to determine potential differences among conditions before the lesson, after the lesson, during the learning activity, and after the learning activity. Results showed no differences of variables collected before the lesson (age, gender, prior knowledge, grade in biology, self-concept in biology, self-efficacy in explaining), $F(2, 126) = 1.07$, p = .385, nor after the lesson (mental effort, subjective difficulty), $F(2, 126) = 0.90$, p = .464. During the learning activity (explaining vs. retrieving), time on task differed naturally across conditions, $F(2, 126) = 31.38$, $p < .001$, η^2_p = .332: Students who wrote an explanation needed significantly more time than students who explained orally ($p < .001$) or retrieved the materials ($p < .001$). However, time did not predict the learning outcome (conceptual knowledge: $r(127) = -0.10$, $p = .265$; transfer knowledge: $r(127) = -0.03$, $p = .738$). See Table 5 for descriptive statistics. Finally, results showed no differences of variables collected after the learning activity (engagement, interest in the learning activity), $F(2, 126) = 1.01$, $p = .363$.

Table 5

Summary of Means and Standard Deviations (in Parentheses) for all Measurements

	Retrieval Practice	Written Explanation	Oral Explanation
Self-Concept in Biology (1-4)	2.88 (0.40)	2.81 (0.47)	2.93 (0.50)
Self-Efficacy in Explaining (1-4)	2.88 (0.49)	2.83 (0.45)	2.75 (0.41)
Prior knowledge (0-5)	2.56 (1.30)	2.40 (1.23)	2.24 (1.18)
Learning Outcomes			
Conceptual knowledge (0-9)	6.53 (2.22)	6.24 (2.10)	6.37 (2.09)
Transfer knowledge (0-8)	4.30 (1.95)	4.22 (1.99)	4.12 (1.87)
Perceived Cognitive Load			
Perceived mental effort (1-9)	4.98 (2.22)	5.78 (2.53)	6.27 (2.23)
Perceived difficulty (1-9)	2.77 (1.94)	4.73 (2.35)	5.66 (1.84)
Monitoring Accuracy	-0.93 (2.05)	-1.10 (2.79)	-0.98 (2.66)
Feelings while Explaining			
Engagement (1-9)	2.69 (0.58)	2.84 (0.68)	2.76 (0.68)
Interest (1-9)	2.52 (0.73)	2.83 (0.64)	2.63 (0.75)
Characteristics of Explanations			
Personal references	0.72 (0.91)	2.04 (2.39)	2.71 (3.67)
Concepts	2.44 (2.68)	8.07 (5.71)	9.68 (4.57)
Elaborations	0.02 (0.15)	0.31 (0.56)	0.17 (0.44)
Time of Learning Activity (in min.)	4.25 (2.54)	10.67 (6.39)	4.47 (2.65)

10.4.2 RQ 1: Is Explaining Effective for School Students?

First, we tested for potential differences between the explaining conditions (oral versus written explanations) and the retrieval condition. Regarding students' learning outcomes, the contrast showed not to be significant, neither for conceptual knowledge, $F(1, 127) = 0.18$, $p = .669$, $\eta_p^2 = .001$, nor for transfer, $F(1, 127) = 0.18$, $p = .674$, $\eta_p^2 = .001$. Regarding students' monitoring accuracy, the contrast analysis revealed also no significance, $F(1, 126) = 0.03$, $p = .865$, $\eta_p^2 < .001$ (we additionally controlled for monitoring judgments during the lesson to consider for students' intra-individual differences; see Hertzog et al., 2013; Lachner et al., 2020). Therefore, our hypothesis was not confirmed, as explaining did not contribute more to learning outcomes and monitoring accuracy than retrieving the material.

10.4.3 RQ 2: Does the Implementation of Explaining Activities Constrain the Effectiveness of Explaining?

Next, we tested for potential differences within the explaining conditions. Following previous evidence (Hoogerheide et al., 2016; Jacob et al., 2020; Lachner et al., 2018), we tested whether oral explaining contributed more to (meta-)comprehension than writing explanations. Again, we did not obtain significant findings regarding conceptual knowledge, $F(1, 127) = 0.08$, $p = .781$, $\eta_p^2 = .001$, or their transfer, $F(1, 127) = 0.05$, $p = .818$, $\eta_p^2 < .001$. Relatedly, the contrast was also not significant for monitoring accuracy, $F(1, 126) = 0.07$, $p = .787$, $\eta_p^2 < .001$. Apparently, the modality of explaining did not constitute a boundary condition of students' learning in the present study.

10.4.4 RQ 3: Do Students' Prerequisites Affect the Effectiveness of Learning-by-Explaining?

To examine our research questions regarding student-related boundary conditions, we explored interaction effects between the learning activities, and prior knowledge and their academic self-concept. Table 6 shows the entire analysis.

10.4.4.1 Prior Knowledge

None of the moderation analyses approached significance, neither for learning outcomes nor for their monitoring accuracy (Table 6), indicating that prior knowledge had no impact on the effectiveness of the learning activities.

Table 6

Summary of the Moderation Analyses with Prior knowledge and Academic Self-Concept regarding Learning Outcome and Monitoring Accuracy

	b	*se*	*t*	*p*
Prior knowledge as Moderator				
Conceptual knowledge				
Explaining contrast	-0.76	0.84	-0.91	.364
Modality contrast	-0.34	0.96	-0.35	.725
Prior knowledge	0.48	0.18	2.55	**.012**
Interaction (Explaining contrast × Prior knowledge)	0.34	0.30	1.13	.261
Interaction (Modality contrast × Prior knowledge)	0.23	0.37	0.62	.543
Transfer knowledge				
Explaining contrast	-0.20	0.74	-0.27	.787
Modality contrast	-0.53	0.84	-0.63	.529
Prior knowledge	0.64	0.16	3.95	**.001**
Interaction (Explaining contrast × Prior knowledge)	0.07	0.27	0.25	.803
Interaction (Modality contrast × Prior knowledge)	0.23	0.32	0.71	.481
Monitoring Accuracy				
Explaining contrast	-0.39	0.87	-0.44	.658
Modality contrast	-0.20	1.00	-0.20	.845
Prior knowledge	-0.06	0.19	-0.34	.736
Interaction (Explaining contrast × Prior knowledge)	0.18	0.31	0.56	.579
Interaction (Modality contrast × Prior knowledge)	-0.11	0.38	-0.28	.783
Monitoring accuracy after lesson	0.69	0.09	7.67	**.001**
Academic Self-Concept as Moderator				
Conceptual knowledge				
Explaining contrast	-6.34	2.58	-2.46	**.015**
Modality contrast	1.76	2.62	0.67	.502
Academic Self-concept	0.83	0.45	1.84	.068
Interaction (Explaining contrast × Academic self-concept)	-2.28	0.89	2.56	**.012**
Interaction (Modality contrast × Academic self-concept)	-0.61	0.90	-0.67	.502
Transfer knowledge				
Explaining contrast	-3.53	2.31	-1.52	.130
Modality contrast	-2.17	2.35	-0.92	.357
Academic Self-concept	1.22	0.40	3.04	**.003**
Interaction (Explaining contrast × Academic self-concept)	1.28	0.80	1.60	.111
Interaction (Modality contrast × Academic self-concept)	0.67	0.81	0.83	.408
Monitoring Accuracy				
Explaining contrast	0.95	2.69	0.35	.724
Modality contrast	-4.00	2.71	-1.47	.143

Academic self-concept	-0.18	0.47	-0.39	.701
Interaction (Explaining contrast × Academic self-concept)	-0.30	0.93	-0.33	.744
Interaction (Modality contrast × Academic self-concept)	1.25	0.93	1.34	.182
Monitoring accuracy after lesson	0.69	0.09	7.71	**.001**

10.4.4.2 Academic Self-Concept

We proceeded similarly for academic self-concept (Table 6). We did obtain a significant interaction effect between the explaining contrast and academic self-concept (b = 2.28; p = .012), indicating that academic self-concept differentially affected the effect of generating an explanation (in oral or in written form) versus retrieving the material. The interaction between the modality contrast and academic self-concept was not significant (b = -0.61; p = .502).

The Johnson-Neyman-Technique (Hayes & Montoya, 2017) revealed that students with a self-concept below 2.11 (on a scale from 1 – 4) benefited more from generating an explanation (in written or in oral from) than retrieving, whereas for students with a self-concept greater than 3.20 benefited more from retrieving the materials compared to generating an explanation (see Figure 15 for the significant areas). Interestingly, there was a full reversal effect, as explaining was only beneficial for students had low self-concepts but detrimental for students who had high self-concepts regarding their acquisition of conceptual knowledge. Regarding transfer knowledge and monitoring accuracy, results showed no significant interaction.

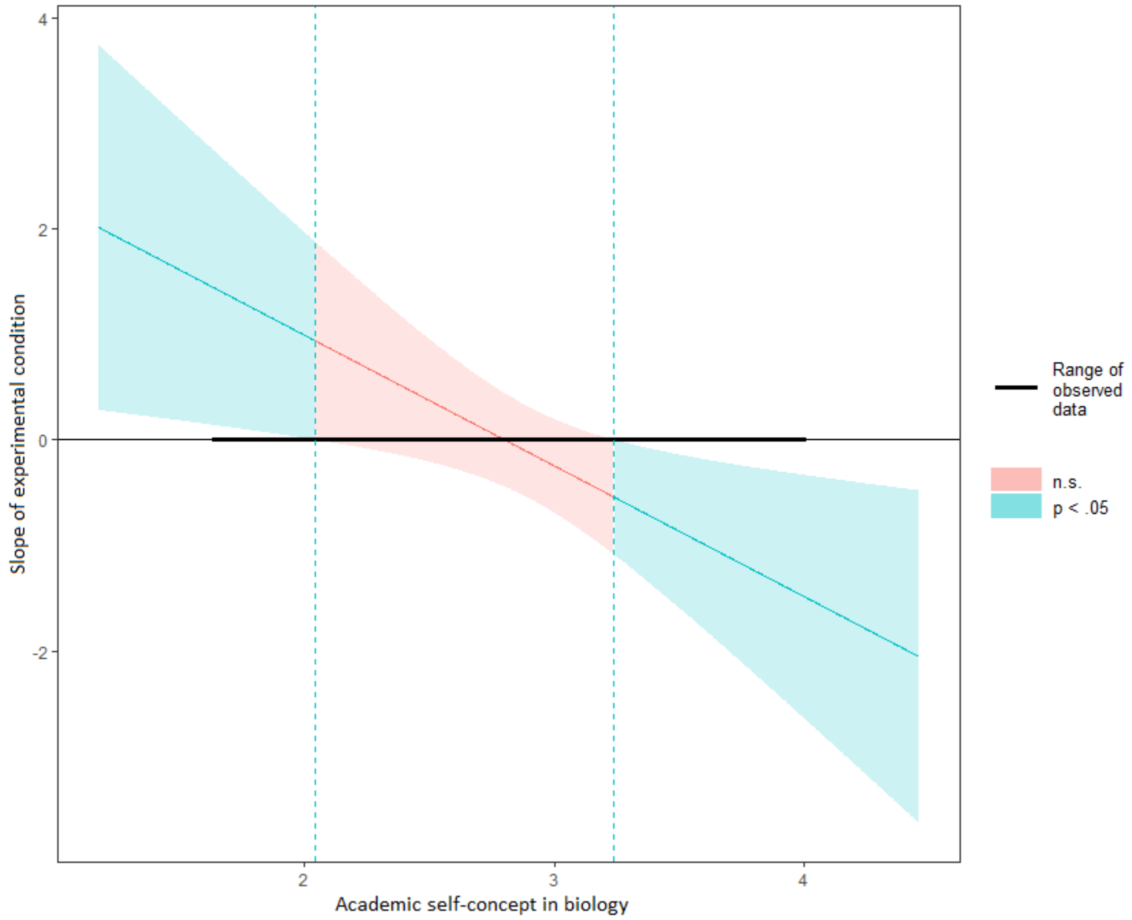

Figure 15. Confidence interval of Johnson-Neyman-Technique regarding students' conceptual knowledge. Effect of the intervention is represented in unstandardized beta coefficients. The band reflects the 95% confidence interval which can be used to demonstrate significant regions. When the confidence intervals do not include zero (as indicated by the horizontal line), the moderation effect of self-concept was significant (i.e., green band). Red bands mark non-significant areas as the confidence intervals include zero.

10.4.5 Further Explorative Analyses

10.4.5.1 Characteristics of The Explanations

We analyzed three characteristics (personal references, concepts, elaborations) of the generated written and oral explanations to investigate potential differences among the explaining tasks. Thus, these analyses were based on data from 86 students. See Table 5 for descriptive statistics. Correlations with learning outcomes can be found in Table 7. Results indicated no differences among explanatory modality regarding personal references, $F(2, 84) = 1.00$, $p = .320$, $\eta_p^2 = .012$, number of concepts, $F(2, 84) = 2.07$, $p < .154$, $\eta_p^2 = .024$, and number of elaborations, $F(2, 84) = 1.66$, $p = .202$, $\eta_p^2 = .019$, indicating that students generated comparable explanations regarding these characteristics. Self-concept did not moderate these findings (Appendix B).

Table 7

Correlations Between Characteristics of the Generated Explanations and Students’ Outcomes across Conditions

	Personal References	Concepts	Elaborations
Conceptual knowledge	(-.05|.08|-.01)	(.37*|.09|.29*)	(.10|.19|-.02)
Transfer knowledge	(.01|.12|-.13)	(.27|.44*|.18)	(.06|.37*|.16)
Perceived mental effort	(-.52*|-.02|.11)	(.01|.07|.16)	(.21|.22|-.26)
Perceived difficulty	(-.22|.07|.07)	(-.15|-.05|-.04)	(-.07|-.06|.17)
Monitoring accuracy	(-.23|-.11|-.19)	(-.19|.12|-.17)	(-.01|-.08|-.09)
Perceived presence	(.07|.04|.25)	(.36*|.24|.24)	(.29|.03|-.17)
Interest	(.11|.12|.27)	(.19|.45*|.21)	(.28|.34*|-.22)
Personal References	-	(.28|.32*|.41*)	(.05|.38*|-.09)
Concepts	-	-	(.56*|.51*|.15)
Elaborations	-	-	-

Note. Numbers in brackets represent the correlations separated for experimental conditions: Left = retrieval condition; middle = writing condition; right = oral condition. *indicates p < .05.

10.4.5.2 Mental Effort

To estimated differences in mental effort during the learning activity (explaining; retrieving), we used an ANOVA with the respective ratings during the learning task as dependent variable and experimental condition as independent variable. Results indicated no main effect of experimental condition, $F(2, 123) = 2.85$, $p = .062$, $\eta_p^2 = .044$, indicating that students rated their effort during the learning task comparably across the conditions.

10.4.5.3 Subjective Difficulty

Similar to the previous analysis on mental effort, we conducted an ANOVA to analyze students’ perceived subjective difficulty during the learning activity. Results indicated a main effect of experimental condition, $F(2, 123) = 20.08$, $p < .001$, $\eta_p^2 = .246$. Post-hoc comparisons (Scheffé) indicated that students reported that explaining (orally, $p < .001$; and in written form, $p < .001$) was more difficult than retrieving the material. The difficulty was perceived to be comparable between oral and written explanations ($p = .105$).

10.5 Discussion

The aims of the current study were to investigate a) whether previous findings on the effectiveness of generating explanations would replicate with secondary school students, as well as b) potential implementation-related and student-related boundary conditions of generating explanations. As such, we investigated the explanatory modality (i.e., oral versus written explanations), and prior knowledge and academic self-concept as crucial boundary conditions. Our findings demonstrated no differences among conditions. Therefore, we had to reject the first and second hypothesis. These results highlight two central aspects: First, technology-mediated explaining is not necessarily an effective learning strategy for younger populations, such as school students. This result appears to be in contrast to the recent study by Hoogerheide, Visee, et al. (2019) at first glance. However, we want to note that the control conditions largely differed. Whereas Hoogerheide, Visee, et al. (2019) realized a plain control condition (restudying the contents), we implemented retrieval practice, which is commonly regarded as powerful consolidation activity (see Fiorella & Mayer, 2016; Waldeyer et al., 2020, for an overview). Retrieval practice is an effective strategy for students, as retrieval supports learning since it enhances memory, provides access to retrieval cues (Rowland, 2014), and enables students to construct new retrieval cues (i.e., spreading activation, see

Carpenter, 2009; Endres et al., 2017; Roediger & Butler, 2011; Rowland, 2014). Thus, our study extends prior research as results indicate that explaining might be more beneficial for school students than restudying materials, but not compared to another generative learning strategy, such as retrieval practice. Second, our results revealed that technology-mediated explaining is not necessarily an effective learning strategy for school students, since retrieval was as effective as explaining regarding both, learning outcomes and monitoring accuracy. Additionally, students in the explaining conditions perceived higher levels of difficulty, as well as required more time to generate the explanations, as compared to retrieval practice. These findings suggest that if learning environments have distinct time constraints, realizing retrieval practice may be the more efficient strategy.

Additionally, results revealed no main effect of modality, as generating written versus oral explanations was comparably effective, suggesting that the modality was no main boundary condition of explaining for school students. In contrast to previous studies, we implemented our explaining condition in an instant messaging mock-up chat, which likely inclined the social presence of a fictitious audience, as represented by the comparable number of personal references within students' explanations. The comparable effects of social presence could have reduced differences between written and oral modality, and result in comparable learning outcomes.

Interestingly, our analysis of student-related boundary conditions revealed that the effectiveness of technology-mediated explaining strongly depended on academic self-concept, particularly for the acquisition of conceptual knowledge. Students with low self-concept benefited more from generating explanations than retrieving the contents. Contrarily, students with high self-concept profited more from retrieving than explaining the contents. Which mechanism underlies this reversal effect deserves more attention, however. We attribute this finding to the alleged presence of a social audience, which likely inclined students to make upward and downward comparisons (e.g., Becker & Neumann, 2016; Marsh et al., 2008; Trautwein et al., 2006) during explaining. Students were instructed to explain the contents to a fictitious student Lisa having low prior knowledge about this topic. Students with low self-concept benefitted more from explaining than retrieval. Thus, explaining might trigger a social situation of productive knowledge exchange because both, Lisa and the explainer had comparable prerequisites. This might incline a situation of information sharing (Ray et al., 2013). Students with high self-concept, however, may no longer profit from explaining because they explained the contents to a less knowledgeable peer. This may create a social situation implying downward comparisons, which have been demonstrated to yield lower levels of information sharing (Ray et al., 2013), and likely lower learning outcomes. Future research is needed that includes students’ perceived information sharing as additional measures to provide empirical evidence for this assumption.

A further unexpected result was that prior knowledge did not moderate the explaining effect on comprehension. This finding contradict previous findings by Hoogerheide et al. (2019), which showed that explaining was effective for students with low- but not high prior knowledge. In contrast to the students by Hoogerheide et al. (2019), who possessed very low prior knowledge (14 % correct in the prior knowledge test), our students had more prior knowledge (25% correct in the prior knowledge test), as we decided to use a curricular learning topic, namely, photosynthesis. Thus, we rather collected data from advanced school students and not from novice students, which might make it difficult to obtain potential effects on prior knowledge. Hence, the definition of low prior knowledge is always relative to the sample investigated, which is a generic problem in research on aptitude-treatment interactions.

10.5.1 Study Limitations and Future Research

One limitation related to measuring characteristics of students’ explanations as potential proxy for underlying mechanisms of the learning activity (Fiorella & Kuhlmann, 2020; Jacob et al., 2020; Lachner et al., 2018). Therefore, we did not have any online measures which could be crucial to understand the underlying processes while generating explanations. It could be that students in the writing condition did more elaborations that they, however, did not externalize in writing. Future research should consider alternative approaches to elicit the underlying cognitive and metacognitive processes during explaining, for instance, by means log-files (Hoogerheide, Renkl, et al., 2019).

Additionally, we aimed at testing our hypotheses in a practice-oriented learning environment without sacrificing internal validity. We scripted the lesson and conducted the experiment in our laboratory classroom to avoid potential confounding variables, and to keep the procedure comparably across classes. However, this procedure likely reduced the external validity as the standardization might have affected students’ behavior and learning process. Future research is, therefore, should replicate these findings in more authentic settings.

Lastly, we would like to point out that our results regarding moderation effects were based on explorative analyses. Thus, future studies need to replicate these results through hypotheses testing. Additionally, since academic self-concept moderated the explaining effect, further research should analyze its causality on learning outcome in more detail, for instance, by systematically inducing prior knowledge and self-concept.

10.5.2 Conclusion

To our knowledge, this study was the first to investigate implementation- and student-related boundary conditions of technology-mediated learning-by-explaining in a field experiment with secondary school students. Our results indicated that technology-mediated explaining is not an efficient learning activity per se, but rather depends on motivational student characteristics, that is, their academic self-concept. These findings complement prior research as they demonstrate that not only implementation-related (modality) but also students-related (academic self-concept) factors may determine the effectiveness of generative learning activities.

10.6 Appendix

10.6.1 Appendix A

Overview of the Overall Dataset

Variable	Number/type of items	Used in present paper	Used in AUTHORS (20xx)
Demographical data			
Age	1 item, open answer format	×	×
Sex	1 item, closed answer format	×	×
Mother language	1 item, open answer format	×	×
Grades in biology			
Last report	1 item, open answer format	×	×
Last class test	1 item, open answer format	×	×
Grades in German			
Last report	1 item, open answer format	×	×
Last class test	1 item, open answer format	×	×
Perceived ratings			
Academic self-concept in biology	11 items, rated on a 4-point Likert scale	Moderator variable	Independent variable
Use of digital media for learning	5 items, rated on a 4-point Likert scale	-	×
Self-efficacy of digital media	5 items, rated on a 4-point Likert scale	-	×
Self-efficacy in explaining	3 items, rated on a 4-point Likert scale	×	-
Prior knowledge	5 items, single-choice-format	Moderator variable	Independent variable

Quality of lesson			
Classroom management	8 items, rated on a 4-point Likert scale	-	×
Cognitive activation	8 items, rated on a 4-point Likert scale	-	×
Constructive support	8 items, rated on a 4-point Likert scale	-	×
(Meta-)Cognitive judgements after lesson			
Perceived mental effort	1 item, rated on a 9-point Likert scale	×	
Perceived difficulty	1 item, rated on a 9-point Likert scale	×	Dependent variable
Monitoring rating	1 item, rated from 0 to 9 points	Control variable	Dependent variable
Conceptual knowledge posttest after lesson	9 items, single-choice-format	-	Dependent variable
(Meta-)Cognitive judgements after learning activity			
Perceived mental effort	1 item, rated on a 9-point Likert scale	×	-
Perceived difficulty	1 item, rated on a 9-point Likert scale	×	-
Monitoring rating	1 item, rated from 0 to 9 points	Dependent variable	-
Perceived feelings during learning activity			
Perceived engagement	1 item, rated on a 4-point Likert scale	-	-
Perceived interest	1 item, rated on a 4-point Likert scale	-	-
Posttest after learning activity			
Conceptual knowledge test	9 items, single-choice-format	Dependent variable	-
Transfer test	4 items, open answer format	Dependent variable	-

Note. All variables are listed in the correct chronological order.

10.6.2 Appendix B

Summary of the Moderation Analyses with Prior knowledge and Academic Self-Concept Regarding Explanatory Characteristics

	ss	*df*	*MS*	*F*	*p*	η_p^2
Prior knowledge as Moderator						
Personal references						
Explanatory modality	9.40	1	9.43	1.00	.320	.012
Prior knowledge	6.50	1	6.57	0.69	.410	.008
Interaction (Modality × Prior knowledge)	11.70	1	11.71	1.24	.268	.015
Concepts						
Explanatory modality	56.0	1	56.04	2.10	.151	.024
Prior knowledge	78.40	1	78.44	2.94	.090	.034
Interaction (Modality × Prior knowledge)	4.20	1	4.22	0.16	.692	.002
Elaborations						
Explanatory modality	0.42	1	0.42	1.81	.182	.019
Prior knowledge	1.53	1	1.53	6.57	**.012**	.070
Interaction (Modality × Prior knowledge)	0.80	1	0.80	3.43	.068	.037
Academic Self-Concept as Moderator						
Personal references						
Explanatory modality	9.40	1	9.43	1.02	.317	.012
Academic self-concept	2.10	1	2.12	0.23	.634	.003
Interaction (Modality × Self-concept)	27.30	1	27.28	2.94	.090	.034
Concepts						
Explanatory modality	56.00	1	56.04	2.13	.149	.024
Academic self-concept	82.60	1	82.56	3.13	.081	.035
Interaction (Modality × Self-concept)	26.70	1	26.69	1.01	.317	.011
Elaborations						
Explanatory modality	0.42	1	0.42	1.62	.207	.019
Academic self-concept	0.03	1	0.03	0.13	.719	.002
Interaction (Modality × Self-concept)	0.02	1	0.02	0.08	.778	.001

Note. Significant results are highlighted in bold letters.

11

GENERAL DISCUSSION

11. General Discussion

Acquiring deep conceptual knowledge in science is a fundamental condition in order to gain scientific literacy and it, therefore, presents a crucial aim in education (BMBF, 2019; KMK, 2009). By and large, students, however, have displayed only a rudimentary and superficial conceptual knowledge in science subjects (i.e., STEM subjects, see MINT Nachwuchsbarometer, 2020; Reiss et al., 2019). Additionally, students often inaccurately judge their own current understanding (Griffin et al., 2008; Kiewra, 2005; Maki et al., 1994; Maki & Berry, 1984; Maki & McGuire, 2009), even though accurate monitoring is a crucial prerequisite for acquiring knowledge (Anderson & Thiede, 2008; Schleinschok et al., 2017; Thiede et al., 2003a, 2003b). Students' inability to build deep conceptual knowledge (i.e., transfer knowledge) and their insufficient metacognitive monitoring abilities (i.e., judgement of learning) conflict with the aim of science education to enable students to achieve high levels in science literacy, as well as highlight the need for additional learning support which engages students in deep learning processes. Generative learning strategies are very well suited to meet these challenges, as they lead students deeply to elaborate provided content and thus result in meaningful learning (Fiorella & Mayer, 2016; Mayer, 1996; Wittrock, 2010). In this context, learning by explaining to fictitious peers might be an excellent generative learning strategy (Brod, 2020; Fiorella & Mayer, 2016). The act of explaining enables students to select the most relevant information, to organize and restructure the information, and to integrate the new content with their prior knowledge. These processes are regarded to result in meaningful learning and may help students to monitor their understanding (Brod, 2020; Mayer, 1996; Wittrock, 2010). Against these theoretical assumptions concerning the positive effects of explaining, recent studies have demonstrated only mixed findings when it comes to the impact of explaining to fictitious others on students' own comprehension and monitoring accuracy (Fukaya, 2013; Hoogerheide et al., 2016; Hoogerheide, Visee, et al., 2019; Lachner et al., 2018; Lachner et al., 2020). Meta-analyses on the effect of explaining to fictitious showed a large variance among studies, which is an indicator that learning by explaining is not necessarily effective but rather depends on further boundary conditions (Kobayashi, 2019; Lachner et al., 2021). For instance, the explanatory modality seemed to play a crucial part in the effectiveness of explaining, as mainly oral explaining but not writing explanations resulted in higher learning outcomes (Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021). Such boundary conditions have rarely been explored by prior research and need further investigation to provide evidence for the effectiveness of explaining to fictitious peers. Another research gap is the investigation of the underlying mechanism of explaining. Theoretical assumptions are rarely empirically tested and need more exploration in future experiments.

Against this background, I have developed the ISEO framework model that contains potential boundary conditions and underlying mechanisms of explaining based on reviewing prior research. The aims of this dissertation were to empirically test this framework model, aiming at contributing to prior research by providing an empirical validated extension of prior assumptions. In three experimental studies, I investigated both implementation-related (i.e., Study 1: text complexity; Study 2: social presence) and student-related (i.e., Study 3: prior knowledge and academic self-concept) boundary conditions. As explanatory modality showed itself to be a crucial implemented-boundary condition of explaining, I compared oral explaining and writing out explanations in all three studies. In order to investigate the boundary conditions that are included in the ISEO framework model, I formulated five hypotheses:

1. *Explaining hypothesis:* Students who generate an explanation outperform students who retrieve the content in regard to students' comprehension (RQ 1a) and their monitoring accuracy (RQ 1b), regardless of the text complexity.
2. *Modality hypothesis:* Students who generate an oral explanation outperform students who generate a written explanation in regard to students' comprehension (RQ 2a) and monitoring accuracy (RQ 2b); but only when studying from highly complex learning material.
3. *Social presence hypothesis:* Inducing social presence during explaining results in higher outcomes regarding students' comprehension (RQ 3a) and their monitoring accuracy (RQ 3b).
4. *Prior knowledge hypothesis:* Students with low prior knowledge benefit more from explaining than students with high prior knowledge in regard to their comprehension (RQ 4a) and their monitoring accuracy (RQ 4b).
5. *Academic self-concept hypothesis:* Students with low academic self-concept benefit more from explaining than students with high academic self-concept regarding their comprehension (RQ 5a) and their monitoring accuracy (RQ 5b).

Additionally, I systematically analyzed students' generated explanations based on my own coding scheme to explore the mechanism of explaining (i.e., retrieval practice hypothesis: concepts; generative learning hypothesis: elaborations; social presence hypothesis: personal references).

11.1 Summary of Studies

Aiming at solving open research questions regarding boundary conditions and the underlying mechanism of explaining to fictitious peers, I conducted three experimental studies in which I systematically tested the ISEO framework model. In Study 1, I analyzed the effect of text complexity on the effectiveness of explaining in oral versus written form for students' comprehension and their monitoring accuracy (RQ 1; RQ 2). University students (*N* = 115) randomly read a limitedly (topic: reproduction) versus a highly (topic: immunology) complex text and explained the content afterwards in either oral or written form to a fictitious peer. In a control group, students retrieved the provided content as a free recall activity. Results showed no significant difference between conditions regarding the text with low complexity on students' comprehension and their monitoring accuracy. This contradicts with the first hypothesis (*explaining hypothesis*) that claims that students benefit more from explaining than from retrieving the content. Regarding the highly complex text, analyses revealed that students who explained indeed showed higher learning outcomes and more accurate monitoring judgements than students who retrieved the learning content. Hence, the *explaining hypothesis* (RQ 1) could only be partly confirmed, as explaining resulted in higher learning outcomes only when the material was complex. Additional, results revealed that oral explaining resulted in higher learning outcomes and more accurate monitoring judgements than writing explanations. These results provide empirical evidence that oral explaining is indeed more efficient compared to writing out an explanation, which is in line with the *modality hypothesis* (RQ2). Besides these investigations of boundary conditions, the underlying mechanism of explaining was also explored. A mediation analysis revealed that the effect of explanatory modality in the highly complex text condition was mediated by personal references and concepts within students' generated explanations. Hence, the results of the analysis support the *social presence hypothesis* as the underlying mechanism of explaining, as personal references serve as an indicator for perceived social presence.

As a follow-up study, I analyzed the influence of social presence, when explanations were being produced in written form, as additional support to enhance students' learning in Study 2 (RQ 3: *social presence hypothesis*). University students (*N* = 137) read the highly complex text (i.e., immunology), which was used in Study 1, and then, again, explained its content in either oral or written form. A second writing condition was implemented in which social presence was induced. Again, students retrieved the provided content in the control condition. Results revealed no significant differences among conditions, neither regarding students' comprehension nor regarding their monitoring accuracy. Additionally, analyses showed that explaining was more enjoyable, effortful, and interesting than retrieval practice. As explaining resulted in outcomes comparable to retrieving the content (RQ 1: *explaining hypothesis*), and oral explaining was as beneficial as writing an explanation (RQ 2: *modality hypothesis*), the first and second hypothesis had to be rejected. Additionally, even though students who wrote an explanation in a situation with induced social presence mentioned more personal references in their explanation, they showed comparable outcomes to the students who wrote an explanation without induced social presence. Thus, the *social presence hypothesis* had to be rejected as well (RQ 3).

In Study 3, I investigated the impact of students' prerequisites and tested, besides research RQ 1 (*explaining hypothesis*) and RQ 2 (*modality hypothesis*), whether students with low prior knowledge (RQ 4: *prior knowledge hypothesis*) and low academic self-concept (RQ 5: *academic self-concept hypothesis*) might especially benefit from explaining. Secondary school students (N = 129) participated in a scripted lesson about photosynthesis and then randomly explained the content to a fictitious peer either in oral or in written form. Again, students in a control group retrieved the content. Results showed no significant differences among conditions, neither regarding students' comprehension nor regarding their monitoring accuracy. Thus, the *explaining hypothesis* (RQ 1) and the *modality hypothesis* (RQ 2) had to be rejected. Additional moderation analyses revealed that students' academic self-concept but not that their prior knowledge moderated learning by explaining, considering that only students with low academic self-concept benefited from explaining. Students with high academic self-concept benefited the most from retrieving the content. Thus, the *prior knowledge hypothesis* (RQ 4) had to be rejected but results of this study provided evidence for the *academic self-concept hypothesis* (RQ 5).

11.2 Is Explaining an Effective Learning Strategy?

One main aim of this dissertation has been to investigate whether certain boundary conditions determine the effectiveness of explaining on students' comprehension and their monitoring accuracy. In this context, prior research highlighted the potential impact of explanatory modality and of text complexity (Hoogerheide et al., 2016; Lachner et al., 2018). The results of Study 1 showed that text complexity was indeed an important boundary condition of learning by explaining, as explaining was only effective when the learning content was complex and challenging to understand. These results are in line with prior research which discussed students' dependence on additional support, for instance, in the form of learning strategies when content is complex and challenging, such as in STEM subjects (Brod, 2020; Heublein et al., 2017; MINT Nachwuchsbarometer, 2020; Reiss et al., 2019). Thus, explaining might appear as an effective learning strategy to support students during processing difficult learning content. However, these results must be interpreted with caution as they could not be replicated in the follow-up study. Using the same materials and age group, results of Study 2 showed no significant differences between explaining and retrieval practice. Interestingly, students in Study 2 showed slightly lower learning outcomes than in Study 1. Additionally, the results revealed a higher correlation between students' prior knowledge and their learning outcomes, which might be an indicator that the content to be learned was too difficult for the second student population. Apparently, students have problems in applying this learning strategy successfully when the study material is too difficult (cf. Brod, 2020). In regard to handling difficult content, such as in STEM subjects, these results highlight the need of additional support during explaining, such as pre-trainings, prompts, or feedback (Chi et al., 1994; Lachner & Neuburg, 2019; McNamara, 2004; Schworm & Renkl, 2007). Nevertheless, since the beneficial effects of explaining over retrieval practice in Study 1 could not be replicated neither in Study 2 nor in Study 3, the question arises whether explaining actually is an effective learning strategy. This question is particularly related to the increasing number of studies that show only null-findings of explaining to fictitious peers (Fiorella et al., 2017; Fukaya, 2013; Lachner et al., 2020; Lachner et al., 2021). One explanation about the null findings may refer to the implementation of the control conditions which differed widely across prior studies. Whereas prior research mainly realized a plain control condition (i.e., restudying the content; see Fiorella & Kuhlmann, 2020; Fiorella & Mayer, 2013; Fukaya, 2013; Hoogerheide et al., 2014; Hoogerheide et al., 2016; Hoogerheide, Renkl, et al., 2019; Hoogerheide, Visee, et al., 2019), I implemented a retrieval practice condition, which is commonly regarded as a powerful learning strategy (see Fiorella & Mayer, 2016; Waldeyer et al., 2020, for an overview). Retrieval practice is considered as an effective strategy for students since retrieval enhances memory (Atkinson & Shiffrin, 1968), provides access to retrieval cues (Rowland, 2014), and enables students to construct new retrieval cues (i.e., spreading activation, see Carpenter, 2009; Endres et al., 2017; Roediger & Butler, 2011; Rowland, 2014). Thus, the studies conducted for this dissertation extended prior research, considering that the results indicated that explaining is more effective than restudying content, but not compared to retrieval practice (see results of Study 2 and Study 3).

As the mixed findings of the dissertational studies raise the question whether explaining is an effective strategy across all prior studies, I conducted two continuously cumulating meta-analyses (i.e., factual knowledge, transfer knowledge) based on prior experiments on explaining to fictitious others and the dissertational studies. In total, fifteen published studies have been included that analyze the effectiveness of explaining to fictitious others by contrasting an explaining condition (oral or written form) to a control condition in which students either restudied or retrieved the provided content. In total, these studies contain twenty-eight experiments that were conducted between 2013 and 2021 (see Chapter 4 for more details). In the meta-analyses, students' comprehension (i.e., factual knowledge, transfer knowledge) was the dependent variable and learning strategy the independent variable. The effect sizes across experiments were combined using one standardized metric (g) based on means and standard deviations of each single experiment (see Borenstein et al., 2009, for more details). For studies in which a within-subjects-design was conducted, the standardized mean change (SMCR) was used, while for studies with between-subjects-designs, the standardized mean difference (SMD) was applied instead.

Regarding factual knowledge, all studies (N = 28) were included in the analysis, as in all studies the effect of explaining on students' factual knowledge was investigated (see Table 8). These studies included nineteen possible comparisons regarding oral explaining and nine comparisons regarding writing explanations. These comparisons were based on data of 1,870 students (explaining conditions: 944; control condition: 926). Results of the meta-analysis showed a small significant combined effect of explaining to a fictitious peer on students' factual knowledge: $g = 0.145$, 95% CI [0.002, 0.288], $p = .047$ (see Table 8, for the single effect sizes, and Figure 16 for graph representation). In line with prior research that demonstrated a large variance across studies on explaining (Kobayashi, 2019; Lachner et al., 2021), the heterogeneity index was significant, as well, $Q(27) = 67.26$, $p < .001$, indicating significant heterogeneity across studies. To test whether explanatory modality indeed influenced the effectiveness of explaining across studies, I conducted an additional moderation analysis with explaining modality as moderator. Results revealed a significant moderation effect of explanatory modality on students' factual knowledge, $QM(1) = 11.05$, $p < .001$, indicating that the effectiveness of explaining depended on whether students generated oral or written explanations. Separate meta-analyses showed a significant effect of oral explaining, $g = 0.294$, 95% CI [0.129, 0.459], $p < .001$, but not of writing explanations, $g = -0.144$, 95% CI [-0.310, 0.023], $p = .091$, indicating that explaining is only more beneficial than restudying or retrieval practice when students explain in oral form (see Table 8 for the effect sizes).

Table 8

Effects of Explaining on Students' Factual Knowledge

Authors	Standardized mean difference	95% CI lower limit	95% CI upper limit
Fiorella & Kuhlmann, 2019 (oral)	0.45	-0.07	0.96
Fiorella & Mayer, 2013, Exp. 1 (oral)	0.81	0.28	1.33
Fiorella & Mayer, 2014, Exp. 2 (oral)	0.55	-0.02	1.11
Fiorella et al., 2017, Exp. 2 (oral)	-0.13	-0.64	0.37
Fiorella et al., 2017, Exp. 2 (oral)	-0.15	-0.66	0.36
Fukaya, 2013, Exp. 1 (oral)	0.93	0.12	1.74
Fukaya, 2013, Exp. 2 (oral)	0.57	-0.16	1.30
Hoogerheide et al., 2014, Exp. 1 (oral)	0.42	-0.13	0.98
Hoogerheide et al., 2014, Exp. 2 (oral)	0.87	0.36	1.38
Hoogerheide et al., 2016, Exp. 1 (written)	-0.06	-0.56	0.44
Hoogerheide et al., 2016, Exp. 2 (oral)	0.62	0.19	1.05
Hoogerheide et al., 2016, Exp. 2 (written)	0.39	-0.04	0.81
Hoogerheide, Renkl, et al., 2019 (oral)	0.43	-0.08	0.94
Hoogerheide, Visee, et al., 2019 (oral)	0.71	0.27	1.14
Koh et al., 2018 (oral)	-0.11	-0.61	0.39
Lachner et al., 2020, Exp. 1 (oral)	-0.04	-0.55	0.47
Lachner et al., 2020, Exp. 2 (oral)	0.13	-0.37	0.62
Lachner et al., 2021, Exp. 1 (written)	0.08	-0.41	0.56
Lachner et al., 2021, Exp. 2 (written)	-0.31	-0.60	-0.03
Jacob et al., 2020, Study 1 (oral, easy text)	-0.07	-0.52	0.38
Jacob et al., 2020, Study 1 (oral, complex text)	0.11	-0.34	0.56
Jacob et al., 2020, Study 1 (written, easy text)	-0.21	-0.67	0.24
Jacob et al., 2020, Study 1 (written, complex text)	-0.16	-0.61	0.30
Jacob et al., 2021, Study 2 (written chat condition)	-0.47	-0.94	0.01
Jacob et al., 2021, Study 2 (oral)	0.06	-0.42	0.53
Jacob et al., 2021, Study 2 (written)	-0.32	-0.80	0.16
Jacob et al., sub., Study 3 (oral)	-0.07	-0.50	0.35
Jacob et al., sub., Study 3 (written)	-0.13	-0.55	0.29

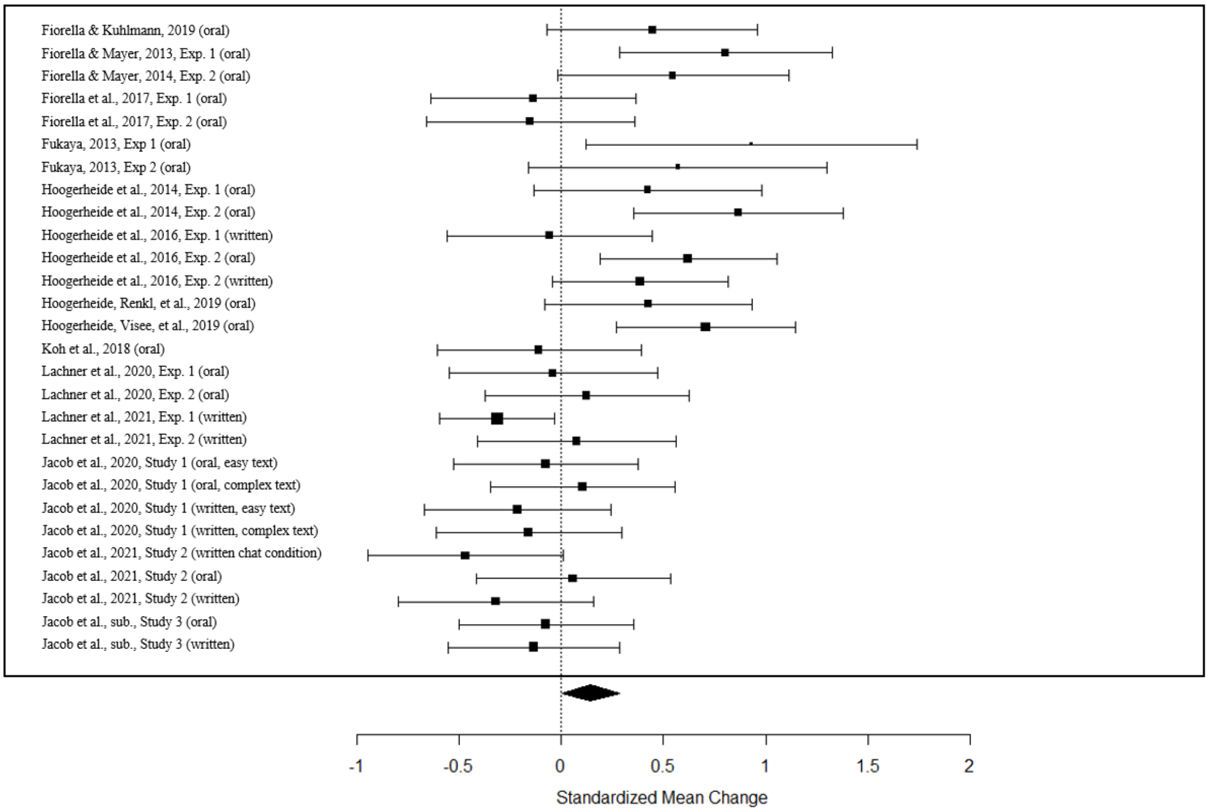

Figure 16. Effects of Explaining on Factual Knowledge Across Studies.

I conducted a similar meta-analysis regarding students' transfer knowledge. This analysis was based on eighteen experiments which contained nine possible comparisons regarding oral explaining and nine comparisons concerned with writing explanations (see Table 9). These comparisons were based on the data of 1,320 students (explaining conditions: 672; control condition: 648). In contrast to factual knowledge, the results revealed no significant combined effect of explaining on students' transfer knowledge, $g = 0.120$, 95% CI [-0.013, 0.253], $p = .078$ (see Table 9, for the single effect sizes, and Figure 17 for graphical representation). Additionally, the heterogeneity index was not significant either, $Q(17) = 27.15$, $p = .056$, indicating that the samples were rather homogeneous regarding students' transfer knowledge. In sum, across studies, explaining indeed had a positive effect on students' learning; interestingly, and in contrast to the generative learning theory, explaining did not result in meaningful learning (i.e., transfer knowledge), since the results only showed effects on students' factual knowledge but not their transfer knowledge (see Chapter 11.4 for potential explanations).

Table 9

Effects of Explaining on Students' Transfer Knowledge

Author	Standardized mean difference	95% CI lower limit	95% CI upper limit
Fiorella & Kuhlmann, 2020 (oral)	0.75	0.23	1.28
Hoogerheide et al., 2016, Exp. 1 (written)	0.21	-0.29	0.71
Hoogerheide et al., 2016, Exp. 2 (written)	0.13	-0.29	0.56
Hoogerheide et al., 2016, Exp. 2 (oral)	0.31	-0.12	0.74
Hoogerheide, Renkl, et al., 2019 (oral)	0.51	0.00	1.02
Lachner et al., 2020, Exp. 1 (oral)	0.00	-0.51	0.51
Lachner et al., 2020, Exp. 2 (oral)	0.20	-0.30	0.70
Lachner et al., 2021, Exp. 1 (written)	-0.39	-0.87	0.10
Lachner et al., 2021, Exp. 2 (written)	0.25	-0.03	0.53
Jacob et al., 2020, Study 1 (oral, easy text)	0.51	0.06	0.97
Jacob et al., 2020, Study 1 (oral, complex text)	-0.20	-0.65	0.25
Jacob et al., 2020, Study 1 (written, easy text)	0.19	-0.26	0.65
Jacob et al., 2020, Study 1 (written, complex text)	-0.45	-0.91	0.01
Jacob et al., 2021, Study 2 (written chat condition)	0.26	-0.22	0.73
Jacob et al., 2021, Study 2 (oral)	0.08	-0.39	0.55
Jacob et al., 2021, Study 2 (written)	-0.04	-0.51	0.44
Jacob et al., sub., Study 3 (oral)	-0.04	-0.46	0.38
Jacob et al., sub., Study 3 (written)	-0.09	-0.52	0.33

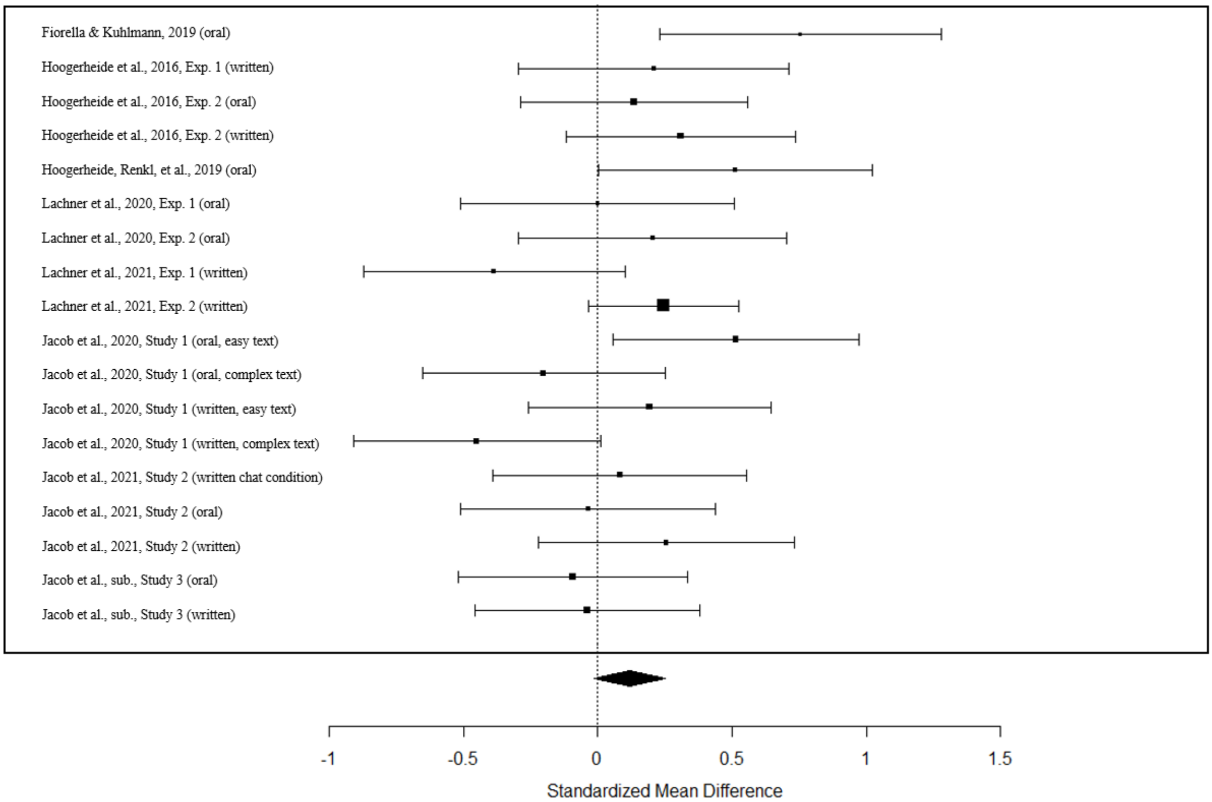

Figure 17. Effects of Explaining on Transfer Knowledge Across Studies.

11.3 Social Presence as the Underlying Mechanism of Explaining

The results of the present meta-analyses show that explaining to a fictitious peer had a positive effect on students' factual knowledge. However, little is still known about the underlying mechanism of explaining. Exploring this open research question in more detail was another main aim of the present dissertation. Aiming at solving this questions, three potential mechanisms were investigated based on prior research. First, the *retrieval practice hypothesis* claims that the act of explaining is a product of a retrieval process, as students are forced to retrieve all previously learned content for generating an explanation (Koh et al., 2018; Lachner et al., 2021). Second, in line with the generative learning theory (Mayer, 1996; Wittrock, 1974), the *generative learning hypothesis* extends the *retrieval practice hypothesis* and states that explaining is a generative learning strategy which helps students to deeply elaborate new content. Deeply elaborating provided content is considered to result in new mental knowledge representations that enhance meaningful learning (Chi, 2009; Chi & Wylie, 2014; Fiorella & Mayer, 2016). Lastly, the research highlights the important role of perceived social presence during explaining which is the *social presence hypothesis* (Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021). Perceived social presence might either motivate students to generate an explanations (*motivational perspective*, see Hoogerheide, Visee, et al., 2019), or it leads students to adapt their knowledge to the audience (*audience design perspective*, see Lachner et al., 2021) which may result in elaborative processes and more comprehensive explanations and, in turn, in higher learning outcomes. To test these hypotheses as potential mechanisms of explaining, students' generated explanations were analyzed in all the three studies of this dissertation. *Personal references* were used as an indicator for perceived social presence, the number of *concepts* represented students' retrieval during explaining, and the number of *elaborations* was used as an indicator for students' elaborative processes.

Regarding the *retrieval practice hypothesis*, the results of the studies reveal that oral explaining led the students to explain more concepts within their explanations compared to written explanations (see Study 1 and Study 2). Thus, oral explaining seemingly induced retrieval practice. However, these retrieval processes did not result in higher learning outcomes, as the mediation analyses with concepts as mediator were not significant. Therefore, the *retrieval practice hypothesis* must be rejected as a single mechanism of explaining.

Regarding the *generative learning hypothesis*, the results of Study 1 show that the students' level of elaboration was rather low and comparable across conditions. An additional mediation analysis reveals no significant effect of elaborations as mediator of the explaining effect. These results are in line with the results of Study 2 and Study 3, which also showed no effects of the students' level of elaboration on their comprehension. Thus, based on the results of the presented dissertation, the *generative learning hypothesis* must be rejected as a potential mechanism of explaining, as a deep elaboration of new content did not result in higher learning outcomes. This conclusion contradicts the generative learning theory (Chi, 2009; Chi & Wylie, 2014; Mayer, 1996) and also the findings of Lachner et al. (2018) who demonstrated meditating effects of elaborating on the effectiveness of explaining to fictitious others. Apparently, the implementation of the explaining task in the present dissertation did not help students to deeply elaborate the content, which highlights the need for additional support, such as pre-training before explaining (McNamara, 2004, 2017), specific prompts

(Chi et al., 1994; Schworm & Renkl, 2007), or feedback on their explanations (Lachner & Neuburg, 2019).

Regarding the *social presence hypothesis*, results of Study 1 showed that students who explained orally mentioned more personal references (which is an indicator for perceived social presence during explaining) than students who wrote an explanation. Seemingly, producing oral language increased the awareness of social presence of the (fictitious) peer, which is in line with prior research (Akinnaso, 1985; Chafe, 1982; Chafe & Tannen, 1987; Einhorn, 1978). Interestingly, higher levels of perceived social presence resulted in more retrieval processes (but not in more elaborative processes), as the serial mediation analysis with personal references and concepts on explaining was significant. The results provide empirical evidence for the *social presence hypothesis* as the underlying mechanism of explaining to fictitious others (see the revised ISEO model in Figure 18). However, these results must be interpreted with caution as the *social presence hypothesis* could not be replicated in Study 2 and Study 3. Nevertheless, Study 2 provided further evidence for the *social presence hypothesis*, as students who generated an explanation for another peer found completing the task more enjoyable than students who retrieved the content. Enjoyment was an indicator for the *motivational perspective* in prior research which might be triggered by higher levels of perceived social presence (see Hoogerheide, Visee, et al., 2019; Lachner et al., 2021). Interestingly, an additional explorative mediation analysis revealed that enjoyment indeed mediated the effect of learning by explaining on students' learning. Thus, Study 2 supports the *social presence hypothesis*, as explaining to a fictitious peer seems to motivate students to complete the task, which, in turn, resulted in higher learning outcomes. In this context, it has to be noted that prior research highlighted the crucial role of excitement to measure students' motivation during explaining (Hoogerheide, Visee, et al., 2019; Okita & Schwartz, 2013; Somerville et al., 2013), which was not measured in the present dissertation. Information on students' excitement would allow further insights into the mechanism of social presence and should be considered in future studies. In sum, future research should test the *social presence hypothesis* in more detail, such as measuring students' excitement during explaining, as the results of this dissertation highlight the potential of social presence as the mechanism of explaining.

11.4 Theoretical Implications

This dissertation provides two main theoretical implications that contribute to prior research. First, it offers a theoretical framework model of potential boundary conditions of learning by explaining to a fictitious peer on students' comprehension and monitoring accuracy. Second, it has explored the underlying mechanism of explaining. Combining the results of both, this dissertation provides specific implications for theory. To begin with, in line with prior research (Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021), the results of Study 1 show that the effectiveness of explaining strongly depends on the explanatory modality and text complexity, as explaining orally was more beneficial than writing out explanations, albeit only when the content was challenging. As these results could not be replicated in further studies, two meta-analyses were conducted to investigate the overall effect of explaining to fictitious others on students' factual knowledge and transfer knowledge across studies. In line with the *modality hypothesis* and results of Study 1, the analyses reveal that explaining not only had a positive effect on students' factual knowledge but was also moderated by the explanatory modality, as oral explaining was more beneficial than writing explanations (see Figure 18 for the revised ISEO model in which the explanatory modality is highlighted as an important boundary condition of explaining). Interestingly, the effect of explaining only occurred when it came to the students' factual knowledge but not regarding their transfer knowledge. These results might be contradictory at first glance, as explaining is regarded to initiate elaborative processes which should result in transfer knowledge (Chi, 2009; Chi & Wylie, 2014; Fiorella & Mayer, 2016). However, a consideration of the results related to the underlying mechanism of explaining might provide an explanation. Analysis of three potential mechanisms of explaining (i.e., *retrieval practice hypothesis*, *generative learning hypothesis*, *social presence hypothesis*) revealed that the *social presence hypothesis* is most likely to drive the explaining effect. Results showed that students who perceived higher levels of social presence were more motivated to retrieve provided content (i.e., more concepts), which resulted in higher learning outcomes. Interestingly, explaining did not enhance elaborative processes, which contradicts with the generative learning theory (Chi, 2009; Chi & Wylie, 2014; Mayer, 1996). In summation, the results not only highlight the crucial role of explanatory modality but also provide evidence for the *social presence hypothesis* as the underlying mechanism of explaining (see Figure 18). Since explaining did not lead the students to engage in elaborative processes but only in retrieval processes, future research should investigate how students could be supported to generate elaborated explanations (see Lachner & Neuburg, 2019, for an example).

In contrast to explanatory modality, the effect of text difficulty as the second implementation-related boundary condition of explaining did not result in a coherent pattern. In Study 1, students only benefited from explaining when studying complex content. These findings are in line with prior research, which highlighted the need of additional learning activities, such as explaining, for processing difficult learning content (Lachner & Nückles, 2015; McNamara et al., 1996; McNamara & Knitsch, 2009; Ozuru et al., 2010). However, using the same materials in Study 2, explaining did not enhance learning. Thus, analyses only partly provided empirical evidence for text complexity as a boundary condition for explaining; therefore, empirical evidence is only demonstrated in grey in the ISEO framework model (see Figure 18). These results highlight the importance of replication studies, as demonstrated within this dissertation, to provide sufficient evidence for empirical effects (cf., Merkt et al., 2020).

Regarding the last implementation-related boundary condition, I assumed that inducing social presence in the written chat condition would trigger students to generate more comprehensive explanations that, in turn, result in higher learning outcomes. In line with this assumption, results showed that students in the chat condition perceived higher levels of social presence compared to the other explaining conditions, as they mentioned significantly more personal references within their explanations. However, perceived social presence resulted neither in more comprehensive explanations nor in higher learning outcomes (see results of Study 2). These results highlight that simply inducing social presence, at least in this form, does not necessarily enhance learning. This contradicts prior research that claimed that perceived social presence is a crucial factor for students' satisfaction and perceived learning (Kim, 2013; Richardson et al., 2017; Russo & Benson, 2005; Wang & Antonenko, 2017; Weidlich & Bastiaens, 2019). However, as prior studies mainly investigated the impact on students' *perceived* learning, little is yet known about the effect of social presence on students' actual learning outcomes. Thus, future research is needed to systematically investigate the effect of social cues within learning activities, such as generating an explanation to a fictitious peer.

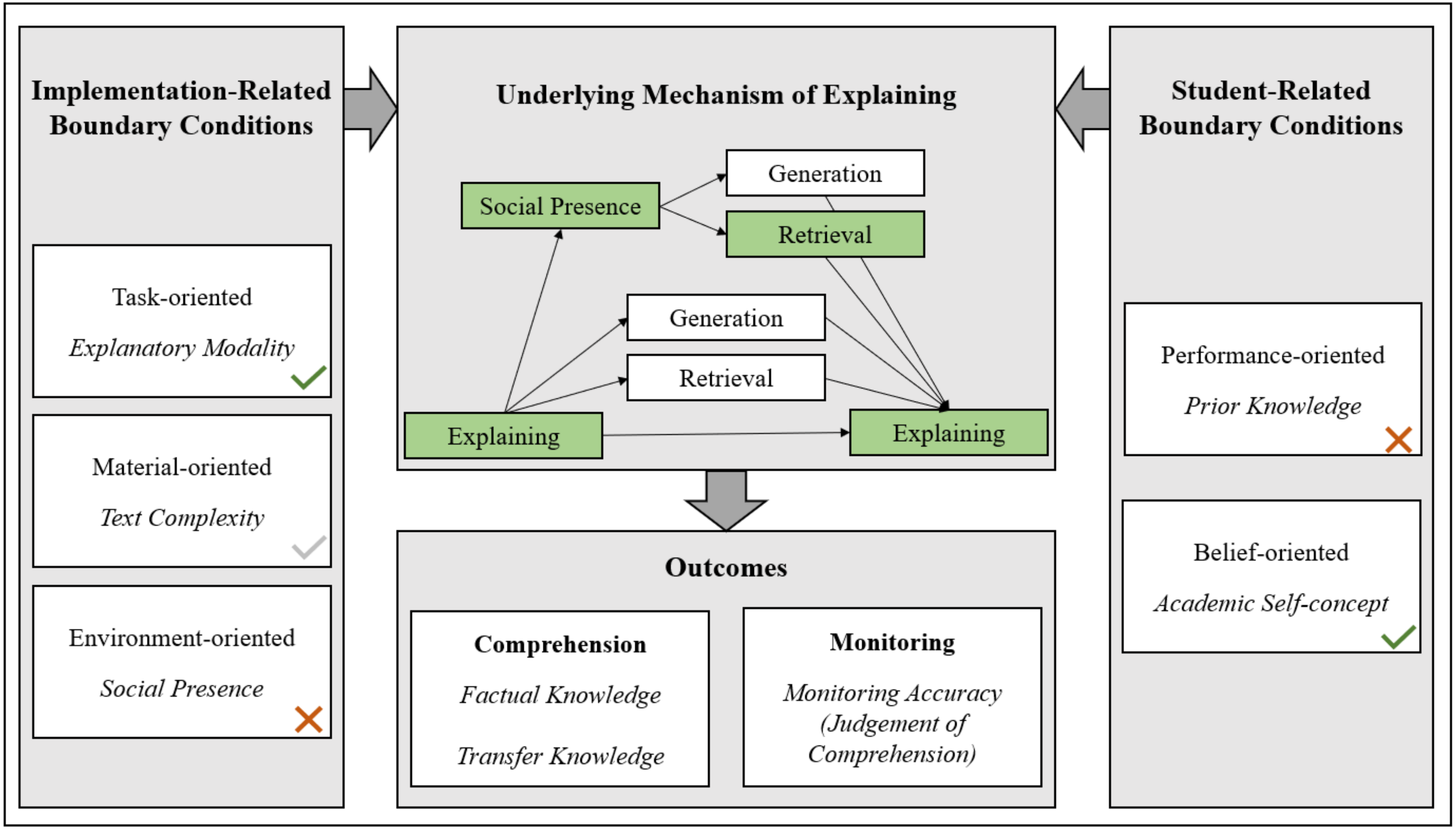

Figure 18. Revised Model of Implementation- and Student-related boundary conditions of Explaining on students' Outcome (ISEO Model)

In addition to implementation-related boundary conditions, Study 3 investigated student-related boundary conditions. Prior research demonstrated that students with low prior knowledge and a low academic self-concept especially benefited from learning strategies (Hoogerheide, Renkl, et al., 2019; Roelle & Renkl, 2020). Against this background, I have investigated whether learning by explaining is an effective learning strategy for school students, and additionally, whether it is especially beneficial for students with low prior knowledge and a low academic self-concept. Results of Study 3 show that academic self-concept, but not prior knowledge, moderates the effectiveness of explaining (see Figure 18). In line with prior research, students with low academic self-concept benefited most from explaining, whereas students with high academic self-concept benefited most from retrieving the content. A possible reason for this reversal effect might refer to the alleged presence of a social audience, which likely inclined students to make upward and downward comparisons during explaining (e.g., Becker & Neumann, 2016; Marsh et al., 2008; Trautwein et al., 2006). Students were instructed to explain the content to a fictitious student Lisa who had low prior knowledge about the topic. As students with a low self-concept benefitted more from explaining than retrieval, explaining might trigger a social situation of productive knowledge exchange because both Lisa and the student had comparable prerequisites, which may bring about a situation of information sharing (Ray et al., 2013). Students with high self-concept, however, may no longer profit from explaining because they explained the content to a less knowledgeable peer. This might create a social situation implying downward comparisons, which have been demonstrated to yield lower levels of information sharing (Ray et al., 2013) and likely lower learning outcomes. Future research is needed that includes students' perceived information sharing as an additional measure to provide empirical evidence for this assumption. Additionally, further experiments are needed that systematically investigate the effect of an academic self-concept on students' learning, for instance, by manipulatively increasing students' academic self-concept in the corresponding topic.

In contrast to the academic self-concept, students' prior knowledge did not moderate the effectiveness of explaining and it, therefore, contradicts prior research (Hoogerheide, Renkl, et al., 2019), which showed that explaining was effective for students with low but not high prior knowledge. In contrast to the students, described by Hoogerheide et al. (2019), who showed rather low prior knowledge, students in the studies behind this dissertation had more prior knowledge, which might be due to the curricular nature of the topic, namely, photosynthesis. Therefore, the data reflected rather advanced school students and not novice students, which might make it difficult to obtain potential effects on prior knowledge. Hence, the definition of low prior knowledge is always relative to the sample investigated, which is a generic problem in research on aptitude-treatment interactions. Future research should investigate the effect of prior knowledge on learning by explaining by conducting experiments with students who show very little prior knowledge of the related topic.

In summation, the studies behind this dissertation present new information about the impact of boundary conditions of learning through explanation. Regarding implementation-related boundary conditions, the explanatory modality resulted in being an impactful boundary condition of explaining, whereas inducing social presence did not enhance students' learning. Mixed findings occurred regarding text complexity, which highlights the need for further investigations. Regarding student-related boundary conditions, the academic self-concept but not prior knowledge moderated the effectiveness of explaining and, therefore, should be investigated in further research in more detail. This dissertation, additionally, provides crucial insights into the mechanism of explaining, as it shows evidence for the *social presence hypothesis*. Students who generated an explanation perceived higher levels of social presence and, therefore, generated more comprehensive explanations or reported higher levels of enjoyment, which resulted in higher learning outcomes. Hence, the dissertation not only provides insights into the mechanism of explaining but also provides empirical evidence on boundary conditions which offers future research the possibility to further investigate the mechanism and boundary conditions in more detail.

11.5 Practical Implications

School students' superficial and rudimentary knowledge in STEM subjects and high dropout rates at universities confront the educational system with an urgent need to act (Heublein et al., 2017; MINT Nachwuchsbarometer, 2020). In this context, generative learning strategies are considered to engage students in deep learning processes and to enhance their metacognitive monitoring, which should result in meaningful knowledge and higher levels of scientific literacy (Brod, 2020; Fiorella & Mayer, 2016). This dissertation aimed at investigating potential boundary conditions and the underlying mechanism of the learning strategy *learning by explaining to fictitious peers*. Analyzing this learning strategy in more detail yields several meaningful practical implications.

First, results of Study 1 and the meta-analyses showed that explaining can be an effective learning strategy to enhance students' learning in STEM subjects. Interestingly, this explaining effect was moderated by the modality of explaining, as students who explained orally showed higher learning outcomes than students who wrote out an explanation (see also Hoogerheide et al., 2016; Lachner et al., 2018; Lachner et al., 2021). Therefore, teachers and other instructors should prefer oral explaining as a strategy to implement in class rather than asking students to provide written explanations. Using digital technologies, students have ample opportunities to generate oral explanations, for instance, by creating an explanatory video or audio message in which they explain the learned content to their (fictitious) peers (see Hoogerheide, Visee, et al., 2019, for an example).

Second, social presence during explaining seemed to be a crucial factor since the dissertation studies have provided empirical evidence for perceived social presence as the underlying mechanism of explaining. Thus, higher levels of social presence resulted in more comprehensive explanations (Study 1) and enjoyment (Study 2), which, in turn, enhanced students' learning. However, this does not mean that inducing social presence within an explanatory activity automatically results in higher learning outcomes. Results of Study 2 clearly demonstrated that implementing social cues within the explanatory activity did not result in higher levels of comprehension than explanatory activities without such social cues. As this dissertation indicates that inducing social cues during explaining does not necessarily enhance students' learning, teachers should not rely only on learning environments that provide social cues when aiming at supporting students' STEM comprehension. Such social cues might only positively affect students' satisfaction and perceived learning (Kim, 2013; Richardson et al., 2017; Russo & Benson, 2005; Wang & Antonenko, 2017) but not necessarily their actual comprehension. Further research is needed to investigate the impact of social cues and social learning environments in more detail to provide clear evidence for practical implementations in STEM subjects.

Third, the dissertation provides empirical evidence that students' academic self-concept should be particularly considered when applying learning strategies (see Study 3). In general, teachers mainly assess students' prior knowledge to adapt their instructions or learning strategies to the students' needs (Dumont, 2019; Nouri, 2016). Students' academic self-concept, in contrast, is mainly assessed in research contexts but not in the classroom (e.g., Marsh et al., 2017; Marsh & Shavelson, 1985; Shavelson et al., 1976; Trautwein & Möller, 2016). Following the results of the present dissertation, teachers should consider measuring students' academic self-concept in the corresponding topic to adapt and tailor adequate learning strategies to students' needs. Students who show a low academic self-concept may benefit most from being provided with generative learning strategies, such as generating explanations. In contrast, students with a high academic self-concept seem to profit more from retrieving the learned content in the form of free recall activities.

Fourth, the results indicate that students' prior knowledge is not a boundary condition of explaining. These findings contradict the *prior knowledge hypothesis*, which claims that explaining is especially beneficial for students with low prior knowledge. However, students in the present study showed advanced levels of prior knowledge, as they reported rather high scores in the prior knowledge test. Thus, the conclusion that prior knowledge is not a boundary conditions of explaining should be interpreted with caution as the *prior knowledge hypothesis* was not tested with low achievers. Just as prior research showed beneficial effects of additional support for students with low prior knowledge (Roelle & Renkl, 2020), so also further research should investigate the impact of prior knowledge with low achievers in more detail. Further replication studies are needed before conclusions for practical implications can be drawn.

Lastly, the dissertation demonstrated that explaining is not necessarily effective for all students. As explaining was implemented as a pure explaining activity without any further information on how students should generate an explanation (e.g., structure, components, examples), these null findings indicate that students do depend on further support, such as pre-trainings or prompts (Chi et al., 1994; McNamara, 2004; Schworm & Renkl, 2007), that guide them through the activity and enable them to adequately apply the learning strategy (Brod, 2020). Thus, for practical implementation, it is essential to consider students' general ability in generating explanations (e.g., quality of explanations) before implementing this strategy in the classroom; it is likewise essential to provide further details on how to generate effective explanations (e.g., elaborative processes, see Chi, 2009; Chi & Wylie, 2014).

11.6 Conclusion

This dissertation provides deep insights into the generative learning strategy of *learning by explaining to a fictitious peer* – which the educational system should consider to better prepare students for successful STEM education. Generating and testing the ISEO framework model reveals that explaining to a fictitious peer is not necessarily an effective learning strategy to enhance students' learning but it rather depends on implementation- and student-related boundary conditions. In this context, the explanatory modality showed to be a crucial implementation-related boundary condition, as students benefited more from explaining orally than from providing written explanations. Additionally, the explaining effect was moderated by students' academic self-concept but not their prior knowledge, which provides first evidence that students' academic self-concept plays a crucial role in the effectiveness of explaining to a fictitious peer. Analyzing students' generated explanations revealed that social presence is most likely to be the underlying mechanism of explaining, as perceived social presence led the students to generate more comprehensive explanations and higher levels of enjoyment which enhanced students' learning. Nevertheless, inducing only social presence during explaining does not necessarily result in higher learning outcomes and, therefore, should be approached with caution. The dissertation thus not only supplies crucial theoretical implications, as it presents a theoretical framework model which was empirically tested within three experiments, but it also offers practical implications that may help teachers to adequately implement this generative learning strategy in the STEM subjects. In conclusion, this dissertation highlights that both implementation-related and student-related boundary conditions affect the explanatory outcome, and it reveals that the explanatory modality and students' academic self-concept, in particular, seem to play an important role in the effectiveness of learning by explaining to fictitious peers.

12

REFERENCES

12 References

Aiello, J. R., & Douthitt, E. A. (2001). Social facilitation from triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice*, *5*(3), 163–180. https://doi.org/10.1037//1089-2699.5.3.163

Ainsworth, S. E., & Scheiter, K. (2021). Learning by drawing visual representations: Potential, purposes, and practical implications. *Current Directions in Psychological Science*, *30*(1), 61–67. https://doi.org/10.1177/0963721420979582

Akinnaso, F. N. (1985). On the similarities between spoken and written language. *Language and Speech*, *28*, 323–359.

Allison, P. D. (2009). *Missing data*. *Quantitative applications in the social sciences: Vol. 136*. Sage.

Alon, I., & Herath, R. K. (2014). Teaching international business via social media projects. *Journal of Teaching in International Business, 25*(1), 44–59. https://doi.org/10.1080/08975930.2013.847814

Anderson, M. C. M., & Thiede, K. W. (2008). Why do delayed summaries improve metacomprehension accuracy? *Acta Psychologica*, *128*(1), 110–118. https://doi.org/10.1016/j.actpsy.2007.10.006

Annis, L. F. (1983). The processes and effects of peer tutoring. *Human Learning: Journal of Practical Research & Applications*, *2*(1), 39–47.

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology*, *94*(2), 416–427. https://doi.org/10.1037//0022-0663.94.2.416

Atkinson, R. K., & Shiffrin, R. M. (1968). Human Memory: A proposed system and its control processes. *Psychology of Learning and Motivation*, *2*, 89–195. https://doi.org/10.1016/S0079-7421(08)60422-3

Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. The City University of New York.

Baadte, C., & Schnotz, W. (2014). Feedback effects on performance, motivation and mood: Are they moderated by the learner's self-concept? *Scandinavian Journal of Educational Research*, *58*(5), 570–591. https://doi.org/10.1080/00313831.2013.781059

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

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), R136-40. https://doi.org/10.1016/j.cub.2009.12.014

Bandura, A. (1989). Human agency in social cognitive theory. *The American Psychologist*, *44*(9), 1175–1184. https://doi.org/10.1037/0003-066X.44.9.1175

Barlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. University Press.

Bascandziev, I., Tardiff, N., Zaitchik, D., & Carey, S. (2018). The role of domain-general cognitive resources in children's construction of a vitalist theory of biology. *Cognitive Psychology*, *104*, 1–28. https://doi.org/10.1016/j.cogpsych.2018.03.002

Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U., Krauss, S., Kunter, M., Löwen, K., Neubrand, M., & Tsai, Y.-M. (2008). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV): Dokumentation der Erhebungsinstrumente*. Max-Planck-Institut für Bildungsforschung. http://hdl.handle.net/hdl:11858/00-001M-0000-0023-998B-4

Becker, M., & Neumann, M. (2016). Context-related changes in academic self concept development: On the long-term persistence of big-fish-little-pond effects. *Learning and Instruction*, *45*, 31–39. https://doi.org/10.1016/j.learninstruc.2016.06.003

Berendes, K., Vajjala, S., Meurers, D., Bryant, D., Wagner, W., Chinkina, M., & Trautwein, U. (2018). Reading demands in secondary school: Does the linguistic complexity of textbooks increase with grade level and the academic orientation of the school track? *Journal of Educational Psychology*, *110*(4), 518–543. https://doi.org/10.1037/edu0000225

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Biel, A. L., & Friedrich, E. V. C. (2018). Why you should report bayes factors in your transcranial brain stimulation studies. *Frontiers in Psychology*, *9*, 1125. https://doi.org/10.3389/fpsyg.2018.01125

Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing self-explanation: A meta-analysis. *Educational Psychology Review*, *30*(3), 703–725. https://doi.org/10.1007/s10648-018-9434-x

Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Memory in everyday life: Vol. 1. Practical aspects of memory: Current research and issues* (Vol. 1, pp. 396–401).

BMBF. (2019). *Mit MINT in die Zukunft! Der MINT-Aktionsplan des BMBF*.

Bond, C. F., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin*, *94*(2), 265–292. https://doi.org/10.1037/0033-2909.94.2.265

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley.

Boshuizen, H. P. A., & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, *16*(2), 153–184. https://doi.org/10.1016/0364-0213(92)90022-M

Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology*, *29*(5), 591–620. https://doi.org/10.1080/00207599408248175

Bourdin, B., & Fayol, M. (2002). Even in adults, written production is still more costly than oral production. *International Journal of Psychology*, *37*(4), 219–227. https://doi.org/10.1080/00207590244000070

Breitwieser, J., & Brod, G. (2021). Cognitive prerequisites for generative learning: Why some learning strategies are more effective than others. *Child Development*, *92*(1), 258–272. https://doi.org/10.1111/cdev.13393

Brod, G. (2020). Generative learning: Which strategies for what age? *Educational Psychology Review*. Advance online publication. https://doi.org/10.1007/s10648-020-09571-9

Brod, G., & Breitwieser, J. (2019). Lighting the wick in the candle of learning: Generating a prediction stimulates curiosity. *NPJ Science of Learning*, *4*, 17. https://doi.org/10.1038/s41539-019-0056-y

Brown, L. V., & Acosta, A. (Eds.). (2007). *Psychology of motivation*. Nova Science Publisher.

Bruin, A. B. H. de, Dunlosky, J., & Cavalcanti, R. B. (2017). Monitoring and regulation of learning in medical education: The need for predictive cues. *Medical Education*, *51*(6), 575–584. https://doi.org/10.1111/medu.13267

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(6), 1563–1569. https://doi.org/10.1037/a0017021

Castro-Alonso, J. C., deKoning, B. B., Fiorella, L., & Paas, F. (2021). Five strategies for optimizing instructional materials: Instructor- and learner-managed cognitive load. *Educational Psychology Review*, 1–29. https://doi.org/10.1007/s10648-021-09606-9

Chafe, W. (1981). *Differences between colloquial and ritual Seneca or how oral literature is literary*. https://escholarship.org/uc/item/9hq7p33b

Chafe, W. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 35–54). Ablex.

Chafe, W., & Tannen, D. (1987). The relation between written and spoken language. *Annual Review of Anthropology*, *16*(1), 383–407. https://doi.org/10.1146/annurev.anthro.16.1.383

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

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*, 145–182.

Chi, M. T. H., Leeuw, N. de, Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439–477. https://doi.org/10.1016/0364-0213(94)90016-7

Chi, M. T. H., & Ohlsson, S. (2005). Complex declarative learning. In K. J. Holyoak & R. G. Morrison (Eds.), *Cambridge Handbook of Thinking and Reasoning.* Cambridge University Press.

Chi, M. T. H., & VanLehn Kurt, A. (1991). The content of physics self-explanations. *The Journal of the Learning Science, 1*(1), 69–105.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Chin, D. B., Dohmen, I. M., Cheng, B. H., Oppezzo, M. A., Chase, C. C., & Schwartz, D. L. (2010). Preparing students for future learning with teachable agents. *Educational Technology Research and Development*, *58*(6), 649–669. https://doi.org/10.1007/s11423-010-9154-5

Cho, H.-C., & Abe, S. (2013). Is two-tailed testing for directional research hypotheses tests legitimate? *Journal of Business Research*, *66*(9), 1261–1266. https://doi.org/10.1016/j.jbusres.2012.02.023

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). American Psychological Association. https://doi.org/10.1037/10096-006

Clark, H. H., & Schaefer, E. F. (1993). Dealing with overhearers. In H. H. Clark (Ed.), *Arenas of language use.* The University of Chicago Press.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). Lawrence Erlbaum Associates, Inc.

Cook, A. E., Halleran, J. G., & O'Brien, E. (1998). What is readily available during reading? A memory-based view of text processing. *Discourse Processes*, *26*(2&3), 109–129.

Corno, L. Y. N. (2008). On teaching adaptively. *Educational Psychologist, 43*(3), 161–173. https://doi.org/10.1080/00461520802178466

Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety: Experiencing flow in work and play*. Jossey-Bass Publishers.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, *100*(1), 223–234. https://doi.org/10.1037/0022-0663.100.1.223

Dentakos, S., Saoud, W., Ackerman, R., & Toplak, M. E. (2019). Does domain matter? Monitoring accuracy across domains. *Metacognition and Learning*, *14*(3), 413–436. https://doi.org/10.1007/s11409-019-09198-4

Dillenbourg, P. (Ed.). (1999). *Collaborative-learning: Cognitive and Computational Approaches*. Elsevier.

Dumont, H. (2019). Neuer Schlauch für alten Wein? Eine konzeptuelle Betrachtung von individueller Förderung im Unterricht. *Zeitschrift Für Erziehungswissenschaft*, *22*(2), 249–277. https://doi.org/10.1007/s11618-018-0840-0

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. https://doi.org/10.1111/bjop.12046

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*(3), 83–87. https://doi.org/10.1111/1467-8721.01235

Duran, D. (2016). Learning-by-teaching: Evidence and implications as a pedagogical mechanism. *Innovations in Education and Teaching International*, *54*(5), 476–484. https://doi.org/10.1080/14703297.2016.1156011

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Eickelmann, B., Bos, W., & Gerick, J. (Eds.). (2019). *ICILS 2018 #Deutschland: Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking*. Waxmann.

Eid, M., Geiser, C., Koch, T., & Heene, M. (2017). Anomalous results in G-factor models: Explanations and alternatives. *Psychological Methods*, *22*(3), 541–562. https://doi.org/10.1037/met0000083

Einhorn, L. (1978). Oral and written style: An examination of differences. *Southern Speech Communication Journal*, *43*(3), 302–311. https://doi.org/10.1080/10417947809372388

Eitel, A. (2016). How repeated studying and testing affects multimedia learning: Evidence for adaptation to task demands. *Learning and Instruction*, *41*, 70–84. https://doi.org/10.1016/j.learninstruc.2015.10.003

Enders, C. K. (2010). *Applied missing data analysis*. *Methodology in the social sciences*. Guilford Press.

Endres, T., Carpenter, S., Martin, A., & Renkl, A. (2017). Enhancing learning by retrieval: Enriching free recall with elaborative prompting. *Learning and Instruction*, *49*, 13–20. https://doi.org/10.1016/j.learninstruc.2016.11.010

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance* (pp. 683–704). Cambridge University Press. https://doi.org/10.1017/CBO9780511816796.038

Ericsson, K. A., Charness, N. H., Feltovich, P. J., & Hoffman, R. R. (Eds.). (2006). *The cambridge handbook of expertise and expert performance*. Cambridge Univ. Press.

Ericsson, K. A., & Simon, H. A. (1999). *Protocol analysis: Verbal reports as data* (Rev. ed., 3. print). *A Bradford book*. MIT Press.

Fiorella, L., & Kuhlmann, S. (2020). Creating drawings enhances learning by teaching. *Journal of Educational Psychology*, *112*(4), 811–822. https://doi.org/10.1037/edu0000392

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, *38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001

Fiorella, L., & Mayer, R. E. (2014). Role of expectations and explanations in learning by teaching. *Contemporary Educational Psychology*, *39*(2), 75–85. https://doi.org/10.1016/j.cedpsych.2014.01.001

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9

Fiorella, L., van Gog, T., Hoogerheide, V., & Mayer, R. E. (2017). It's all a matter of perspective: Viewing first-person video modeling examples promotes learning of an assembly task. *Journal of Educational Psychology, 109*(5), 653–665. https://doi.org/10.1037/edu0000161

Fonseca, B. A., & Chi, M. T. H. (2011). Instruction based on self-explanation. In R. E. Mayer & P. A. Alexander (Eds.), *Educational psychology handbook series. Handbook of research on learning and instruction.* Routledge.

Frank, B., & Kluge, A. (2014). Development and first validation of the PLBMR for lab-based microworld research. *Proceedings of the International Society for Presence Research*, 17–19.

Fukaya, T. (2013). Explanation generation, not explanation expectancy, improves metacomprehension accuracy. *Metacognition Learning*, *8*(1), 1–18. https://doi.org/10.1007/s11409-012-9093-0

Furr, R. M., & Rosenthal, R. (2003). Repeated-measures contrasts for "multiple-pattern" hypotheses. *Psychological Methods, 8*(3), 275–293. https://doi.org/10.1037/1082-989X.8.3.275

Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, *62*(3), 378–391. https://doi.org/10.1037/0022-3514.62.3.378

Golke, S., & Wittwer, J. (2018). High-performing readers underestimate their text comprehension: Artifact or psychological reality? *Proceedings of the 39th Annual Conference of the Cognitive Science*.

Grabowski, J. (2010). Speaking, writing, and memory span in children: Output modality affects cognitive performance. *International Journal of Psychology*, *45*(1), 28–39. https://doi.org/10.1080/00207590902914051

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Gräsel, Cornelia, Bruhn, Johannes, Mandl, Heinz, Fischer, & Frank (1997). Lernen mit Computernetzen aus konstruktivistischer Perspektive. *Unterrichtswissenschaft*, *25*(1), 4–18.

Griffin, T. D., Jee, B. D., & Wiley, J. (2009). The effects of domain knowledge on metacomprehension accuracy. *Memory & Cognition*, *37*(7), 1001–1013. https://doi.org/10.3758/MC.37.7.1001

Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, *36*(1), 93–103. https://doi.org/10.3758/mc.36.1.93

Gruber, H., Mandl, H., & Renkl, A. (2000). Was lernen wir in Schule und Hochschule: Träges Wissen? In H. Mandl & J. Gerstenmaier (Eds.), *Die Kluft zwischen Wissen und Handeln* (pp. 139–156). Hogrefe.

Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *American Journal of Distance Education*, *11*(3), 8–26. https://doi.org/10.1080/08923649709526970

Hancke, J., Vajjala, S., & Meurers, D. (2012). Readability classification for German using lexical, syntactic, and morphological features. *Proceedings of COLING 2012: Technical Papers*, 1063–1080.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hayes, A. F., & Montoya, A. K. (2017). A tutorial on testing, visualizing, and probing an interaction involving a multicategorical variable in linear regression analysis. *Communication Methods and Measures*, *11*(1), 1–30. https://doi.org/10.1080/19312458.2016.1271116

Hertzog, C., Hines, J. C., & Touron, D. R. (2013). Judgments of learning are influenced by multiple cues in addition to memory for past test accuracy. *Archives of Scientific Psychology*, *1*(1), 23–32. https://doi.org/10.1037/arc0000003

Heublein, U., Ebert, J., Hutzsch, C., Isleib, S., König, R., Richter, J [Johanna], & Woisch, A. (2017). *Zwischen Studienerwartungen und Studienwirklichkeit: Ursachen des Studienabbruchs, beruflicher Verbleib der Studienabbrecherinnen und Studienabbrecher und Entwicklung der Studienabbruchquote an deutschen Hochschulen*.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, *70*(2), 151–179. https://doi.org/10.3102/00346543070002151

Hinds, P. J., Patterson, M., & Pfeffer, J. (2001). Bothered by abstraction: The effect of expertise on knowledge transfer and subsequent novice performance. *Journal of Applied Psychology*, *86*(6), 1232–1243. https://doi.org/10.1037/0021-9010.86.6.1232

Hoijtink, H., Mulder, J., van Lissa, C., & Gu, X. (2019). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods*, *24*(5), 539–556. https://doi.org/10.1037/met0000201

Hoogerheide, V., Deijkers, L., Loyens, S. M., Heijltjes, A., & van Gog, T. (2016). Gaining from explaining: Learning improves from explaining to fictitious others on video, not from writing to them. *Contemporary Educational Psychology*, *44-45*, 95–106. https://doi.org/10.1016/j.cedpsych.2016.02.005

Hoogerheide, V., Loyens, S. M., & van Gog, T. (2014). Comparing the effects of worked examples and modeling examples on learning. *Computers in Human Behavior*, *41*, 80–91. https://doi.org/10.1016/j.chb.2014.09.013

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

Hoogerheide, V., Renkl, A., Fiorella, L., Paas, F., & van Gog, T. (2019). Enhancing example-based learning: Teaching on video increases arousal and improves problem-solving performance. *Journal of Educational Psychology*, *111*(1), 45–56. https://doi.org/10.1037/edu0000272

Hoogerheide, V., Renkl, A., Glogger-Frey, I., Lachner, A., Ravensbergen, S., & van Gog, T. (2020). Testing the social presence hypothesis: Teaching a real student impairs problem-solving performance. In V. Hoogerheide & J. Roelle (Eds.), *When and why does learning by teaching (not) forster learning?*. EARLI SIG 6-7 Meeting.

Hoogerheide, V., Staal, J., Schaap, L., & van Gog, T. (2019). Effects of study intention and generating multiple choice questions on expository text retention. *Learning and Instruction*, *60*, 191–198. https://doi.org/10.1016/j.learninstruc.2017.12.006

Hoogerheide, V., Visee, J., Lachner, A., & van Gog, T. (2019). Generating an instructional video as homework activity is both effective and enjoyable. *Learning and Instruction*, *64*, 101226. https://doi.org/10.1016/j.learninstruc.2019.101226

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology*, *49*(5), 505–528. https://doi.org/10.1016/j.jsp.2011.07.001

Hurd, S. (2006). Individual learner differences and distance language learning: An overview. *RTVU ELT Express*, *12*(4).

IET. (2008). *Studying stem: What are the barriers?*

Jacob, L., Lachner, A., & Scheiter, K. (2020). Learning by explaining orally or in written form? Text complexity matters. *Learning and Instruction*, *68*, 101344. https://doi.org/10.1016/j.learninstruc.2020.101344

Jairam, D., & Kiewra, K. A. (2010). Helping students soar to success on computers: An investigation of the SOAR study method for computer-based learning. *Journal of Educational Psychology*, *102*(3), 601–614. https://doi.org/10.1037/a0019137

Jerusalem, M., & Schwarzer, R. (1999). Allgemeine Selbstwirksamkeitserwartung. In R. Schwarzer & M. Jerusalem (Eds.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen* (pp. 13–14).

Jonassen, D. H. (2005). Tools for representing problems and the knowledge required to solve them. In S.-O. Tergan & T. Keller (Eds.), *Knowledge and information visualization: Searching for synergies* (pp. 82–94). Springer.

Jong, T. de, & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*(2), 105–113.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, *19*(4), 509–539. https://doi.org/10.1007/s10648-007-9054-3

Katona, G. (1940). *Organizing and memorizing: Studies in the psychology of learning and teaching*. Columbio University Press.

Kiewra, K. A. (2005). *Learn how to study and SOAR to success*. Pearson, Prentice Hall.

Kim, Y. (2013). Digital peers to help children's text comprehension and perceptions. *Educational Technology & Society*, *16*(4), 59–70.

Kim, Y., & Baylor, A. (2007). Pedagogical agents as social models to influence learner attitudes. *Instructional Technology*, *47*(1), 23–28.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W., & van Dijk, T. A [Teun A.] (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Klemke, A. (2017). IWM-Study 2.0: A generic learning environment for online studies. *Leibniz-Institut Für Wissensmedien*.

KMK. (2009). *Empfehlung der Kultusministerkonferenz zur Stärkung der mathematisch-naturwissenschaftlich-technischen Bildung: Beschluss der Kultusministerkonferenz vom 07.05.2009*.

KMK. (2020). *Bildungsstandards im Fach Biologie für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.06.2020*.

Knogler, M., Harackiewicz, J. M., Gegenfurtner, A., & Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology*, *43*, 39–50. https://doi.org/10.1016/j.cedpsych.2015.08.004

Kobayashi, K. (2019). Learning by preparing-to-teach and teaching: A meta-analysis. *Japanese Psychological Research*, *61*(3), 1–12. https://doi.org/10.1111/jpr.12221

Koh, A. W. L., Lee, S. C., & Lim, S. W. H. (2018). The learning benefits of teaching: A retrieval practice hypothesis. *Applied Cognitive Psychology*, *32*(3), 401–410. https://doi.org/10.1002/acp.3410

Köller, O., Magenheim, J., Molitor, H., Ramseger, J., Steffensky, M., Winther, E., Wollring, B., & Wiesmüller, C. (Eds.). (2019). *Wissenschaftliche Untersuchungen zur Arbeit der Stiftung "Haus der Kleinen Forscher": Band 11. Zieldimensionen für Multiplikatorinnen und Multiplikatoren früher MINT-Bildung*. Verlag Barbara Budrich.

Kong, Y., Seo, Y. S., & Zhai, L. (2018). Comparison of reading performance on screen and on paper: A meta-analysis. *Computers & Education*, *123*(1), 138–149.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037//0022-3514.77.6.1121

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Eds.). (2013). *Mathematics teacher education: Vol. 8*. *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project*. Springer.

Lachner, A., Backfisch, I., Hoogerheide, V., van Gog, T., & Renkl, A. (2020). Timing matters! Explaining between study phases enhances students’ learning. *Journal of Educational Psychology*, *112*(4), 841–853. https://doi.org/10.1037/edu0000396

Lachner, A., Burkhart, C., & Nückles, M. (2017a). Mind the gap! Automated concept map feedback supports students in writing cohesive explanations. *Journal of Experimental Psychology: Applied*, *23*(1), 29–46. https://doi.org/10.1037/xap0000111

Lachner, A., Burkhart, C., & Nückles, M. (2017b). Mind the gap! Automated concept map feedback supports students in writing cohesive explanations: A validation study of the CohViz Feedback System. *Journal of Experimental Psychology. Applied*, *23*(1), 29–46. https://doi.org/10.1037/xap0000111

Lachner, A., Hoogerheide, V., van Gog, T., & Renkl, A. (in prep.). When and why does learning by non-interactive teaching work?

Lachner, A., Jacob, L., & Hoogerheide, V. (2021). Learning by writing explanations: Is explaining to a fictitious student more effective than self-explaining? *Learning and Instruction*.

Lachner, A., Ly, K.-T., & Nückles, M. (2018). Providing written or oral explanations? Differential effects of the modality of explaining on students' conceptual learning and transfer. *Journal of Experimental Education*, *86*(3), 344–361. https://doi.org/10.1080/00220973.2017.1363691

Lachner, A., & Neuburg, C. (2019). Learning by writing explanations: Computer-based feedback about the explanatory cohesion enhances students' transfer. *Instructional Science*, *47*(1), 19–37. https://doi.org/10.1007/s11251-018-9470-4

Lachner, A., & Nückles, M. (2015). Bothered by abstractness or engaged by cohesion? Experts' explanations enhance novices' deep-learning. *Journal of Experimental Psychology. Applied*, *21*(1), 101–115. https://doi.org/10.1037/xap0000038

Lakoff, R. T. (1982). Some of my favorite writers are literate: The mingling of oral and literate strategies in written communication. In D. Tannen (Ed.), *Spoken and written language: Exploring orality and literacy* (pp. 239–260). Ablex.

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

Leinhardt, G. (2002). Instructional explanations: A commonplace for teaching and location for contrast. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 333–357). American Educational Research Assoc.

Liberman, A. (1994). Why is speech so much easier than reading and writing? In C. A. Fowler (Ed.), *Status report on speech research* (pp. 175–183).

Maki, R. H., & Berry, S. L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 663–679. https://doi.org/10.1037/0278-7393.10.4.663

Maki, R. H., Jonas, D., & Kallod, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin & Review*, *1*(1), 126–129. https://doi.org/10.3758/BF03200769

Maki, R. H., & McGuire, M. J. (2009). Metacognition for text: Findings and implications for education. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied Metacognition* (Vol. 39, pp. 39–67). Cambridge University Press. https://doi.org/10.1017/CBO9780511489976.004

Mandl, H., Gruber, H., & Renkl, A. (1994). Zum Problem der Wissensanwendung. *Zeitschrift Für Lernforschung*, *22*(3), 233–242.

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *The British Journal of Educational Psychology*, *81*(1), 59–77. https://doi.org/10.1348/000709910X503501

Marsh, H. W., Martin, A. J., Yeung, A. S., & Craven, R. G. (2017). Competence self-perceptions. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation: Theory and application* (pp. 85–115). The Guilford Press.

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality & Social Psychology Bulletin*, *34*(4), 542–552. https://doi.org/10.1177/0146167207312313

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist*, *20*(3), 107–123. https://doi.org/10.1207/s15326985ep2003_1

Marsh, H. W., Trautwein, U., Lüdtke, O., & Köller, O. (2008). Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs: Role of generalized and specific others. *Journal of Educational Psychology*, *100*(3), 510–524. https://doi.org/10.1037/0022-0663.100.3.510

Mayer, R. E. (1996). Learning strategies for making sense out of expository text: The SOI model for guiding three cognitive processes in knowledge construction. *Educational Psychology Review*, *8*(4), 357–371. https://doi.org/10.1007/BF01463939

Mayer, R. E. (2009). *Multimedia learning*. Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 43–71). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.005

Mayer, R. E., Stull, A. T., Campbell, J., Almeroth, K., Bimber, B., Chun, D., & Knight, A. (2007). Overestimation bias in self-reported SAT scores. *Educational Psychology Review*, *19*(4), 443–454. https://doi.org/10.1007/s10648-006-9034-z

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes*, *38*(1), 1–30. https://doi.org/10.1207/s15326950dp3801_1

McNamara, D. S. (2013). The epistemic stance between the author and reader: A driving force in the cohesion of text and writing. *Discourse Studies*, *15*(5), 579–595. https://doi.org/10.1177/1461445613501446

McNamara, D. S. (2017). Self-explanation and reading strategy training (SERT) improves low-knowledge students’ science course performance. *Discourse Processes*, *54*(7), 479–492. https://doi.org/10.1080/0163853X.2015.1101328

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*(1), 1–43. https://doi.org/10.1207/s1532690xci1401_1

McNamara, D. S., & Knitsch, W. (2009). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, *22*(3), 247–288.

McNamara, D. S., O'Reilly, T. P., Best, R. M., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with Istart. *Journal of Educational Computing Research*, *34*(2), 147–171. https://doi.org/10.2190/1RU5-HDTJ-A5C8-JVWE

McNamara, D. S., & Scott, J. L. (1999). Training self explanation and reading strategies. *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*, 1156–1160.

Merkt, M., Lux, S., Hoogerheide, V., van Gog, T., & Schwan, S. (2020). A change of scenery: Does the setting of an instructional video affect learning? *Journal of Educational Psychology*, *112*(6), 1273–1283. https://doi.org/10.1037/edu0000414

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

MINT Nachwuchsbarometer. (2020). *MINT Nachwuchsbarometer 2020*.

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, *79*(3), 1129–1167. https://doi.org/10.3102/0034654309337522

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, *90*(3), 376–419. https://doi.org/10.3102/0034654320919354

Moreno, M. A., & van Orden, G. C. (2001). Word recognition, cognitive psychology of. *International Encyclopedia of the Social & Behavioral Sciences*, 16556–16561. https://doi.org/10.1016/B0-08-043076-7/01553-9

Neighbors, C., Lewis, M. A., Fossos, N., & Grossbard, J. R. (2007). Motivation and risk behaviors: A self-determination perspective. In L. V. Brown & A. Acosta (Eds.), *Psychology of motivation* (pp. 99–113). Nova Science Publisher.

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect". *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *14*(4), 676–686. https://doi.org/10.1037//0278-7393.14.4.676

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framewok and new findings. *The Psychology of Learning and Motivation*, *26*, 125–173.

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, *125*(6), 737–759. https://doi.org/10.1037/0033-2909.125.6.737

Nouri, J. (2016). The flipped classroom: For active, effective and increased learning - especially for low achievers. *International Journal of Educational Technology in Higher Education*, *13*(1). https://doi.org/10.1186/s41239-016-0032-z

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

Nückles, M., Winter, A., Wittwer, J., Herbert, M., & Hübner, S. (2006). How do experts adapt their explanations to a layperson’s knowledge in asynchronous communication? An experimental study. *User Modeling and User-Adapted Interaction*, *16*(2), 87–127. https://doi.org/10.1007/s11257-006-9000-y

O’Reilly, T., & McNamara, D. S. (2007a). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional “high-stakes” measures of high school students’ science achievement. *American Educational Research Journal*, *44*(1), 161–196. https://doi.org/10.3102/0002831206298171

OECD. (2016). *PISA 2015 Results: Excellence and equity in education*.

OECD. (2017). *How does PISA for development measure scientific literacy?* OECD Publishing. https://doi.org/10.1787/9789264208780-en

(2018a). *PISA. Pisa for development assessment and analytical framework: Reading, mathematics and science*. OECD Publishing. https://doi.org/10.1787/19963777

OECD. (2018b). Pisa for development science framework: Reading, mathematics and science. In *PISA. PISA for development assessment and analytical framework: Reading, mathematics and science* (pp. 71–98). OECD Publishing. https://doi.org/10.1787/9789264305274-6-en

Oh, C. S., Bailenson, J. N., & Welch, G. F. (2018). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI*, *5*, 114. https://doi.org/10.3389/frobt.2018.00114

Okita, S. Y., & Schwartz, D. L. (2013). Learning by teaching human pupils and teachable agents: The importance of recursive feedback. *Journal of the Learning Sciences*, *22*(3), 375–412. https://doi.org/10.1080/10508406.2013.807263

O'Reilly, T., Best, R., & McNamara, D. S. (2004). Self-explanation reading training: Effects for low-knowledge readers. *Journal of the Annual Meeting of the Cognitive Science Society*, *26*(26), 1053-1058.

O'Reilly, T., & McNamara, D. S. (2007b). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes*, *43*(2), 121–152. https://doi.org/10.1080/01638530709336895

Orhan Özen, S. (2017). The effect of motivation on student achievement. In E. Karadag (Ed.), *The factors effecting student achievement* (pp. 35–56). Springer International Publishing. https://doi.org/10.1007/978-3-319-56083-0_3

Orús, C., Barlés, M. J., Belanche, D., Casaló, L., Fraj, E., & Gurrea, R. (2016). The effects of learner-generated videos for YouTube on learning outcomes and satisfaction. *Computers & Education*, *95*, 254–269. https://doi.org/10.1016/j.compedu.2016.01.007

Ozuru, Y., Briner, S., Best, R., & McNamara, D. S. (2010). Contributions of self-explanation to comprehension of high- and low-cohesion texts. *Discourse Processes*, *47*(8), 641–667. https://doi.org/10.1080/01638531003628809

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, *19*(3), 228–242. https://doi.org/10.1016/j.learninstruc.2008.04.003

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

Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*(2), 117–175. https://doi.org/10.1207/s1532690xci0102_1

Pereira, J., Echeazarra, L., Sanz-Santamaría, S., & Gutiérrez, J. (2014). Student-generated online videos to develop cross-curricular and curricular competencies in nursing studies. *Computers in Human Behavior*, *31*, 580–590. https://doi.org/10.1016/j.chb.2013.06.011

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology*, *58*, 193–198. https://doi.org/10.1037/h0049234

Piaget, J. (1926). *The language and thought of the child*. Routledge & Kegan Paul.

Pilegard, C., & Fiorella, L. (2016). Helping students help themselves: Generative learning strategies improve middle school students' self-regulation in a cognitive tutor. *Computers in Human Behavior*, *65*, 121–126. https://doi.org/10.1016/j.chb.2016.08.020

Plötzner, R., Dillenbourg, P., Praier, M., & Traum, D. (1999). Learning by explaining to oneself and to others. In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp. 103–121). Elsevier.

Pressley, M., Wood, E., Woloshyn, V. A., 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.

Prinz, A., Golke, S., & Wittwer, J. (2018). The double curse of misconceptions: Misconceptions impair not only text comprehension but also metacomprehension in the domain of statistics. *Instructional Science*, *46*(5), 723–765. https://doi.org/10.1007/s11251-018-9452-6

Prinz, A., Golke, S., & Wittwer, J. (2020). To What Extent Do Situation-Model-Approach Interventions Improve Relative Metacomprehension Accuracy? Meta-Analytic Insights. *Educational Psychology Review*, *32*(4), 917–949. https://doi.org/10.1007/s10648-020-09558-6

R Core Team. (2018). *R: A language and environment for statistical* [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ranganathan, P., Pramesh, C. S., & Buyse, M. (2016). Common pitfalls in statistical analysis: The perils of multiple testing. *Perspectives in Clinical Research*, *7*(2), 106–107. https://doi.org/10.4103/2229-3485.179436

Ray, D. G., Neugebauer, J., Sassenberg, K., Buder, J., & Hesse, F. W. (2013). Motivated shortcomings in explanation: The role of comparative self-evaluation and awareness of explanation recipient's knowledge. *Journal of Experimental Psychology. General*, *142*(2), 445–457. https://doi.org/10.1037/a0029339

Reiss, K., Weis, M., Klieme, E., & Köller, O. (2019). *PISA 2018*. Waxmann Verlag GmbH. https://doi.org/10.31244/9783830991007

Renkl, A. (1996). Träges Wissen: Wenn Erlerntes nicht genutzt wird. *Psychologische Rundschau*, *47*(2), 78–92.

Renkl, A. (2008). Lernen und Lehren im Kontext der Schule. In A. Renkl (Ed.), *Lehrbuch Pädagogische Psychologie* (pp. 109–153). Huber.

Renkl, A. (2009). Lehren und Lernen. In R. Tippelt & B. Schmidt-Hertha (Eds.), *Handbuch Bildungsforschung* (pp. 737–751). Verlag für Sozialwissenschaften.

Rhoads, J. A., Daou, M., Lohse, K. R., & Miller, M. W. (2019). The effects of expecting to teach and actually teaching on motor learning. *Journal of Motor Learning and Development*, *7*(1), 84–105. https://doi.org/10.1123/jmld.2017-0052

Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior*, *71*, 402–417. https://doi.org/10.1016/j.chb.2017.02.001

Richter, J [Juliane], Scheiter, K., & Eitel, A. (2016). Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. *Educational Research Review*, *17*, 19–36. https://doi.org/10.1016/j.edurev.2015.12.003

Richter, J [Juliane], Scheiter, K., & Eitel, A. (2018). Signaling text-picture relations in multimedia learning: The influence of prior knowledge. *Journal of Educational Psychology*, *110*(4), 544–560. https://doi.org/10.1037/edu0000220

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Roediger, H. L., & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*(3), 181–210.

Roediger, H. L., & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roelle, J., & Nückles, M. (2019). Generative learning versus retrieval practice in learning from text: The cohesion and elaboration of the text matters. *Journal of Educational Psychology*, *111*(8), 1341–1361. https://doi.org/10.1037/edu0000345

Roelle, J., & Renkl, A. (2020). Does an option to review instructional explanations enhance example-based learning? It depends on learners' academic self-concept. *Journal of Educational Psychology*, *112*(1), 131–147. https://doi.org/10.1037/edu0000365

Rogiers, A., Merchie, E., & van Keer, H. (2019). Learner profiles in secondary education: Occurrence and relationship with performance and student characteristics. *The Journal of Educational Research*, *112*(3), 385–396. https://doi.org/10.1080/00220671.2018.1538093

Roscoe, R. D. (2014). Self-monitoring and knowledge-building in learning by teaching. *Instructional Science*, *42*(3), 327–351. https://doi.org/10.1007/s11251-013-9283-4

Roscoe, R. D., & Chi, M. T. H. (2008). Tutor learning: The role of explaining and responding to questions. *Instructional Science*, *36*(4), 321–350. https://doi.org/10.1007/s11251-007-9034-5

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*(5), 521–562. https://doi.org/10.1207/s15516709cog2605_1

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581. https://doi.org/10.2307/2335739

Russo, T., & Benson, S. (2005). Learning with invisible others: Perceptions of online presence and their relationship to cognitive and affective learning. *International Forum of Educational Technology & Society*, *8*(1), 54–62.

Ryokai, K., Vaucelle, C., & Cassell, J. (2003). Virtual peers as partners in storytelling and literacy learning. *Journal of Computer Assisted Learning*, *19*(2), 195–208. https://doi.org/10.1046/j.0266-4909.2003.00020.x

Savary, J., Kleiman, T., Hassin, R. R., & Dhar, R. (2015). Positive consequences of conflict on decision making: When a conflict mindset facilitates choice. *Journal of Experimental Psychology. General*, *144*(1), 1–6. https://doi.org/10.1037/a0038551

Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Cambridge monographs and texts in applied psycholinguistics. Advances in applied psycholinguistics: Reading, writing, and language learning* (pp. 142–175). Cambridge University Press.

Schleinschok, K., Eitel, A., & Scheiter, K. (2017). Do drawing tasks improve monitoring and control during learning from text? *Learning and Instruction*, *51*, 10–25. https://doi.org/10.1016/j.learninstruc.2017.02.002

Schneider, W. (2015). *Memory development from early childhood through emerging adulthood*. Springer International Publishing. https://doi.org/10.1007/978-3-319-09611-7

Schneider, W., Schlagmüller, M., & Ennemoser, M. (2007). *LGVT 6-12: Lesegeschwindigkeits- und -verständnistest für die Klassen 6-12*. Hogrefe.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning*, *4*(1), 33–45. https://doi.org/10.1007/s11409-008-9031-3

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. https://doi.org/10.1037/0022-0663.99.2.285

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education*, *27*(2), 259–267. https://doi.org/10.1016/j.tate.2010.08.009

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*(3), 407–441. https://doi.org/10.3102/00346543046003407

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. Wiley.

Sindoni, M. G. (2013). *Spoken and written discourse in online interactions: A multimodal approach*. Routledge.

Singer Trakhman, L. M., Alexander, P. A., & Silverman, A. B. (2018). Profiling reading in print and digital mediums. *Learning and Instruction*, *57*, 5–17. https://doi.org/10.1016/j.learninstruc.2018.04.001

Snow, C. E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Rand.

Somerville, L. H., Jones, R. M [Rebecca M], Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. J. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science*, *24*(8), 1554–1562. https://doi.org/10.1177/0956797613475633

Streiner, D. L. (2015). Best (but oft-forgotten) practices: The multiple problems of multiplicity-whether and how to correct for many statistical tests. *The American Journal of Clinical Nutrition*, *102*(4), 721–728. https://doi.org/10.3945/ajcn.115.113548

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction*, *8*(4), 351–362. https://doi.org/10.1207/s1532690xci0804_5

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tannen, D. (Ed.). (1982). *Spoken and written language: Exploring orality and literacy*. Ablex.

Tannen, D. (1983). Oral and literate strategies in spoken and written discourese. In R. M. Bailey & R. M. Fosheim (Eds.), *Literacy for life: The demand for reading and writing* (pp. 79–96). The Modern Language Association.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003a). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*(1), 66–73. https://doi.org/10.1037/0022-0663.95.1.66

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003b). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*(1), 66–73. https://doi.org/10.1037/0022-0663.95.1.66

Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, *25*(6), 631–645. https://doi.org/10.1080/01443410500345172

Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept. *Journal of Personality and Social Psychology*, *90*(2), 334–349. https://doi.org/10.1037/0022-3514.90.2.334

Trautwein, U., & Möller, J. (2016). Self-concept: Determinants and consequences of academic self-concept in school contexts. In A. A. Lipnevich, F. Preckel, & R. D. Roberts (Eds.), *The Springer Series on Human Exceptionality. Psychosocial skills and school systems in the 21st century* (Vol. 84, pp. 187–214). Springer International Publishing. https://doi.org/10.1007/978-3-319-28606-8_8

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3). https://doi.org/10.18637/jss.v045.i03

van de Pol, J., van Loon, M., van Gog, T., Braumann, S., & Bruin, A. de (2020). Mapping and drawing to improve students' and teachers' monitoring and regulation of students' learning from text: Current findings and future directions. *Educational Psychology Review*, *32*(4), 951–977. https://doi.org/10.1007/s10648-020-09560-y

van Dijk, T. A., & Kintsch, W. (1992). *Strategies of discourse comprehension*. Academic Press.

van Gog, T., Hoogerheide, V., & van Harsel, M. (2020). The role of mental effort in fostering self-regulated learning with problem-solving tasks. *Educational Psychology Review.* Advance online publication. https://doi.org/10.1007/s10648-020-09544-y van Loon, M. H., Bruin, A. B. H. de, van Gog, T., van Merriënboer, J. J. G., & Dunlosky, J. (2014). Can students evaluate their understanding of cause-and-effect relations? The effects of diagram completion on monitoring accuracy. *Acta Psychologica*, *151*, 143–154. https://doi.org/10.1016/j.actpsy.2014.06.007

VanLehn, K., Jones, R. M [Randolph M.], & Chi, M. T. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences*, *2*(1), 1–59. https://doi.org/10.1207/s15327809jls0201_1

Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, F., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., Klimmt, C., Laarni, J., Ravaja, N., Sacau, A., Baumgartner, T., & Jäncke, P. (2004). *MEC Spatial Presence Questionnaire (MEC-SPQ): Short documentation and instructions for application*. Report to the European Community. http://www.ijk.hmt-hannover.de/presence

Waldeyer, J., Heitmann, S., Moning, J., & Roelle, J. (2020). Can generative learning tasks be optimized by incorporation of retrieval practice? *Journal of Applied Research in Memory and Cognition*, *9*(3), 355–369. https://doi.org/10.1016/j.jarmac.2020.05.001

Wang, J., & Antonenko, P. D. (2017). Instructor presence in instructional video: Effects on visual attention, recall, and perceived learning. *Computers in Human Behavior*, *71*, 79–89. https://doi.org/10.1016/j.chb.2017.01.049

Wassenburg, S. I., Koning, B. B. de, Koedinger, K. R., & Paas, F. (sub.). Limits of learning by teaching: Explaining to self versus explaining to others.

Watson, G. R. (1989). What is... Concept mapping? *Medical Teacher*, *11*(3-4), 265–269. https://doi.org/10.3109/01421598909146411

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, *13*(1), 21–39. https://doi.org/10.1016/0883-0355(89)90014-1

Webb, N. M., Franke, M. L., De, T., Chan, A. G., Freund, D., Shein, P., & Melkonian, D. K. (2009). 'Explain to your partner': Teachers' instructional practices and students' dialogue in small groups. *Cambridge Journal of Education*, *39*(1), 49–70. https://doi.org/10.1080/03057640802701986

Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology*, *87*(3), 406–423. https://doi.org/10.1037/0022-0663.87.3.406

Weidlich, J., & Bastiaens, T. J. (2017). Explaining social presence and the quality of online learning with the SIPS model. *Computers in Human Behavior*, *72*, 479–487. https://doi.org/10.1016/j.chb.2017.03.016

Weidlich, J., & Bastiaens, T. J. (2019). Designing sociable online learning environments and enhancing social presence: An affordance enrichment approach. *Computers & Education*, *142*, 103622. https://doi.org/10.1016/j.compedu.2019.103622

Weinhuber, M., Lachner, A., Leuders, T., & Nückles, M. (2019). Mathematics is practice or argumentation: Mindset priming impacts principle- and procedure-orientation of teachers' explanations. *Journal of Experimental Psychology. Applied*, *25*(4), 618–646. https://doi.org/10.1037/xap0000227

Weinstein, Y., Madan, C. R., & Sumeracki, M. A. (2018). Teaching the science of learning. *Cognitive Research: Principles and Implications*, *3*(1), 2. https://doi.org/10.1186/s41235-017-0087-y

Wharton, C., & Kintsch, W. (1991). An overview of the construction-integration model: A theory of comprehension as a foundation for a new cognitive architecture. *SIGART Bulletin*, *2*(4), 169–173.

Wienke, B., & Jekauc, D. (2016). A qualitative analysis of emotional facilitators in exercise. *Frontiers in Psychology*, *7*, 1296. https://doi.org/10.3389/fpsyg.2016.01296

Wiggins, B. L., Eddy, S. L., Grunspan, D. Z., & Crowe, A. J. (2017). The ICAP active learning framework predicts the learning gains observed in intensely active classroom experiences. *AERA Open*, *3*(2), 233285841770856. https://doi.org/10.1177/2332858417708567

Wiley, J., Griffin, T. D., Jaeger, A. J., Jarosz, A. F., Cushen, P. J., & Thiede, K. W. (2016). Improving metacomprehension accuracy in an undergraduate course context. *Journal of Experimental Psychology. Applied*, *22*(4), 393–405. https://doi.org/10.1037/xap0000096

Wiley, J., Griffin, T. D., & Thiede, K. W. (2005). Putting the comprehension in metacomprehension. *The Journal of General Psychology*, *132*(4), 408–428. https://doi.org/10.3200/GENP.132.4.408-428

Williams, J. J., & Lombrozo, T. (2010). The role of explanation in discovery and generalization: Evidence from category learning. *Cognitive Science*, *34*(5), 776–806. https://doi.org/10.1111/j.1551-6709.2010.01113.x

Wittrock, M. C. (1974). A generative model of mathematics learning. *Journal for Research in Mathematics Education*, *5*(4), 181. https://doi.org/10.2307/748845

Wittrock, M. C. (1985). Teaching learners generative strategies for enhancing reading comprehension. *Theory into Practice*, *24*(2), 123–126. https://doi.org/10.1080/00405848509543158

Wittrock, M. C. (2010). Learning as a generative process. *Educational Psychologist*, *45*(1), 40–45. https://doi.org/10.1080/00461520903433554

Wittwer, J., Nückles, M., Landmann, N., & Renkl, A. (2010). Can tutors be supported in giving effective explanations? *Journal of Educational Psychology*, *102*(1), 74–89. https://doi.org/10.1037/a0016727

Wylie, R., & Chi, M. T. H. (2014). The self-explaination principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 413–432). Cambridge University Press.

Yi, M. Y., & Hwang, Y. (2003). Predicting the use of web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International Journal of Human-Computer Studies*, *59*(4), 431–449. https://doi.org/10.1016/S1071-5819(03)00114-9

Zaitchik, D., Iqbal, Y., & Carey, S. (2014). The effect of executive function on biological reasoning in young children: An individual differences study. *Child Development*, *85*(1), 160–175. https://doi.org/10.1111/cdev.12145