



**Die Kunst, Concept Mapping zu trainieren und einzusetzen:
Ein Vergleich unterschiedlicher Ansätze und ihre
Bedeutung für kognitive Prozesse, kognitive Belastung und
Lernleistung**

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Lukas Bernhard Becker, M. Ed.

aus Berlin

Berichterstatter (Gutachter): Prof. Dr. Jörg Großschedl
Prof. Dr. Till Bruckermann

Datum der Disputation: 01. März 2022

Inhaltsverzeichnis

Abkürzungsverzeichnis	IV
Zusammenfassung	V
Abstract	VI
1 Einleitung	1
2 Theoretischer Hintergrund	9
2.1 Wissen und Lernstrategien	9
2.1.1 Wissenserwerb und Wissensorganisation	9
2.1.2 Klassifizierung von Wissen.....	11
2.1.3 Grundlagen des bedeutungsvollen Lernens.....	14
2.1.4 Klassifizierung von Lernstrategien	17
2.1.5 Lernstrategietraining	21
2.2 Untersuchungsgegenstand Concept Mapping	26
2.2.1 Anwendungsgebiete von Concept Maps	26
2.2.2 Wissenserwerb mithilfe von Concept Mapping	36
2.3 Trainingsansätze und Einsatzmöglichkeiten von Concept Mapping	43
2.3.1 Trainingsansätze im Concept Mapping.....	43
2.3.2 Einsatzmöglichkeiten von Concept Mapping.....	54
3 Forschungsfragen und Hypothesen	58
3.1 Studie 1.....	59
3.2 Studie 2.....	60
4 Studie 1: Untersuchung unterschiedlicher Trainingsansätze.....	63
5 Studie 2: Untersuchung der Faktoren <i>Training</i> und <i>Setting</i>	121
6 Diskussion	167
6.1 Concept Mapping-Trainingsbezogene Aspekte	170
6.1.1 Kognitive Prozesse	170
6.1.2 Kognitive Belastung	171
6.1.3 Metakognitive Vorhersage	172

6.1.4	Lernleistung.....	174
6.2	Einsatz von Concept Mapping in Elaborations- und Retrievalsettings.....	178
6.2.1	Kognitive Prozesse.....	179
6.2.2	Kognitive Belastung.....	180
6.2.3	Metakognitive Vorhersage.....	180
6.2.4	Lernleistung.....	181
6.3	Limitationen der beiden Studien.....	183
7	Zusammenfassung.....	184
8	Ausblick.....	186
	Literaturverzeichnis.....	190
	Danksagung.....	227
	Erklärung zur Verfügbarkeit von Primärdaten.....	229
	Erklärung zur Dissertation.....	230

Abkürzungsverzeichnis

aQCM Index	absolute Quality of Concept Map Index (dt. aQCM-Index)
bQCM Index	balanced Quality of Concept Map Index (dt. bQCM-Index)
CLT	Cognitive Load Theory
CM	Concept Mapping
ECL	Extraneous Cognitive Load
E-proposition	Elaboration-suggesting proposition
GCL	Germane Cognitive Load
ICL	Intrinsic Cognitive Load
JOL	Judgment of Learning
LGVT	Lesegeschwindigkeits- und -verständnisstest
O-proposition	Organization-suggesting proposition
R-proposition	Recall-suggesting proposition
SJT	Similarity Judgments Test
T++	Concept Mapping-Trainingsgruppe, die während der Trainingsphase Scaffolding und Feedback erhielt
T+	T+ in Studie 1: Concept Mapping-Trainingsgruppe, die in der Trainingsphase Concept Mapping ohne Lernhilfen wiederholt übte T+ in Studie 2: Concept Mapping-Trainingsgruppe, die während der Trainingsphase Scaffolding und Feedback erhielt (Training der Gruppe T++ in Studie 1)
T-	Kontrolltrainingsgruppe
T+E	Concept Mapping-Trainingsgruppe, die während der Trainingsphase Scaffolding und Feedback erhielt und in der Lernphase eine Concept Map mit dem verfügbaren Lernmaterial erstellte
T-E	Kontrolltrainingsgruppe, die in der Lernphase eine Concept Map mit dem verfügbaren Lernmaterial erstellte
T+R	Concept Mapping-Trainingsgruppe, die während der Trainingsphase Scaffolding und Feedback erhielt und in der Lernphase eine Concept Map ohne das Lernmaterial erstellte
T-R	Kontrolltrainingsgruppe, die in der Lernphase eine Concept Map ohne das Lernmaterial erstellte

Zusammenfassung

Concept Mapping wird als Lernmethode eingesetzt, um kognitive Prozesse der Organisation und Elaboration sowie metakognitive Prozesse bei Lernenden anzuregen und dadurch ihre Lernleistung zu verbessern. Erfolgreiches Lernen mit Concept Mapping erfordert Übung, da durch die unvertraute Anwendung einer Methode häufig eine hohe kognitive Belastung entsteht. Allerdings sind wichtige Parameter wie die Dauer und Intensität eines angemessenen Trainings für eine lerneffektive Anwendung von Concept Mapping noch umstritten. Die Methode wurde zudem bisher meist in einem Setting eingesetzt, in dem der Text den Lernenden während des Concept Mappings zur Verfügung stand (Elaborationssetting). In anderen Studien wurde die Methode jedoch basierend auf den Erkenntnissen zur Lernwirksamkeit von Abrufprozessen in einem Setting eingesetzt, in dem der Text während der Concept Map-Erstellung nicht verfügbar war (Retrievalsetting). Aufgrund uneinheitlicher Lernwirksamkeitsbefunde zum Concept Mapping ist davon auszugehen, dass das Training und die Textverfügbarkeit die für das Lernen notwendigen (meta)kognitiven Prozesse beeinflussen könnten. Zur Steigerung der Lernwirksamkeit des Trainings und Einsatzes von Concept Mapping wurde in zwei Studien untersucht, ob Lernende, die ein spezifisches Training erhalten, erfolgreicher im Lernen mit Concept Mapping sind als solche, die auf andere Weise an die Methode herangeführt werden. Ferner sollte herausgefunden werden, wie lerneffektiv der Einsatz des Concept Mappings in einem Elaborationssetting im Vergleich zum Retrievalsetting ist. Die erste Studie wurde mit $N = 73$ Studierenden durchgeführt und diente der Untersuchung von drei unterschiedlichen Trainingsansätzen. In der zweiten Studie mit $N = 93$ Studierenden wurden die Erkenntnisse der ersten Studie aufgegriffen und zwei Trainingsgruppen verglichen. Außerdem wurde die Textverfügbarkeit in der anschließenden Lernphase manipuliert, um herauszufinden, ob mögliche Effekte der beiden Settings von der Dauer und Intensität eines vorangegangenen Methodentrainings im Concept Mapping abhängen. Die Ergebnisse beider Studien geben Hinweise darauf, dass sich ein mehrwöchiges Training mit Scaffolding und Feedback positiv auf den Wissenserwerb und die Concept Mapping-Fertigkeiten auswirkt. Die Ergebnisse hinsichtlich der kognitiven Belastung sind uneindeutig. Ein Retrievalsetting führte, entgegen den Erwartungen, zu mehr Elaboration, jedoch hatte dies keinen Einfluss auf die Lernleistung. Dies könnte, trotz signifikanter Gruppenunterschiede, an dem generell niedrigen Maß an Elaboration gelegen haben. Ferner deuten die Ergebnisse der zweiten Studie darauf hin, dass ein Retrievalsetting metakognitive Prozesse anregt und zu einer akkurateren Einschätzung der eigenen Lernleistung führt. Concept Mapping sollte daher extensiv und mit Lernhilfen trainiert und für eine Förderung metakognitiver Prozesse ein Retrievalsetting gewählt werden.

Abstract

Concept mapping is used as a learning method to stimulate cognitive processes of organization and elaboration as well as metacognitive processes in learners and thereby improve their learning performance. Successful learning with concept mapping requires practice, as unfamiliarity with a method often results in high cognitive load. However, important parameters such as the duration and intensity of appropriate training for a learning effective use of concept mapping are still debated. Moreover, the method has mostly been used in a setting where the text was available to learners during concept mapping (elaboration setting). However, in other studies, based on the findings on the learning efficacy of retrieval processes, the method has been used in a setting where the text was not available during concept mapping (retrieval setting). Based on inconsistent learning efficacy findings on concept mapping, it can be assumed that training and text availability might influence the (meta)cognitive processes necessary for learning. To increase the learning effectiveness of the training and use of concept mapping, two studies investigated whether learners who receive specific training are more successful in learning with concept mapping than those who are introduced to the method in other ways. Furthermore, the aim was to find out how effective the use of concept mapping is in learning in an elaboration setting compared to a retrieval setting. The first study was conducted with $N = 73$ students and served to investigate three different training approaches. In the second study with $N = 93$ students, the findings of the first study were taken up and two training groups were compared. In addition, text availability in the subsequent learning phase was manipulated to determine whether possible effects of the two settings depended on the duration and intensity of a previous method training in concept mapping. The results of both studies provide evidence that several weeks of training with scaffolding and feedback have a positive effect on knowledge acquisition and concept mapping skills. Results regarding cognitive load are equivocal. A retrieval setting led to more elaboration, contrary to expectations, but this had no effect on learning performance. This could have been due to the generally low level of elaboration, despite significant group differences. Furthermore, the results of the second study suggest that a retrieval setting stimulates metacognitive processes and leads to a more accurate assessment of one's own learning performance. Concept mapping should therefore be trained extensively and with learning aids, and a retrieval setting should be chosen to promote metacognitive processes.

1 Einleitung

In einer zunehmend naturwissenschaftlich und technologisch bestimmten Welt sollen Lernende in der Schule und Universität eine Wissensbasis aufbauen, die ihnen eine Teilnahme am gesellschaftlichen und wissenschaftlichen Diskurs ermöglicht (z. B. Klieme et al., 2000; Ufer & Neumann, 2018). Eine solche nutzbare und fundierte Basis besteht in jedem Wissenschaftsbereich aus Fachbegriffen bzw. Konzepten, die für das jeweilige Fach spezifisch sein können und für ein Verständnis des Bereichs essenziell sind (Morse & Jutras, 2008; s. auch Kattmann, 1993). Dieses Verständnis setzt jedoch nicht nur die Kenntnis der einzelnen Konzepte voraus, sondern auch von den Beziehungen zwischen diesen Konzepten und erfordert ein Denken in Zusammenhängen bzw. ein vernetztes Denken. Das vernetzte Denken steht auch in Verbindung mit dem Ziel heutiger Lehre, Lernende nicht mehr ausschließlich zu befähigen, sich an Fakten zu erinnern, sondern eine effektive Speicherung von Wissen und die Unterstützung von Transferleistungen zu fördern, um mit dem erworbenen Wissen neue Probleme zu lösen (Dobson et al., 2019; Fraune, 2013; Hammann, 2019; Mayer, 2002b; Renkl, 2020).

Allerdings wird in zahlreichen Studien (Anderman et al., 2012; Assaraf et al., 2013; Grotzer et al., 2017; Hashem & Mioduser, 2013; Hmelo-Silver & Azevedo, 2006; Jacobson & Wilensky, 2006; Plate, 2010; Proctor & van Zandt, 2018; Raia, 2005, 2008; Tripto et al., 2017; Yoon et al., 2017) berichtet, dass Lernende häufig Probleme im Umgang mit komplexen Sachverhalten und dem dafür erforderlichen vernetzten Denken haben. Probleme der Lernenden äußern sich symptomatisch z. B. darin, dass sie Prozesse oder Sachverhalte inadäquat analysieren und rekonstruieren. Infolgedessen nehmen sie einen komplexen Sachverhalt nicht als Ganzes wahr, sondern fokussieren sich auf einzelne Komponenten, die sie in einfachen kausalen und linearen Beziehungen miteinander zu verbinden suchen (Eilam, 2012; Fraune, 2013; Raia, 2008; Reiner & Eilam, 2001). Die Gründe hierfür sind Gegenstand der didaktischen Forschung, um den Unterricht und das Lernen zu verbessern, wobei dies insbesondere für Naturwissenschaften, wie die Biologie, Physik und Chemie, wichtig erscheint: Die Inhalte dieser Wissenschaften zeichnen sich nämlich durch solche komplexen Beziehungen zwischen Konzepten aus, die es sinnvoll miteinander zu verbinden gilt, um eine strukturierte und flexibel einsetzbare Wissensbasis zu erlangen (Haugwitz & Sandmann, 2009; Hmelo-Silver & Azevedo, 2006; Holovatch et al., 2017; Lasker, 2019; Schmid & Telaro, 1990; Tripto et al., 2017; Yoon, 2008; Yoon et al., 2017).

Von diesen Naturwissenschaften steht im Fokus dieses Projektes die Biologie, in der Ursprung, Entwicklung, Erscheinungsformen und Eigenschaften von Lebewesen sowie ihre Struktur- und Funktionsvielfalt untersucht werden (Campbell et al., 2021; Sadava et al., 2019; Urry et al., 2019). Aufgrund dieser Untersuchungsgegenstände und ihren Beziehungen zueinander ist die Biologie durch unzählige Verbindungen zwischen einzelnen Komponenten gekennzeichnet. Diese Komponenten werden in biologischen Systemen geordnet: Ein biologisches System bildet ein Ganzes in Bezug auf seine Bestandteile, wie z. B. das System „Zelle“ als Ganzes den Zellkern und die Proteine als Bestandteile beinhaltet. Gleichzeitig ist jedes System Teil eines größeren Systems und steht im Austausch mit anderen Systemen, weshalb die Zelle als ein Teil des Organs und des Organismus gesehen werden muss, die wiederum Teile anderer Systeme sind (Brown & Schwartz, 2009). Aufgrund dieser Vielzahl an Verbindungen gleicht die Biologie „einem Gewebe miteinander vernetzter Konzepte ohne bestimmten Ausgangspunkt oder einen fest vorgeschriebenen Weg“ (Urry et al., 2019, S. XLIV) und verlangt von den Lernenden die Berücksichtigung dieser komplexen Beziehungen (Assaraf et al., 2013; Bramwell-Lalor & Rainford, 2014; Feigenspan & Rayder, 2017; Hegarty-Hazel & Prosser, 1991b; Heinze-Fry & Novak, 1990; Johnstone, 1991; Kinchin et al., 2005; Schmid & Telaro, 1990; Tibell & Rundgren, 2010; Tripto et al., 2017; Verhoeff et al., 2008).

Der Biologieunterricht hat folglich die Aufgabe der Vermittlung komplexer biologischer Inhalte, einschließlich miteinander verbundener Fakten, biologischer Konzepte, Prinzipien und disziplinärer Kernideen (Kultusministerkonferenz [KMK], 2020; Nawani et al., 2016; Wadouh et al., 2014). Die Aufgabe der Lernenden besteht dementsprechend darin, diese Inhalte im jeweiligen Bereich (z. B. Zellbiologie) zu lernen, aber auch mit anderen Bereichen (z. B. Physiologie und Evolution) zu vernetzen, wobei Hammann (2019) darauf verweist, dass insbesondere die Wissensvernetzung „seit vielen Jahren als eine grundsätzliche Möglichkeit angesehen [wird], die Effizienz des Unterrichts in den naturwissenschaftlichen Fächern zu verbessern“ (S. 78). Die Aufgabe der Lehrenden hingegen besteht darin, einerseits geeignetes Lernmaterial zur Verfügung zu stellen und andererseits den Lernenden Werkzeuge wie geeignete Lernstrategien an die Hand zu geben, damit sie sich Wissensinhalte aneignen und diese vernetzen können (s. auch Friedrich & Mandl, 2006).

Innerhalb der Biologie kommt der Zellbiologie, in der die Struktur, Funktion und das Verhalten von Zellen untersucht werden, eine besondere Rolle zu, da Lebewesen aus Zellen bestehen und die Zelle die „Grundeinheit der Biologie“ (Hardin et al., 2015, S. 8) darstellt. Daher gilt die Zellbiologie weithin als das wichtigste Fach zur Förderung der heutigen modernen Biologie (Suwono et al., 2021). Die biologische Zelle gilt insofern als repräsentativ

für biologische Systeme, da sie aus miteinander in Verbindung stehenden Komponenten (z. B. dem Zellkern und den Proteinen) besteht, die sowohl eigene Funktionen besitzen, aber auch für die Zelle als Ganzes von Bedeutung sind (Becker & Großschedl, 2019; Khodor et al., 2004; Morse & Jutras, 2008). Ein Verständnis der Zelle als Gesamtheit umfasst die Kenntnis über ihre Komponenten und deren Beziehungen zueinander und setzt vernetztes Denken voraus. Dieses Verständnis ist nicht nur wichtig, um Zellen zu verstehen, sondern auch viele Phänomene und biologische Prozesse sowie das allgemeine Funktionieren von Lebewesen (Barak et al., 1999; Brown & Schwartz, 2009; M. T. Chi et al., 1994; Del Fernández Fernández & Jiménez Tejada, 2019; Songer & Mintzes, 1994; Verhoeff, 2003; Verhoeff et al., 2008). Doch auch für den Bereich der Zellbiologie werden Probleme wie eine zu simple Herangehensweise an komplexe Sachverhalte berichtet, die dazu führt, dass Fakten isoliert voneinander gelernt und komplexe Zusammenhänge, z. B. zwischen einzelnen Zellbestandteilen wie Organellen, missachtet werden. Das Verständnis der Zelle, inklusive ihrer Bestandteile und Funktionen, bleibt dadurch limitiert und Lernende erwerben eine rudimentäre Wissensbasis, die wiederum zu niedrigen Lernerfolgen führen kann (z. B. Bramwell-Lalor & Rainford, 2014; Dreyfus & Jungwirth, 1989; Flores et al., 2003; Großschedl & Harms, 2013b; Hesse, 2002; Lewis et al., 2000a, 2000b; Novak, 2003; Songer & Mintzes, 1994).

Als mögliche Ursache für die Probleme Lernender im vernetzten Denken wird in der Forschung die Tendenz zum Auswendiglernen diskutiert (z. B. McClure et al., 1999; Novak, 2003; Watson et al., 2016). Auswendiglernen als Prozess kann ohne oder mit wenig Vorwissen stattfinden und ohne oder mit geringer Integration neuen Wissens in die bestehende Wissensstruktur. Damit soll nicht ausgedrückt werden, dass das Auswendiglernen keine Relevanz für das Lernen besitzt (s. auch Garvin-Doxas et al., 2007), allerdings hat sich gezeigt, dass die auswendig gelernten Informationen unverbunden gespeichert werden, sich damit nur schwer abrufen lassen und schneller vergessen werden (z. B. Novak, 1998; Okebukola, 1990). Das Auswendiglernen ist jedoch selbst nur ein Symptom eines größeren Problems, nämlich der Tendenz zur Abfrage reinen Faktenwissens, z. B. mittels Multiple-Choice-Tests, welches an Schulen und Universitäten dominiert (DiSessa, 2001; Hasnain & Bhamani, 2014; Hasnor et al., 2013; Novak, 2003). Derartige Wissenstests gelten zwar als objektiv (Suskie, 2018), aber sie begrenzen die Antworten der Lernenden und geben wenig Einblick in die Struktur ihres Wissens (McClure et al., 1999; Watson et al., 2016). Zudem hat eine solche Praxis zur Folge, dass Lernende nicht zum vernetzten Denken und dem Lernen von Zusammenhängen angeregt werden, sondern oft Fakten isoliert voneinander auswendig lernen, da sie damit häufig Erfolg haben (Mayer, 2002b; Montfort et al., 2009; Novak, 2003).

Dem Auswendiglernen wird von Ausubel (1963, 2000) im Rahmen seiner Assimilationstheorie¹ das bedeutungsvolle Lernen gegenübergestellt. Für effektives bedeutungsvolles Lernen gelten folgende Voraussetzungen: Erstens, gut organisierte und relevante Wissensstrukturen, die eine Grundlage für die Integration und Verbindung neuer Ideen oder Konzepte bilden; zweitens muss eine emotionale Verpflichtung bestehen, Neues in das vorherige Wissen zu integrieren, und drittens muss der Lernstoff für die Lernenden bedeutsam sein, so dass eine Integration neuer Informationen auf nicht-willkürliche, sondern gezielte und aktive Weise erfolgen kann (s. auch Driver, 1989; Novak, 1977). Dies unterstreicht zum einen die Bedeutung des Vorwissens für das Lernen und bedeutet zum anderen einen zusätzlichen kognitiven Aufwand für Lernende (DiCarlo, 2006). Doch dieser zusätzlich geleistete Aufwand wird als Grund dafür gesehen, dass sich bedeutungsvolles Lernen als wirksam erwiesen hat und Lernende deutlich besser abschneiden, wenn sie aufgefordert werden, ihr Wissen nicht nur wiederzugeben, sondern auch auf neue Probleme zu übertragen (Novak et al., 1983).

Die Aufgabe der schulischen und universitären Bildungseinrichtungen wäre folglich die Förderung vernetzten Denkens und bedeutungsvollen Lernens mittels geeigneter Instruktions- und Lernstrategien auf der einen, und die Einforderung solchen Denkens durch entsprechende Leistungsmessinstrumente auf der anderen Seite (Mintzes et al., 2001). Ein Werkzeug, welches in der fachdidaktischen Forschung für die genannten Zwecke empfohlen wird, ist das Erstellen von Concept Maps (Concept Mapping; CM), das sowohl als Lern- als auch Diagnoseinstrument eingesetzt werden kann (Allen & Tanner, 2002; Großschedl & Harms, 2018; Kinchin, 2001; Sumfleth et al., 2010). Concept Maps werden als nützliches Instrument zur Veranschaulichung von Zusammenhängen zwischen Konzepten gesehen, da sie Sachverhalte in Netzwerk-ähnlichen Diagrammen darstellen (Novak & Cañas, 2006, 2008). Das Netzwerk besteht aus „Knoten“ und den dazwischenliegenden Verbindungen. Als Knoten dienen Begriffe oder Konzepte, während in den Verbindungen die Beziehungen zwischen diesen definiert werden. Für die Definitionen der Beziehungen werden beschriftete Pfeile genutzt. Durch die Richtung eines Pfeils wird die Leserichtung vorgegeben und mit der Beschriftung auch die grammatische, semantische und logische Abhängigkeit der beiden Begriffe zueinander. Weiter voneinander entfernte Konzepte können dabei durch sogenannte Querverbindungen (*cross-links*) verbunden werden. Zwei miteinander verbundene Konzepte bilden die kleinste sinnvolle Einheit einer Concept Map,

¹ Im Rahmen der Assimilationstheorie wird davon ausgegangen, dass der wichtigste Faktor, der das Lernen beeinflusst, das Vorwissen ist, ergo das, was die bzw. der Lernende bereits weiß. Beim Lernen werden potenziell bedeutungsvolle Aspekte des Lernmaterials selektiv mit anderen relevanten Informationen aus dem Vorwissen verknüpft und können somit im Gedächtnis verankert werden (Ausubel, 2000).

eine sogenannte Proposition (s. Abbildung 1; Brandstädter et al., 2012; Cañas & Novak, 2014; Chevron, 2014; Cronin et al., 1982; Heinze-Fry & Novak, 1990; Novak, 1990a, 2010; Novak & Cañas, 2008).

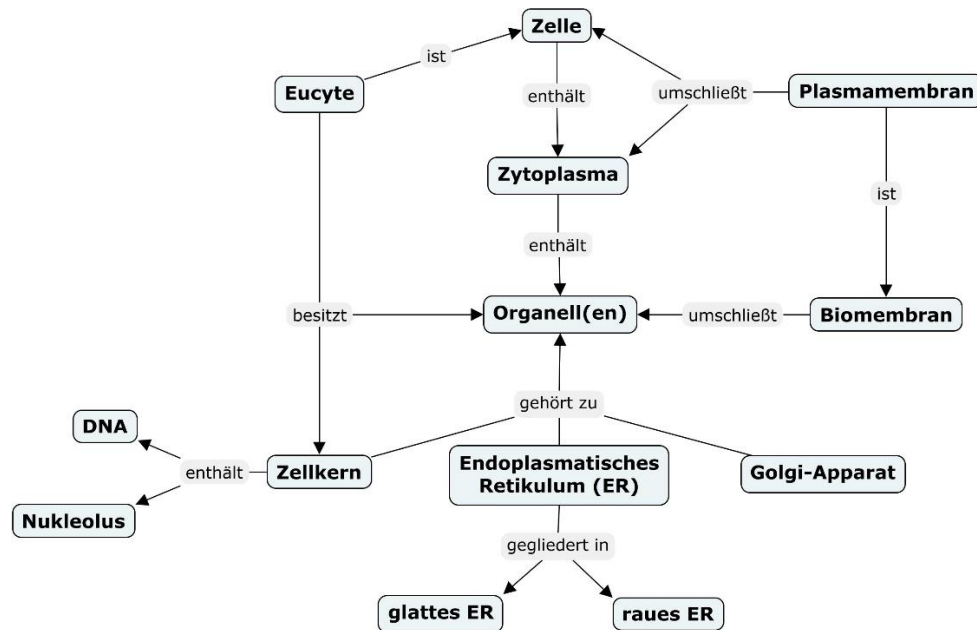


Abbildung 1. Concept Map zum Thema „Zelle“.

Concept Maps wurden als Diagnoseinstrument entwickelt, mit deren Hilfe die Vorstellungen von Lernenden erfasst werden sollten (Novak & Gowin, 1984). Da Concept Maps das individuelle Verständnis der bzw. des Lernenden für eine bestimmte Domäne darstellen (Kapuza, 2020; Ritter et al., 2019), wurden sie in vielen Studien zum bedeutungsvollen Lernen verwendet (Andrews et al., 2008; B. L. Martin et al., 2000; Mintzes et al., 2001; Pearsall et al., 1997; Quinn et al., 2003). Darüber hinaus deuten Befunde darauf hin, dass Concept Maps auch effektive Lerninstrumente darstellen, die das Lernen abstrakter Themen und komplexer Beziehungen erleichtern, indem sie das vernetzte Denken unterstützen und das bedeutungsvolle Lernen fördern (Bramwell-Lalor & Rainford, 2014; Cañas et al., 2003; Hay et al., 2008; Hilbert & Renkl, 2009; Horton et al., 1993; Karakuyu, 2010; Kinchin, 2000b; Morse & Jutras, 2008; Nesbit & Adesope, 2006; Novak, 1990a; Novak & Gowin, 1984; O’Donnell et al., 2002; Okebukola, 1990; Okebukola & Jegede, 1988; Pruettt & Weigel, 2020). Ferner wird davon ausgegangen, dass das Erstellen von Concept Maps metakognitive Prozesse fördert, da es Lernenden durch die Visualisierung ihres Wissens hilft, Wissenslücken oder Verständnisschwierigkeiten selbst zu erkennen (Chularut & DeBacker, 2004; Friedrich & Mandl, 2006; Hilbert & Renkl, 2008; Lumer & Winter, 2019; Nesbit & Adesope, 2006; Novak, 1990a; Novak & Cañas, 2008; Roth & Roychoudhury, 1993). Solche Erkenntnisse können Lernende

anschließend nutzen, um ihre Wissenslücken zu schließen und ihr Lernen zu optimieren (Lumer & Winter, 2019).

Die Lernwirksamkeit von Concept Mapping wird jedoch aufgrund der Ergebnisse verschiedener Studien infrage gestellt, in denen Concept Mapping nicht lernwirksam eingesetzt werden konnte und zu niedrigeren oder ähnlichen Lernerfolgen führte wie bspw. das Notizen Erstellen (z. B. Großschedl & Harms, 2013b; Markow & Lonning, 1998; Reader & Hammond, 1994). Als möglicher Grund für diese Lernwirksamkeitsunterschiede wird das Maß an Vertrautheit der Lernenden mit dem Concept Mapping diskutiert, da bei näherer Betrachtung der methodischen Unterschiede zwischen den Studien auffällt, dass die Trainingsmaßnahmen vor dem Concept Mapping-Einsatz stark variieren (Mintzes et al., 2001; Nesbit & Adesope, 2006; Schroeder et al., 2018; s. auch Eshuis et al., 2021; Jegede et al., 1990). Wie aus Untersuchungen zum Einsatz von Lernstrategien jedoch bekannt ist, kann der ungeübte Einsatz von Lernstrategien das Lernen behindern (z. B. Cañas et al., 2003; Friedrich & Mandl, 2006), da die Lernenden ihre begrenzten kognitiven Kapazitäten nicht nur der Verarbeitung der Lerninhalte, sondern auch der kontrollierten Ausführung der Lernstrategie widmen müssen. Diese Bindung der Kapazitäten durch den Lernstrategieinsatz hat zur Folge, dass die Lernenden sich weniger auf die Lerninhalte selbst konzentrieren können (Klepsch et al., 2017; Sweller et al., 1998; Sweller, 2005).

Dies gilt auch für die Lernmethode des Concept Mappings, da Forscher:innen (z. B. Allen & Tanner, 2003; Cañas et al., 2003; Jüngst & Strittmatter, 1995; Michalak & Müller, 2017; Mintzes et al., 2011) bezüglich der Konfrontation mit Concept Mapping als unbekannter Lernmethode von Problemen Lernender berichten. Bei Michalak und Müller (2017) äußerten sich die Probleme der Lernenden in einer linearen Herangehensweise an die Diagramme, da sie mit der ganzheitlichen Betrachtung einer Grafik nicht vertraut waren. Allen und Tanner (2003) beschreiben, dass ein Mangel an Vertrautheit mit der Methode zu Schwierigkeiten führte, eine sinnvolle Concept Map zu erstellen und die Ersteller:innen auf externe Hilfe angewiesen waren. Die Bedeutung einer methodischen Heranführung in Form von Trainingsmaßnahmen für den souveränen und lernförderlichen Einsatz von Concept Mapping ist daher ersichtlich und wird von vielen Forscher:innen empfohlen (z. B. Haugwitz & Sandmann, 2009; Hilbert, Nückles, Renkl, et al., 2008; Hilbert & Renkl, 2005, 2008; Jegede et al., 1990; Jüngst & Strittmatter, 1995; McCagg & Dansereau, 1991; Morse & Jutras, 2008; Renkl & Nückles, 2006; Roessger et al., 2018; Romero et al., 2017; Sumfleth et al., 2010). Diese Empfehlungen werden durch die Ergebnisse von Metaanalysen zum Concept Mapping gestützt, in denen ein positiver Einfluss

von Trainingsmaßnahmen auf den Lernerfolg festgestellt werden konnte (Nesbit & Adesope, 2006; Schroeder et al., 2018).

Die Trainingsmaßnahmen und -empfehlungen variieren jedoch nicht nur bezüglich der Trainingsdauer und -intensität, sondern auch der vermittelten Inhalte und der Art und Weise, wie sie vermittelt werden. Während in einigen Studien eine Kurzeinführung in das Concept Mapping gegeben wurde (z. B. Blunt & Karpicke, 2014; Karpicke & Blunt, 2011b), nutzten andere Forscher:innen ein extensives Training, welches sich oft über mehrere Sitzungen erstreckte und zusätzliche Lernhilfen wie z. B. Feedback zu den erstellten Concept Maps beinhaltete (z. B. Jegede et al., 1990; Roessger et al., 2018). Aus den Befunden lassen sich daher keine klaren Empfehlungen bezüglich dieser Trainingsparameter ableiten, da sie mit unterschiedlichen Ergebnissen im Sinne des Lernerfolgs assoziiert sind (z. B. Ajaja, 2013; Chiu, 2004; Chularut & DeBacker, 2004; Jegede et al., 1990). Es ist daher weiterhin unklar, wie Concept Mapping vermittelt und trainiert werden soll, damit Lernende ein Maß an Vertrautheit mit der Strategie erlangen, das es ihnen ermöglicht, Concept Mapping lernwirksam zu nutzen.

Neben der Frage, wie Concept Mapping trainiert werden soll, wird in der Forschung ebenfalls diskutiert, zu welchem Zeitpunkt Concept Mapping am lernwirksamsten in Lernsituationen eingesetzt werden sollte (z. B. Blunt & Karpicke, 2014; O'Day & Karpicke, 2020). Die meisten Studien, in denen Concept Mapping als Lernmethode eingesetzt wurde, beinhalteten eine kritische Lernphase, in der die Lernenden mit dem bereitgestellten Material (bspw. einem Text) Concept Maps zu den Lerninhalten erstellten. Das Lernmaterial steht den Lernenden dabei zur Verfügung, um ihnen einerseits zu helfen, z. B. den Text mittels Concept Mapping besser zu verstehen, bevor ihre Lernleistung gemessen wird (z. B. Karpicke & Blunt, 2011b; Nesbit & Adesope, 2006; Sumfleth et al., 2010). Andererseits soll der Abgleich zwischen der eigenen Concept Map und dem Text das Erkennen von Verständnislücken und somit metakognitive Prozesse fördern (Bramwell-Lalor & Rainford, 2014; Kinchin, 2011; Mintzes et al., 1997; Ritchhart et al., 2009; Salmon & Kelly, 2014). In anderen Studien (z. B. Ruiz-Primo, Shavelson, et al., 2001) wurde Concept Mapping ausschließlich ohne verfügbares Lernmaterial eingesetzt, um z. B. die Wissensstrukturen der Lernenden zu erfassen. Da unter diesen Bedingungen jedoch die Schritte der Erstellung einer Concept Map ausgeführt werden und davon ausgegangen werden kann, dass Concept Mapping lernförderlich ist, gehen manche Forscher:innen wie Blunt und Karpicke (2014) davon aus, dass auch das Concept Mapping ohne verfügbares Lernmaterial einen positiven Effekt auf das Lernen haben könnte. Eine Gegenüberstellung der Bedingungen zur Evaluierung, welche Herangehensweise effektiver die Lernleistung steigern kann, wurde allerdings nur in wenigen Studien durchgeführt

(z. B. Blunt & Karpicke, 2014; O'Day & Karpicke, 2020). O'Day und Karpicke (2020) verweisen in ihrem Bericht selbst auf den Mangel an Forschung zur Anwendung von Lernmethoden wie dem Concept Mapping in einer Bedingung, in der den Lernenden das Material nicht zur Verfügung steht. Letztgenannte Bedingung stellte sich in ihrer Studie als effektiver heraus im Vergleich zu Bedingungen, in denen das Lernmaterial verfügbar war, wobei auch hier Nachteile der Concept Mapping-Gruppen bei anschließenden Leistungstests im Vergleich zu anderen Lernmethoden (z. B. Notizen erstellen) gefunden wurden.

Aus fachdidaktischer Perspektive können die Ergebnisse der Studien von Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) jedoch nur begrenzt als verlässlich eingestuft werden. Zum einen kann kritisiert werden, dass die Forscher:innen ihren Proband:innen lediglich eine Kurzeinführung zu den Grundlagen des Concept Mappings gaben, die einen Mangel an Übung und Vertrautheit mit Concept Mapping seitens der Lernenden bedeutet haben könnten. Diese Art von Training wurde schon vorher von Forscher:innen wie Mintzes et al. (2011) moniert, deren Kritik durch die oben erwähnten Erkenntnisse von Einzelstudien und Metaanalysen gestützt wird, die nahelegen, dass die Lernleistung beim Concept Mapping von einer trainingsinduzierten Vertrautheit mit der Methode abhängt (z. B. Schroeder et al., 2018). Zum anderen nutzten Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) sehr kurze Lerntexte (in beiden Studien waren es weniger als 300 Wörter pro Text), welche im Alltag der Studierenden eine Belastung darstellen, die unter dem Level liegt, das sie gewöhnlich bei der Arbeit mit deutlich umfangreicheren Texten leisten müssen.

Aus den diversen Trainingsmaßnahmen und den unterschiedlichen Herangehensweisen, Concept Mapping einzusetzen sowie den daraus potenziell resultierenden Lernwirksamkeitsunterschieden, lassen sich Desiderata zur Wirksamkeit unterschiedlicher Trainingsmaßnahmen und Einsatzmöglichkeiten ableiten. Darauf aufbauend bestehen die Ziele dieser Arbeit darin, methodische Einführungen in das Concept Mapping und dessen Einsatz zu untersuchen, um anhand der Ergebnisse Empfehlungen zu einer effektiveren Lernförderung geben zu können.

1. Eine Gegenüberstellung unterschiedlicher Trainingsansätze im Concept Mapping mit Lerntexten, die für Studierende eine studienalltagsgetreue Belastung darstellen, soll Aufschluss darüber geben, ob die Lernwirksamkeit der Methode von Trainingsparametern wie der Dauer, Intensität und der Bereitstellung von Lernhilfen abhängt.
2. Ein Vergleich der Trainingsansätze mit anschließender Kombination der beiden Ansätze (Concept Mapping bei verfügbarem bzw. nicht-verfügbarem Lernmaterial) soll Klarheit bezüglich der Frage bringen, wie Concept Mapping am wirksamsten eingesetzt werden kann, um den Wissenserwerb und das Verständnis der Lernenden zu fördern.

2 Theoretischer Hintergrund

2.1 Wissen und Lernstrategien

In der (fach)didaktischen Forschung wird versucht herauszufinden, wie Menschen lernen und wie Lehr-Lern-Situationen gestaltet sein müssen, um ihre Lerneffektivität zu steigern (z. B. Helmke, 2021; Jegede et al., 1990; P.-Y. Martin & Nicolaisen, 2015b; Meyer, 2018). Das Ziel, bei Lernenden das vernetzte Denken und bedeutungsvolle Lernen zu fördern, ist in diesem Kontext zu sehen und wird mithilfe der Vermittlung von Lernstrategien zu erreichen versucht, die Lernenden dabei helfen sollen, ihre Lerneffizienz zu optimieren (P.-Y. Martin & Nicolaisen, 2015a). Lernstrategien stellen komplexe kognitive Aktivitäten der Informationsbearbeitung dar, welche zur erfolgreichen Bearbeitung einer lernspezifischen Aufgabe beitragen (Baumert, 1993; Pressley et al., 1985). Friedrich und Mandl (2006) erweiterten diese Sichtweise auf Lernstrategien und definieren sie als „jene Verhaltensweisen und Gedanken, die Lernende aktivieren, um ihre Motivation und den Prozess des Wissenserwerbs zu beeinflussen und zu steuern“ (S. 1). In diesem Sinne stellt die Anwendung von Lernstrategien einerseits ein Handeln dar, das „die Art und Weise der Aneignung von Lerninhalten“ (Wernke, 2013, S. 15) bestimmt. Andererseits berücksichtigen Friedrich und Mandl (2006) im Gegensatz zu Pressley et al. (1985) zusätzlich die motivationale und affektive Komponente im Lernstrategiebegriff. In der vorliegenden Arbeit stellt der Begriff Lernstrategie auch eine Kategorie dar, die wiederum gemäß Weinstein und Mayer (1986) unterteilt wird in kognitive, metakognitive Lernstrategien und Strategien des Ressourcenmanagements (s. auch Baumert, 1993; Pintrich et al., 1991; K. P. Wild & Schiefele, 1994). Zu den Lernstrategien werden konkrete Lernmethoden wie das Concept Mapping gezählt, das u. a. als Organisationsstrategie gilt (Mandl & Friedrich, 2006). Bevor jedoch auf diverse Lernstrategien eingegangen wird, werden zwei Prozesse vorgestellt, die durch den Einsatz von Lernstrategien gefördert werden sollen: der Wissenserwerb und die Wissensorganisation.

2.1.1 Wissenserwerb und Wissensorganisation

Wissenserwerb bzw. Lernen bezeichnet die Aneignung von Wissen, wobei sowohl der Erwerb deklarativen („Wissen, dass“) als auch prozeduralen Wissens („Wissen, wie“, vgl., J. R. Anderson, 1976, S. 219) im Fokus dieser Arbeit stehen. Auf das prozedurale Wissen, welches sich auf das Ausführen einer Aktion bezieht (Renkl, 2020), wird zu einem späteren Zeitpunkt eingegangen (s. Kapitel 2.1.2 und 2.1.4 sowie 2.1.5). Als deklaratives Wissen werden Informationen bezeichnet, die über den bewussten und gezielten Abruf zugänglich sind, sofern sie sich auf Ereignisse und persönliche Erfahrungen (Inhalte des *episodischen Gedächtnisses*) oder auf Fakten, Konzepte, Prinzipien und Regeln (Inhalte des *semantischen Gedächtnisses*)

beziehen (Tulving, 1972). Folglich wird im semantischen Gedächtnis schulisches bzw. universitäres und berufliches Wissen gespeichert (z. B. Zellen besitzen eine Plasmamembran). Während im episodischen Gedächtnis die Speicherung spezifisch abläuft, damit z. B. die chronologische Abfolge von Ereignissen rekonstruiert werden kann, erfolgt die Speicherung im semantischen Gedächtnis insofern unspezifischer, als dass Informationen abstrahiert und nach Bedeutungszusammenhängen dekontextualisiert zusammengeführt werden (Edelmann & Wittmann, 2019; Tulving, 1972). Basierend auf dem Postulat von R. C. Anderson (1984) „the essence of knowledge is structure“ (S. 5) wird angenommen, dass Wissen organisiert ist und beide Gedächtnisse eine Form der Struktur bzw. Organisation erfordern (Goldsmith et al., 1991; T. Gruber, 2018). Die Form der Organisation dient zudem als Basis für die Integration und Speicherung neuer Informationen und ist für das Lernen elementar, da Lernen laut Renkl (2020) letztlich bedeutet „Informationen mit bereits vorhandenen Wissens-elementen zu vernetzen [...] Man könnte auch sagen, Lernen ist Andocken neuer Informationen an das Vorwissen“ (S. 9).

Bezüglich der Frage *wie* Wissen organisiert ist, wird in dieser Arbeit auf das propositionale Netzwerkmodell von Quillian (1968; Collins & Quillian, 1969) Bezug genommen. Im Rahmen von Netzwerkmodellen wird davon ausgegangen, dass das Wissen innerhalb eines Inhaltsbereichs um zentrale Konzepte herum organisiert ist (s. auch Stasz et al., 1976). Diese Konzepte werden dann entsprechend ihrer Verwandtschaft miteinander verbunden, wobei kürzere Abstände zwischen zwei Begriffen enger Verwandtschaft bestehen (*Baum – Wurzeln*) und längere Abstände zwischen ferner verwandten Konzepten (*Baum – Buch*). Diese Beziehungen werden ebenfalls im propositionalen Netzwerkmodell von Quillian (1968; Collins & Quillian, 1969) dargestellt: In diesem Modell wird davon ausgegangen, dass Gedächtnisinhalte als Begriffe mehr oder weniger hierarchisch in semantischen Netzwerken angeordnet werden (Assaraf & Orion, 2005; Ausubel et al., 1978; Buchner & Brandt, 2017; Renkl & Nückles, 2006). Die Struktur des Netzwerks ergibt sich daher einerseits aus der Existenz von Konzepten sowie ihrer Verbindungen zueinander und andererseits aus der Vorstellung darüber, welche Beziehungen zwischen den Konzepten existieren. Diese Beziehungen können semantischer, logischer und grammatischer Natur sein und entsprechend definiert werden. Durch die Beziehung zweier Konzepte und ihrer gemeinsamen Beziehungen zu weiteren Konzepten ergibt sich die semantische Verwandtschaft der Begriffe (Collins & Loftus, 1975; Jonassen et al., 1993). Zwischen den Begriffen *Baum* und *Wurzeln* besteht dementsprechend eine Verbindung wie z. B. *Baum – ist verankert über → Wurzeln*, in der die Beziehung zwischen beiden explizit wird. Dagegen wird die Verbindung zwischen *Baum* und

Buch erst durch weitere Verbindungen über andere Begriffe wie *Holz*, das vom Baum stammt, und *Papier*, das aus Holz gewonnen wird und aus dem das *Buch* besteht, deutlich.

Eine gute Kenntnis in einem Fach oder Bereich setzt demnach eine hochgradig integrierte konzeptuelle Struktur voraus, da sie die viele solcher Verbindungen zwischen fachspezifischen Konzepten besitzt (Ruiz-Primo & Shavelson, 1997). Bezüglich der Aktivierung und Nutzung der Informationen des semantischen Gedächtnisses gehen Collins und Loftus (1975) davon aus, dass jedes Konzept durch äußere Einflüsse aktiviert werden kann. Nach der Aktivierung eines Konzepts können weitere, mit dem ursprünglich aktivierten Konzept verbundene Konzepte, aktiviert werden, die wiederum mit ihnen verbundene Konzepte aktivieren können, weshalb diese Form der fortlaufenden, ausfächernden Aktivierung als *spreading activation* bezeichnet wird. Bezogen auf die gute Kenntnis in einem Fach oder Bereich bedeutet dies, dass durch die Struktur und Organisation der Verbindungen entsprechend viele fachspezifische Konzepte aktiviert werden können. Die logische Konsequenz ist zum einen die Annahme, dass grafische Methoden wie Concept Maps (die den gleichen Prinzipien folgen) in der Lage sind, Strukturen und Inhalte solcher kognitiven Repräsentationen abzubilden und zum anderen ihre Entstehung und bedeutungsvolles Lernen zu erleichtern, da den Begriffen durch die Erstellung von Verbindungen zu anderen ihre Bedeutung beigegeben wird (Mandl & Fischer, 2000).

2.1.2 Klassifizierung von Wissen

Die Ergebnisse des bedeutungsvollen Lernens können mithilfe von Blooms (1956) Taxonomie der kognitiven Lernziele kategorisiert werden (Krathwohl, 2002): Seine Taxonomie besteht aus den kognitiven Bereichen *Wissen*, *Verständnis*, *Anwendung*, *Analyse*, *Synthese* und *Evaluation*, wobei Bloom davon ausgeht, dass die jeweilige nächste Stufe der Taxonomie nur durch erfolgreiche Absolvierung der Vorstufe erreicht werden kann. Dadurch soll die Entwicklung des Wissens vom Aufbau des Vorwissens bis hin zur Evaluation komplexer Denkprozesse verfolgt werden können. L. W. Anderson et al. (2001) überarbeiteten die Bloom'sche Taxonomie und gehen von den kognitiven Prozessdimensionen *Erinnern*, *Verstehen*, *Anwenden*, *Analysieren*, *Beurteilen* und *Schaffen* aus. Zudem erweiterten sie diese um die Wissensdimension mit den vier Hauptkomponenten *deklaratives*, *konzeptuelles*, *prozedurales* und *metakognitives Wissen* (L. W. Anderson et al., 2001; Krathwohl, 2002), die im Folgenden als Wissensdomänen bezeichnet werden. In diesem Zusammenhang wird davon ausgegangen, dass Wissen von eher „konkret“ bis zu eher „abstrakt“ geordnet werden kann, wobei die reine Reproduktion eines gelernten Faktes (*Erinnern*) weniger abstrakt ist als das *Anwenden* von Wissen in einem neuen Bereich (s. auch Chevron, 2014).

Deklaratives Wissen im Sinne von Faktenwissen² entspricht dem Wissen über isolierte Elemente wie Fachbegriffe oder spezifische Details. Dieses Wissen wird von Krathwohl (2002) als wichtig für Lernende angesehen, um mit einem Fachgebiet vertraut zu sein oder um Probleme in diesem Fachgebiet zu lösen. Neben dem deklarativen Wissen ist für dieses Projekt das strukturelle Wissen von Bedeutung, das die Struktur des Wissens beschreibt, und im Rahmen dieses Projektes als eigene „Hauptkomponente der Wissensdomäne“ (um bei der Terminologie von L. W. Anderson et al., 2001, zu bleiben) behandelt wird. Während diese Sichtweise von einigen Forscher:innen geteilt wird (Anohina-Naumeca, 2015; Jonassen et al., 1993; Kinchin & Cabot, 2010), betrachten andere das strukturelle Wissen als ein Merkmal anderer Wissenstypen, z. B. als Struktur des konzeptuellen Wissens (z. B. Großschedl & Harms, 2013b). In diesem Fall wird es als eine Kombination aus Wissenstyp und dessen Qualität definiert (Jong & Ferguson-Hessler, 1996; Shavelson et al., 2005). Ungeachtet dieser Feinheiten kann strukturelles Wissen in beiden Fällen als eine Einheit betrachtet werden (vgl. Jonassen et al., 1993) und als „Struktur des Wissens“ wurde es in Studien z. B. mithilfe von *Similarity Judgments Tests* (SJT; s. Kapitel 4, 2.4.5 und 5, 2.4.3³) erhoben (Goldsmith et al., 1991; Großschedl & Harms, 2008, 2013a). Das strukturelle Wissen bzw. die Struktur des Wissens kann mittels eines SJTs erfasst werden, indem die semantische Verwandtschaft bzw. Nähe zweier gegebener Begriffe (im zellbiologischen Zusammenhang z. B. *Nukleus* – *Biomembran* oder *Ribosom* – *DNA*) auf einer 9-stufigen Skala (1 = geringfügig verwandt bis 9 = stark verwandt) bewertet wird. Durch die Verbindungen zwischen den Begriffen und ihrer Nähe bzw. Distanz entsteht somit ein Netz, das den kognitiven Strukturen der Person ähneln soll. Im Gegensatz zu Concept Maps, die von einem propositionalen Gedächtnismodell ausgehen (s. Kapitel 2.1.1), werden im SJT die Verbindungen nicht explizit definiert, was zumindest als Erklärung für eine Unterscheidung in den Ansichten zum strukturellen Wissen aufgrund der theoretischen Fundierung gesehen werden kann (Royer et al., 1993; Shavelson et al., 2005).

² Um das deklarative Wissen im Sinne von Faktenwissen (Anderson et al., 2001) sprachlich vom deklarativen Wissensbegriff Tulvings (1972; s. Kapitel 2.1) abzugrenzen, nutzen Wissenschaftler:innen, wie z. B. Großschedl (2010), den Terminus *Faktenwissen*. Da in den veröffentlichten Originalarbeiten (s. Kapitel 4 und 5) jedoch der Begriff *declarative knowledge* im Sinne von L. W. Anderson et al. (2001) genutzt wurde, um das Faktenwissen zu beschreiben, wird im Rahmen dieser Arbeit die entsprechende deutsche Übersetzung gewählt.

³ An dieser Stelle wird darauf verwiesen, dass die veröffentlichten Manuskripte, die der Dissertation zur Verfügung gestellt wurden, jeweils in sich geschlossenen Abschnitte darstellen, die eine eigene Kapitel-, Abbildungs- und Tabellenummerierung sowie ein eigenes Literaturverzeichnis besitzen. Im Kapitel 4 ist das veröffentlichte Manuskript über die Studie 1 von Becker, Welter, Aschermann und Großschedl (2021). In Kapitel 5 ist das veröffentlichte Manuskript über die Studie 2 von Becker, Welter und Großschedl (2021). Verweise in der übrigen Arbeit auf einzelne Kapitel, Abbildungen und Tabellen der Manuskripte werden mit einem Verweis auf das Kapitel 4 (Studie 1) bzw. 5 (Studie 2) eingeleitet, worauf die Nummer des Unterkapitels (z. B. „s. Kapitel 4, 3.2“) bzw. der Verweis auf die Abbildung oder Tabelle folgt (z. B. „s. Kapitel 4, Tab. 4“).

Konzeptuelles Wissen schließt das allgemeine Wissen über Fakten ein (Watson et al., 2016), bezieht sich jedoch auch auf das Wissen über Beziehungen zwischen den (vermeintlich) isolierten Fakten bzw. Elementen und wie sie organisiert sind (Hiebert & Lefevre, 1986; Jong & Ferguson-Hessler, 1996). Darüber hinaus beinhaltet konzeptuelles Wissen das Wissen über Kategorien, Prinzipien und Verallgemeinerungen, sowie Theorien, Modelle und Strukturen. Der Vorteil konzeptuellen Wissens liegt für Lernende darin, dass es i. d. R. auf neue Situationen und Kontexte übertragbar (Budé et al., 2011; Paas, 1992) und somit flexibel anwendbar ist (Montfort et al., 2009). Im Vergleich zu reinem Auswendiglernen von Fakten ist der Erwerb konzeptuellen Wissens jedoch oft schwieriger zu erreichen (M. T. H. Chi & Roscoe, 2002; Montfort et al., 2009; Watson et al., 2016), da es voraussetzt, dass Lernende die Verbindungen zwischen Elementen identifizieren und verstehen (McCormick, 1997; Yu et al., 2015). Dies erfordert eine gute Gedächtnisorganisation, in der die zentralen Begriffe einer Domäne in sinnvoller Weise miteinander in Beziehung stehen, die allerdings bei Lernenden nicht immer vorhanden ist (L. W. Anderson et al., 2001; Goldsmith et al., 1991; Krathwohl, 2002).

Während deklaratives Wissen als das „Wissen, dass“ bezeichnet wird, bezieht sich das prozedurale Wissen auf das „Wissen, wie“ (vgl. J. R. Anderson, 1976, S. 219), ergo darauf, wie etwas zu tun ist (s. auch Renkl, 2020; Shavelson et al., 2005). Prozedurales Wissen schließt einerseits die Kenntnis über Techniken und Methoden ein, andererseits die Kenntnis über Kriterien, anhand derer eine Entscheidung getroffen wird, wann ein Verfahren angemessen ist (Krathwohl, 2002; s. auch Kapitel 2.1.4). J. R. Anderson (1983) stellt das deklarative dem prozeduralen Wissen in seiner *Adaptive Control of Thought-Theorie* (kurz: ACT-Theorie) gegenüber und postuliert, dass allem Können (prozeduralem Wissen) das deklarative Wissen vorausgeht. Allerdings kann durch das Sammeln von Erfahrung mit der Ausführung der Aktion im Sinne der Wissenskompilierung deklaratives Wissen in prozedurales Wissen übertragen werden. Die Erfahrungen können z. B. im Rahmen eines Trainings gesammelt werden, in dem Lernende einzelne Prozesse zu größeren Ablaufrastern zusammengefasst und automatisieren können, was kognitive Ressourcen freisetzt, die nun anderweitig zur Verfügung stehen (Royer et al., 1993; Steiner, 2006; s. Kapitel 2.1.5).

Als letzte Domäne nennen L. W. Anderson et al. (2001) das metakognitive Wissen. Dieses Wissen gehört mit der metakognitiven Kontrolle zum Oberbegriff der Metakognition, welche die Fähigkeit bezeichnet, die Prozesse der Kognition, Motivation und des Verhaltens im Rahmen des Lernens und Problemlösens zu steuern und zu regulieren (Flavell, 1971, 1984; s. auch Hasselhorn, 1992; Pintrich, 2000). Metakognitives Wissen umfasst dabei das Wissen über Kognition im Allgemeinen sowie das Bewusstsein über die eigenen kognitiven Prozesse

bzw. deren Produkte (Krathwohl, 2002). Darunter fällt bei L. W. Anderson et al. (2001) auch das metakognitive Strategiewissen (*strategic knowledge*), über das Lernende für einen erfolgreichen Einsatz von Lernstrategien verfügen sollten. Das metakognitive Strategiewissen kann weiter unterteilt werden in das deklarative und prozedurale Strategiewissen⁴: Dabei umfasst das deklarative Strategiewissen Kenntnisse über eingesetzte Lernstrategien wie Nutzen und Funktion der Strategie und Strategieschritte, während das prozedurale Strategiewissen all jenes Wissen einschließt, welches zur erfolgreichen Strategieranwendung vonnöten ist und der Ausführung deklarativen Strategiewissens dient (Harris et al., 2009; Leopold et al., 2006; Lipowsky & Hess, 2019; Philipp, 2021; Sturm & Weder, 2016). Metakognitive Kontrolle bezieht sich auf die Überwachungsvorgänge und Maßnahmen, die im Hinblick auf kognitive Prozesse aktiv ergriffen werden (Hasselhorn, 1992). Nichtsdestotrotz sind beide, metakognitives Wissen und metakognitive Kontrolle, verstärkt im Fokus der Lernforschung, da sich gezeigt hat, dass sie entscheidend für einen bewussten und reflektierten Lernprozess sind und erheblich zum Lernerfolg beitragen (Stark et al., 2008, s. Kapitel 2.2.1).

Der Wissenserwerb in diesen Domänen und die Ausbildung kohärenter Wissensstrukturen sollen, wie zu Beginn dieses Kapitels erwähnt, durch die Anwendung von Lernstrategien (s. Kapitel 2.1.4) unterstützt werden. Hierbei soll v. a. das sogenannte bedeutungsvolle Lernen gefördert werden, das zur Ausführung der von L. W. Anderson et al. (2001) kategorisierten kognitiven Prozesse (*verstehen, anwenden, analysieren, etc.*) befähigt, die über das reine *Erinnern* im Sinne der Faktenwiedergabe hinausgehen (Großschedl, 2010; Krathwohl, 2002). Daher gilt es als wichtiges Erziehungsziel (Mayer, 2002a) und ist insbesondere für die Naturwissenschaften relevant, wie im Folgenden gezeigt wird.

2.1.3 Grundlagen des bedeutungsvollen Lernens

Über den Einsatz von Lernstrategien sowohl an Schulen als auch an Universitäten wird berichtet, dass Lernende in ihrem Lernalltag nur wenige Strategien nutzen und eine Tendenz zum Auswendiglernen zeigen (Novak, 2003). Ob ihre Lernprozesse als erfolgreich gelten, hängt von der Prüfung bzw. dem Test ab, der das Wissen erfasst, weshalb die Tendenz zum reinen Auswendiglernen indiziert, dass es von Lernenden mit erfolgreichem Lernen gleichgesetzt wird (Mayer, 2002a; Montfort et al., 2009; Novak, 2003). Rein auswendiggelernte Informationen können jedoch, wenn überhaupt eine Verankerung im Langzeitgedächtnis stattfindet, nur unverbunden und isoliert von anderen verwandten Informationen gespeichert werden. Mit dem

⁴ Als dritte Komponente des metakognitiven Strategiewissens wird das konditionale Strategiewissen gezählt, mithilfe dessen Lernende ihre Entscheidung fällen können, wann und warum sie eine gewisse Lernstrategie einsetzen, um einen Lernerfolg zu erzielen (Hasselhorn (1992). Da im vorgestellten Projekt die Lernmethode jedoch vorgegeben wurde, wird es nicht näher erläutert.

Vorwissen unverbunden gespeicherte Informationen sind daher, wie in Kapitel 2.1.1 dargelegt, nicht über mehrere Verbindungen zugänglich und lassen sich nur schwer abrufen (Okebukola, 1990). Ein Ziel der Bildung im naturwissenschaftlichen Kontext ist somit die Förderung der Fähigkeit, nicht nur isolierte Fakten zu wissen, sondern auch Zusammenhänge zu erkennen, zu verstehen und das Gelernte im Rahmen eines Wissenstransfers anzuwenden (Mayer, 2002a; Mestre, 2002). Dafür ist eine andere Art des Lernens als das Auswendiglernen notwendig, die Ausubel (1963, 2000) im Rahmen seiner Assimilationstheorie als bedeutungsvolles Lernen bezeichnet und dem Auswendiglernen gegenübergestellt. Beim bedeutungsvollen Lernen werden Verbindungen zwischen (vermeintlich isolierten) Informationen hergestellt und organisiert (kognitive Prozesse der Organisation) sowie in bestehende Wissensstrukturen integriert (kognitive Prozesse der Elaboration; J. R. Anderson, 1983; Weinstein & Mayer, 1986).

Neben dem Lernerfolg, der abhängig vom Diagnoseinstrument unterschiedlich ausfallen kann, berücksichtigt das bedeutungsvolle Lernen zusätzlich die Relevanz der Eigenaktivität der Lernenden: Es setzt eine gut organisierte Wissensstruktur im relevanten Bereich voraus sowie die erwähnten Verknüpfungen zwischen neuen und bestehenden Wissensstrukturen (s. auch Driver, 1989). Auf diese Weise werden neue Konzepte oder Ideen in das Vorwissen integriert (Lawless et al., 1998; Novak, 1990b; Novak & Gowin, 1984). Diese Integrationsvoraussetzung unterstreicht sowohl die Bedeutung des Vorwissens für das Lernen als auch die Eigenaktivität der Lernenden, da die Integration bewusst und aktiv geschehen muss (s. dazu auch: DiCarlo, 2006). Dadurch können Lernende ein tiefes Verständnis der Konzepte erlangen und in einem bestimmten Kontext gelernte Konzepte auf neue Probleme anzuwenden (Novak, 2003). Aus diesem Grund besitzt das bedeutungsvolle Lernen in den Naturwissenschaften und insbesondere in der Biologie eine besondere Relevanz, da es in diesen Bereichen wichtig ist, Zusammenhänge zwischen Konzepten zu erkennen und zu verstehen sowie Wissen in neuen Kontexten anzuwenden (DiCarlo, 2006). Chevron (2014) verweist jedoch darauf, dass solches bedeutungsvolle Lernen schwierig zu erreichen ist, da es mehrere kognitive Schritte erfordert.

Eine entscheidende Rolle bei der Vermittlung und Förderung des bedeutungsvollen Lernens kommt daher der Lehrkraft zu: Sie entscheidet sowohl über Faktoren, wie die Gestaltung des Unterrichts und der Lernmaterialien, als auch über die Wahl der Instrumente, die zur Messung des Lernerfolgs eingesetzt werden. Im Unterricht kann die Lehrkraft bedeutungsvolles Lernen durch eine geeignete Präsentationsreihenfolge der Informationen und die eingesetzten Unterrichtsmethoden fördern und somit einen Beitrag zur Reduzierung der Belastung leisten, die durch ungeeignete Methoden entsteht (s. auch Kapitel 2.1.5.2). Die

adäquate Gestaltung gilt es auch bei der Erstellung von Lernmaterialien zu beachten, wobei explizite Hinweise im Material den Lernenden helfen können, Bezüge zu anderen Elementen des jeweiligen Wissensgebietes zu erkennen und somit ihr Verständnis fördern (Cañas & Novak, 2014; Novak, 1977). Als letzter, aber gleichfalls wichtiger Faktor gilt die Wahl der Bewertungs- oder Beurteilungsmethoden durch die Lehrkraft (s. Kapitel 1): Die Tendenz zur Abfrage reinen Faktenwissens dominiert in Schulen und Universitäten, indem z. B. Multiple-Choice-Tests eingesetzt werden (DiSessa, 2001; Novak, 2003). Derartige Wissenstests gelten zwar als objektiv (Suskie, 2018), allerdings werden durch das Format die Antworten der Lernenden eingeschränkt, weshalb nur wenige Rückschlüsse auf ihre Wissensstrukturen möglich sind (McClure et al., 1999; Watson et al., 2016). Eine solche Praxis hat zusätzlich zur Folge, dass Schüler:innen und Studierende häufig das reine Auswendiglernen wählen, da sie damit oft Erfolg haben (Mayer, 2002a; Montfort et al., 2009; Novak, 2003), obwohl Lernende, die Beziehungen zwischen Fakten bzw. einzelnen Wissens-elementen herstellen, sich erfolgreicher an Wissensinhalte erinnern als jene, die dies nicht tun (Gunstone, 1981; Wadouh et al., 2014). Zudem zeigen diverse Studien, dass Lernende mit dem Ziel, das Gelernte auch zu verstehen, viel besser abschneiden, wenn sie aufgefordert werden, ihr Wissen auf neue Probleme zu übertragen (Bascones & Novak, 1985; Novak et al., 1983).

Um die kognitiven Prozesse der Organisation und Elaboration zu fördern und damit bedeutungsvolles Lernen zu ermöglichen, können Lehrkräfte sowohl Lernstrategien vermitteln als auch Diagnoseinstrumente einsetzen, die von den üblichen Methoden wie Multiple-Choice oder richtig-falsch-Tests abweichen. Zu diesen Alternativen gehört das Concept Mapping, auf das später näher eingegangen wird (s. in Kapitel 2.2). Um die Vor- und Nachteile des Concept Mapping als Lernmethode im Rahmen der Lernstrategien einzuordnen, werden im Folgenden kognitive und metakognitive Strategien vorgestellt, die Lernende nutzen können, um entsprechende Prozesse zu initiieren. Basierend auf den vorgestellten Theorien und Modellen (s. Kapitel 2.1.2) wird davon ausgegangen, dass das prozedurale Wissen und das metakognitive Wissen für die Ausführung der jeweiligen Strategie bedeutsam sind, um einen gut geplanten, strukturierten und überwachten Lernprozess mit entsprechendem Lernerfolg zu ermöglichen.

2.1.4 Klassifizierung von Lernstrategien

In ihrer Schulzeit sollen Schüler:innen gemäß curricularer Vorgaben mit diversen Lernstrategien konfrontiert und in ihrem Einsatz derart geschult werden, dass sie die Strategien kompetent anwenden können (Artelt, 2006). Derartige Lernstrategien werden nach Weinstein und Mayer (1986) in drei Gruppen geordnet: kognitive, metakognitive und ressourcenbezogene Lernstrategien (s. auch Baumert, 1993; Pintrich et al., 1991; K. P. Wild & Schiefele, 1994). Der Einsatz kognitiver Lernstrategien dient der Erarbeitung, Strukturierung und Nutzung von Wissen. Zu den metakognitiven Strategien gehören z. B. das Planen und Überwachen des Lernens, wohingegen die Anstrengungsbereitschaft und die Nutzung von Materialien ressourcenbezogen sind (den Elzen-Rump et al., 2008; Mandl & Friedrich, 2006).

Des Weiteren können Lernstrategien, basierend auf dem Modell der Verarbeitungstiefe von Craik und Lockhart (1972), in Oberflächen- und Tiefenstrategien eingeteilt werden, mit deren Hilfe Unterschiede in der Behaltensleistung von Informationen erklärt werden sollen (Baumert & Köller, 1996; Marton & Säljö, 1997). Hierbei gilt grundsätzlich: Je „tiefer“ Lernende die Inhalte verarbeiten, desto länger werden sie sich daran erinnern. Zu den Oberflächenstrategien gehört bspw. das einfache Lesen eines Textes oder das Vokabellernen im Sinne des Auswendiglernens von isolierten Informationen (Marton & Säljö, 1997; Mayer, 2002a). Dem gegenüber stehen vertiefende Strategien wie z. B. die kritische Reflexion von Textaussagen, deren Umstrukturierung oder die Analyse von Argumentationen. Neues Wissen soll in bereits bestehende Wissensstrukturen, im Sinne der „Tiefe“, integriert werden. Dies wiederum soll die Lernenden dazu befähigen, das neue Wissen auch in anderen Kontexten anwenden zu können (Baumert, 1993; Friedrich & Mandl, 2006; Mayer, 2002a; K. P. Wild & Schiefele, 1994). Diese Unterteilung, sowohl der Verarbeitungsebenen als auch der entsprechenden Lernstrategien, ermöglicht es den Lernenden, unterschiedliche Lernstrategien kontextabhängig zu nutzen und Aspekte ihrer Lernprozesse zu beeinflussen. Auch wenn sich die Funktionalitäten dieser Strategiekategorien nicht eindeutig voneinander trennen lassen, werden innerhalb der sogenannten kognitiven Lernstrategien Organisations-, Elaborations- und Wiederholungsstrategien voneinander unterschieden, die z. B. zur Verarbeitung neuer Informationen oder zur Umgestaltung bestehender kognitiver Strukturen eingesetzt werden.

2.1.4.1 Kognitive Lernstrategien

Organisationsstrategien sollen entsprechende kognitive Prozesse anregen, durch die Lernende Informationen sinnvoll organisieren und Zusammenhänge zwischen Sinneinheiten erkennen. Indem Lernende z. B. Tabellen, Diagramme oder Concept Maps erstellen, werden relevante und verwandte Detailinformationen zu größeren Einheiten zusammengefasst und gruppiert, wodurch sie kognitiv leichter zu verarbeiten sind (Wolters et al., 2005). Zudem werden Informationen in eine bessere Struktur übertragen und Lernende werden beim Aufbau eines kohärenten Beziehungsnetzwerkes der gelernten Informationen unterstützt (Nückles et al., 2008). Durch die Organisation von Informationen und die systematische Verdichtung des Wissens kann zudem die Rekonstruktion von Wissen unterstützt werden, wenn eine Wiedergabe erforderlich ist (Novak & Cañas, 2008).

Zu den Elaborationsstrategien gehören die Aktivierung des eigenen Vorwissens, das Finden von Beispielen und Analogien für einen spezifischen Kontext, das Fragenstellen sowie das Zusammenzufassen des neu Gelernten in eigenen Worten. Mithilfe dieser Aktionen sollen Lernende auch bei der Arbeit mit komplexen Texten dabei unterstützt werden, ihre Aufmerksamkeit auf das Wesentliche zu lenken, ihr Wissen zu organisieren und neue Informationen in das Vorwissen zu integrieren (Isaak et al., 2020; Lumer & Winter, 2019). Eine derartige Aktivierung und Strukturierung des Wissens ist für das bedeutungsvolle Lernen wichtig, da sich die Bedeutung neuer Informationen für die bzw. den Lernende:n erst aus den Verbindungen mit dem Vorwissen ergibt (s. Kapitel 2.1.3). Auf diese Weise begünstigen Elaborationsstrategien ein tiefgründiges Verständnis über einen Sachverhalt (Kalyuga, 2009; K. P. Wild & Schiefele, 1994) und können zusätzlich dazu beitragen, das Lernen im naturwissenschaftlichen Unterricht durch die Förderung des vernetzten Denkens zu erleichtern (Wadouh, 2007).

Wiederholungsstrategien dienen i. d. R. dazu, Wissen zu festigen, zu behalten und im Langzeitgedächtnis zu verankern. Dies geschieht mit dem Lernmaterial (z. B. durch erneutes Lesen) oder ohne verfügbares Lernmaterial (z. B. beim Lernen mit Karteikarten). Da es sich hierbei jedoch um eine buchstäbliche *Wiederholung* handelt, d. h. Prozesse nach der eigentlichen Lernphase (Karpicke & Grimaldi, 2012), gelten Mnemotechniken (umgangssprachlich: *Eselsbrücken*; J. R. Anderson, 2013; Stangl, 2006) sowie Wiederholungsstrategien oft als Oberflächenstrategien. Nichtsdestotrotz sind sie für das Lernen im Sinne von *Behalten* und *Abrufen* (engl. *retrieval*) von Informationen wichtig, da Informationen aus dem Gedächtnis verdrängt werden können, wenn sie nicht aktiv erinnert werden (Wolters et al., 2005). Dadurch können Mnemotechniken und Wiederholungsstrategien einerseits

Auswirkungen auf Lernprozesse haben und andererseits einen Beitrag zur Elaboration und bedeutungsvollem Lernen leisten (Friedrich & Mandl, 2006; Karpicke & Grimaldi, 2012; Steiner, 2006; Sumfleth et al., 2010). Welche Rolle der gezielte Einsatz des Abrufs von Informationen aus dem Gedächtnis (*retrieval practice*) in Kombination mit Organisations- und Elaborationsstrategien wie dem Concept Mapping einnimmt und wie er in einigen Studien (z. B. Blunt & Karpicke, 2014; O'Day & Karpicke, 2020) implementiert wurde, wird ausführlicher in Kapitel 2.3.2 thematisiert.

2.1.4.2 Metakognitive Lernstrategien

Die bisher vorgestellten kognitiven Lernstrategien gelten auch als Primärstrategien (Dansereau, 1978), da sie den Lernfortschritt durch die Auseinandersetzung mit neuen Inhalten selbst unterstützen sollen. Für effektives Lernen sind jedoch auch metakognitive Prozesse und Strategien von Bedeutung (Bannert, 2007; Flavell, 1984; Hasselhorn & Artelt, 2018). Metakognitive Lernstrategien sollen die entsprechenden metakognitiven Prozesse der Planung, Überwachung, Bewertung bzw. Evaluation und die darauf basierte Regulation der Lernprozesse fördern (Gräsel, 1997; Stark et al., 2008). Wenn Lernende metakognitive Lernstrategien einsetzen, übernehmen sie Kontrollaufgaben, die traditionell von Lehrkräften übernommen werden, weshalb Artelt (2000) die Metakognition als zentrales Merkmal des selbstregulierten Lernens ansieht (s. auch Bannert, 2007). Im Idealfall versetzt der Einsatz dieser Strategien Lernende in die Lage, das eigene Lernverhalten an aktuelle Situationen und Aufgaben ohne externe Hilfe oder gar Kontrolle anzupassen und zu optimieren (Baumert, 1993; Friedrich & Mandl, 2006; Hasselhorn & Gold, 2017; K. P. Wild & Schiefele, 1994). Da sie eine übergeordnete Erfolgskontrolle der eigenen Lernschritte darstellen, können sie auch als Stützstrategien bezeichnet werden (vgl. Dansereau, 1978).

Der Einsatz metakognitiver Strategien hat sich in mehreren Studien als wichtiger Faktor für erfolgreiches Lernen erwiesen (P. Chen et al., 2017; Hattie, 2012; Höfer & Steffens, 2012; Lockl & Schneider, 2017). Der Erwerb metakognitiver Strategien ist jedoch in der Regel mühsam und selten beiläufig oder zufällig wie es reiner Wissenserwerb sein kann; zudem hat es sich gezeigt, dass Lernende metakognitive Strategien bei der Nutzung kognitiver Lernstrategien oft nur unzureichend bzw. ineffektiv einsetzen (z. B. Bonner & Holliday, 2006; Großschedl, 2010; Konrad, 2006). Forscher:innen wie Hasselhorn und Gold (2017) weisen deshalb darauf hin, dass der geübte Einsatz metakognitiver Strategien von mehreren Faktoren abhängt, wie der Art und Komplexität der Strategien und den instruktionalen Rahmenbedingungen einer Lernsituation. Derartige Rahmenbedingungen spielen auch bei den ressourcenbezogenen Lernstrategien, die im Folgenden vorgestellt werden, eine zentrale Rolle.

2.1.4.3 Ressourcenbezogene Lernstrategien

Das Ressourcenmanagement wird in ein internes und ein externes unterteilt und beschreibt die Interaktion Lernender mit ihrer Umwelt. Zu den internen Ressourcen werden z. B. die Anstrengungsbereitschaft und das Zeitmanagement gezählt, wohingegen die Gestaltung des Arbeitsplatzes und das Hilfesuchverhalten der Lernenden bei Problemen zu den externen Ressourcen gezählt werden (Baumert, 1993; Friedrich & Mandl, 2006; K. P. Wild & Schiefele, 1994). Die Bedeutung der Nutzung externer Ressourcen wird von Friedrich und Mandl (2006) im Zusammenhang mit der Strategie des kooperativen Lernens betont, da durch diese Form des Lernens ein zusätzlicher Einfluss auf die Motivation und Emotionen genommen wird.

Zusammenfassend kann festgehalten werden, dass das Spektrum an kognitiven, metakognitiven und ressourcenbezogenen Lernstrategien den Lernenden viele Möglichkeiten bietet, ihre Lerneffizienz zu steigern. Viele Lernende haben jedoch bis zum Studieneintritt an einer Universität für sich selbst das Spektrum an verfügbaren Strategien auf wenige reduziert (z. B. Streblov & Schiefele, 2006; K. P. Wild, 2005). Diese Reduktion der Lernstrategien hängt u. a. mit der Vertrautheit bzw. Erfahrung zusammen, die Lernende mit den Strategien haben und aus der sie den Nutzen der Strategien ableiten. Eine unbekannte Strategie, deren Erfolg ungewiss ist, wird von Lernenden nicht gewählt, wohingegen solche Strategien, mit denen sie vertraut sind und die sich in der Praxis für sie bewährt haben, häufig genutzt werden: So berichten zahlreiche Forscher:innen (z. B. Karpicke et al., 2009; Marton & Säljö, 1997; Streblov & Schiefele, 2006; Toppino et al., 2018; Toppino et al., 2021), dass Strategien wie das erneute Lesen viel genutzt, wohingegen aufwändigere Tiefenstrategien weniger genutzt werden. Dies deutet darauf hin, dass Lernende mit diesen „oberflächlicheren“ Strategien vertraut sind, sie als erfolgsversprechend bewertet werden und Lernende sie entsprechend einsetzen. In Konsequenz bedeutet dies, dass Lernende mit Tiefenstrategien entweder weniger vertraut sind oder sie als weniger nutzbringend ansehen. Da sich Tiefenstrategien jedoch generell als lernförderlich erwiesen haben (z. B. Mandl & Friedrich, 2006; Wadouh, 2007), ist davon auszugehen, dass die Lernenden unzureichend vertraut mit dem Strategieeinsatz sind. Friedrich und Mandl (2006) plädieren deshalb dafür, dass die Vertrautheit mit einer Lernstrategie ebenfalls zum Vorwissenstand gezählt und deren Einsatz trainiert werden sollte. Dafür sprechen auch Befunde, dass Lernende, die mit einer Strategie nicht vertraut sind, oft Probleme damit haben, die Strategie in ihr Repertoire zu integrieren (*Produktionsdefizit*), aus dem Strategieeinsatz Nutzen zu ziehen (*Nutzungsdefizit*; Bannert, 2003) oder das Potenzial der Strategie voll auszuschöpfen (z. B. Karpicke, 2009; Kornell & Son, 2009). Diese Probleme

resultieren in Lerneinbußen (z. B. Artelt, 1999; Baumert & Köller, 1996; K. P. Wild, 2005), weshalb Lernende im Vorhinein mit der spezifischen Lernstrategie vertraut gemacht werden sollten, um sie für eine erleichterte Informationsaufnahme und -verarbeitung effektiv einsetzen zu können.

2.1.5 Lernstrategietraining

Die Anwendung einer Lernstrategie erfordert prozedurales Wissen und die Vertrautheit mit der Strategieanwendung kann durch Fördermaßnahmen wie einem Training erlangt werden (Leopold & Leutner, 2002; Renkl, 2020; Sumfleth et al., 2010; s. auch Kapitel 2.1.1; 2.1.2; 4, 1.3). Bisherige Befunde (z. B. Friedrich & Mandl, 2006; K. P. Wild, 2005) deuten darauf hin, dass leistungsstarke Lernende sich im Ausmaß und der Qualität des Strategiegebrauchs von weniger starken Lernenden unterscheiden. Allerdings kann es auch bei normalerweise leistungsstarken Proband:innen zu Leistungseinbrüchen kommen, wenn der Vorwissensstand von Lernenden bei der Einführung einer neuen Lernstrategie durch die Lehrkraft nicht berücksichtigt wird. Solche Einbrüche sind häufig darauf zurückzuführen, dass die Leistungsstarken i. A. bereits über für sie erfolgreiche Lernstrategien verfügen, deren Nutzung und üblicherweise resultierenden Erfolg durch den Einsatz der neueingeführten Strategie unterbrochen wird. Eine erfahrungsgemäß gewinnbringende Strategie soll demnach durch eine neue ersetzt werden, deren Erfolgswahrscheinlichkeit ungewiss ist. In der konkreten Lernsituation müssen sie folglich nicht nur den zu lernenden Inhalt kognitiv verarbeiten, sondern zusätzlich die methodischen Schritte ausführen und überwachen, die mit der Ausführung der neuen Lernstrategie verbunden sind, wodurch ihr Lernen gestört wird. Dieser negative Effekt auf die Leistung wird als *expertise reversal effect* bezeichnet und verdeutlicht die notwendige Anpassung von Instruktionen und Materialien nicht nur an leistungsschwache, sondern auch leistungsstarke Personen (Friedrich, 1992; Kalyuga et al., 2003; Kalyuga, 2007; vgl. Kapitel 4, 3.5).

Für das Training müssen daher mehrere Faktoren an die jeweilige Lerngruppe angepasst werden, so dass Lernende einerseits durch den Strategieeinsatz kognitiv nicht unnötig belastet werden und andererseits die Strategie als für sie relevant und nützlich wahrnehmen (Friedrich & Mandl, 1992). Verfügen Lernende nicht über dieses Wissen, droht neben der zusätzlichen kognitiven Belastung ein Konflikt zwischen bekannten, funktionierenden Lernstrategien und den neuvermittelten, der ebenfalls zu Lerneinbußen führen kann (*mathemathantischer Effekt*; Clark, 1989; s. auch Blunt & Karpicke, 2014).

2.1.5.1 Indirekte vs. direkte Lernstrategietrainings

Lernstrategietrainings können in indirekte und direkte Trainings unterteilt werden. Bei der indirekten Förderung liegt der Fokus speziell auf den Lerninhalten und die Lernstrategie wird beiläufig trainiert. Ein solches Training kann z. B. eine Kurzeinführung darstellen, in der wichtige Informationen zur Lernstrategie in Form eines Lehrvortrags gegeben werden. Da von den Lernenden erwartet wird, dass sie nach der Einführung die Lernstrategie ein- und umsetzen, während sie gleichzeitig die Lerninhalte verarbeiten sollen, erfordern derartige Ansätze eine entsprechend hohe Selbstständigkeit.

Dagegen zeichnet sich ein direktes Strategietraining dadurch aus, dass die Lernstrategie im Fokus steht und die Lerninhalte von sekundärer Bedeutung sind. Zudem werden, im Gegensatz zur indirekten Förderung, sowohl Wirkungsweise als auch Vorteile der trainierten Lernstrategie explizit benannt. Dadurch sollen Lernende einerseits durch die kognitiv anspruchsvolle Lernstrategie nicht überfordert werden (*cognitive load theory* nach Sweller et al., 1998; Sweller, 2005; s. Kapitel 2.1.5.2) und andererseits sollen sie dazu motiviert werden, die trainierte Lernstrategie auch in neuen Anwendungsbereichen einzusetzen (Friedrich & Mandl, 1992). Damit Lernende die Vorteile der Lernstrategie selbst erfahren und sie nachhaltig einsetzen, wird ihnen im Rahmen ausführlicherer Trainings (extensiver Trainings, oft mit mehreren Sitzungen) nicht nur deklaratives Wissen zur Lernstrategie vermittelt, sondern das aktive und angeleitete Ausprobieren der Lernstrategie ermöglicht. Dadurch wird dem Umstand Rechnung getragen, dass die Umsetzung der für die Lernstrategie spezifischen Schritte prozedurales Wissen erfordert und insbesondere die Wiederholung von prozeduralem Wissen erhebliche Lerneffekte hat (vgl. Kap 4, 1.3): Wenn ein:e Schüler:in bspw. zuerst eine kleinschrittige Anleitung verwenden musste, um eine Proposition für die Concept Map zu erstellen, würde sie bzw. er durch wiederholtes Ausführen der Schritte die Aufgabe früher oder später auch ohne die Anleitung durchführen können, weil das Endergebnis als deklaratives Wissen abgerufen werden kann (Renkl, 2015). Dadurch können Routinen entstehen, die weniger kognitive Anstrengung erfordern und im Laufe der Zeit schneller durchgeführt werden können. Gleiches gilt für Fertigkeiten wie das Lesen, bei dem die Wiederholung der Handlung zu mehr Routine führt und der Abruf aus dem Langzeitgedächtnis mit sehr geringem Aufwand erfolgt, was als Automatismus bezeichnet wird (Moors & de Houwer, 2006), und eine Entlastung des Arbeitsgedächtnisses darstellt (Mietzel, 2017; s. auch Kapitel 2.1.5.2).

2.1.5.2 Lernhilfen und kognitive Belastung

Um den unterschiedlichen Anforderungen der Lernenden gerecht zu werden und sie beim Lernstrategieinsatz bestmöglich zu unterstützen, können Lernhilfen wie Scaffolds und Feedback in den Trainings eingesetzt werden. Bezüglich des Einsatzes solcher Maßnahmen verweisen Hasselhorn und Gold (2013) wiederum auf eine Abhängigkeit der Maßnahmen-Effektivität vom Vorwissen der Lernenden. Während solche mit mittlerem Vorwissensstand am meisten von den Lernhilfen profitieren, kann es bei solchen mit wenig und viel Vorwissen zu einer zusätzlichen Belastung durch die Lernhilfen kommen. Lernende mit wenig Vorwissen können die Lernhilfen nicht effektiv nutzen, da sie nicht wissen, wie sie diese einzusetzen haben, während Lernhilfen für Personen mit einem hohen Maß an Vorwissen eine zusätzliche Ablenkung darstellen (vgl. Kapitel 4, 1.2). Diese von den Ausgangsbedingungen der Lernenden abhängig auftretenden Effekte werden als Beispiel für die *Aptitude Treatment Interaction* (Cronbach, 1957, 1975; s. auch Kalyuga, 2007) und als Folge kognitiver Belastung gesehen, wobei letztere in der entsprechenden *cognitive load theory* (Debie & van de Leemput, 2014; Klepsch et al., 2017; Sweller et al., 1998; Sweller, 2005; vgl. Kapitel 4, 1.2) aufgegriffen wird. Im Rahmen von Mehrspeichermodellen des Gedächtnisses wird angenommen, dass ein sogenanntes Arbeitsgedächtnis die kritische Verarbeitungsstelle für Inhalte ist, die durch kognitive Prozesse vom Kurzzeitgedächtnis ins Langzeitgedächtnis übertragen werden sollen oder aus dem Langzeitgedächtnis aufgerufen (Baddeley, 2012; Novak & Cañas, 2008). Die *cognitive load theory* besagt, dass das Arbeitsgedächtnis eine limitierte Kapazität besitzt, weshalb die zur Verfügung stehenden Ressourcen beim Lernen optimal genutzt werden müssen. Bezüglich des Lernprozesses wird von drei Facetten des Arbeitsgedächtnisses ausgegangen, die die Gesamtkapazität des Arbeitsgedächtnisses bestimmen. Diese drei Facetten sind die intrinsische, lernbezogene und extrinsische kognitive Belastung (Klepsch et al., 2017): Den Lernenden sollte ermöglicht werden, ihre Ressourcen einerseits der Ausführung ihrer (meta)kognitiven Lernstrategien widmen zu können (*germane cognitive load*/lernbezogene kognitive Belastung) und andererseits den Anforderungen des eigentlichen Lerngegenstands (*intrinsic cognitive load*/intrinsische kognitive Belastung). Schließlich sollte das Lernmaterial und die Lernumgebung so gestaltet sein, dass durch diese (äußeren) Faktoren keine zusätzliche extrinsische kognitive Belastung (*extrinsic cognitive load*) entsteht. In diesem Fall können Lernende die begrenzten Ressourcen des Arbeitsgedächtnisses dem *germane* und *intrinsic cognitive load* zur Verfügung stellen. Andernfalls können Lernende durch die erhöhte extrinsische Belastung, die z. B. durch ungeeignetes Lernmaterial entsteht, zusätzlich beansprucht werden. Entsprechend werden Ressourcen, die den lerndienlichen Prozessen gewidmet werden

könnten, durch die extrinsische Belastung gebunden. Infolgedessen kann das Lösen der Aufgabe und/oder das Lernen selbst dadurch erschwert werden und die Wahrscheinlichkeit, dass die Lerninhalte behalten werden, verringert (Nückles et al., 2010).

- Scaffolding

Während des Lernens müssen mehrere Aktivitäten überwacht werden, was insbesondere bei der Neueinführung einer Lernstrategie hohe Anforderungen an Lernende stellt (Hmelo-Silver & Azevedo, 2006). Schnotz und Kürschner (2007) schlugen daher vor, Lernenden genügend Unterstützung zu geben, um eine Aufgabe zu bewältigen, die sonst schwer oder unmöglich wäre. Insbesondere in der kritischen Anfangsphase können Scaffolds in Form von stark vorstrukturierten Aufgaben oder Hilfestellungen eingesetzt werden, um Lernende dabei zu unterstützen, die Anforderungen der neuen Lernstrategie zu meistern (Bannert, 2003). Scaffolds haben sich in zahlreichen Studien als effektiv zur Förderung metakognitiver, kognitiver und affektiver Variablen herausgestellt, wie van de Pol et al. (2010) in ihrer Metaanalyse herausfanden. Solche Scaffolds können im weiteren Verlauf des Trainings mit zunehmender Sicherheit der Lernenden reduziert werden, wodurch den Lernenden ein selbstständigeres Arbeiten ermöglicht wird und ihre kognitive Belastung reduziert werden kann (Nückles et al., 2010). Ohne Hilfen wie Scaffolds fällt es Lernenden häufig schwer, eine Lernstrategie aktiv und gezielt in das persönliche Strategierepertoire zu integrieren oder den erzielbaren Nutzen aus der Strategie zu ziehen (Bannert, 2003; s. auch Kapitel 2.1.4). Entsprechendes wird berichtet bezüglich des metakognitiven Strategiewissens und metakognitiven Fertigkeiten zur exekutiven Regulation des eigenen Lernprozesses (Bannert, 2003; Hasselhorn & Gold, 2013; vgl. Kapitel 4, 1.3). Doch insbesondere die metakognitiven Prozesse der Lernenden sind einer der Hauptfaktoren, welche die Lernfähigkeit beeinflussen, wie Hmelo-Silver und Azevedo (2006) betonen.

Um metakognitive Planungs-, Überwachungs- und Evaluationsprozesse zu fördern und somit den Lernerfolg zu steigern, können metakognitive Promptingmaßnahmen eingesetzt werden (Hilbert & Renkl, 2008, 2009; Huang et al., 2015). Promptingmaßnahmen (oder kurz: Prompts) sollen Lernaktivitäten anregen, die Lernende zwar beherrschen, jedoch nicht adäquat anwenden können (Berthold et al., 2007). Der Begriff stammt aus dem Englischen und bedeutet in der Pädagogik so viel wie „Lernhilfe“ oder „Denkanstoß“, die Lernende beim Lösen von Aufgaben navigieren sollen und ihnen helfen, relevante Aspekte des Lösungsprozesses zu erkennen. Prompts haben sich als effektiv erwiesen, Lernaktivitäten zu fördern und oberflächliches Lernen zu verhindern (Pressley et al., 1992; Stark et al., 2008). Einige Studien zeigen jedoch, dass metakognitives Prompting keinen Einfluss auf metakognitive

Lernaktivitäten (Konrad, 2006) oder nicht zu signifikanten Verbesserungen der Gedächtnis- und Wissenstestleistungen führte (Bannert et al., 2009). In der Studie von Bannert et al. (2009) fühlten sich die Studienteilnehmer:innen durch die anspruchsvollen metakognitiven Aufgaben, die von den Prompts gefordert wurden, in ihrem Lernen eingeschränkt und gestört. Einen möglichen Grund liefern z. B. Davis (2003) sowie Chen und Bradshaw (2007), die berichten, dass Lernende in der Lage sein müssen, eine Promptingmaßnahme zu verstehen oder zu interpretieren. Andernfalls würden Lernende diese Hilfen entweder nicht adäquat umsetzen oder ihr keine Aufmerksamkeit schenken. Darüber hinaus kann es zu einem Konflikt zwischen dem Einsatz kognitiver und metakognitiver Strategien kommen, wenn es Lernenden an der Fähigkeit oder dem Vorwissen mangelt, von metakognitiver Unterstützung zu profitieren (Bruckermann et al., 2017). Diese Ergebnisse stehen daher im Zusammenhang mit den zuvor beschriebenen Befunden von Hasselhorn und Gold (2013), dass die Effektivität der Lernhilfen vom Vorwissen der Lernenden abhängt. Promptingmaßnahmen sollten daher mit Bedacht und immer unter Berücksichtigung des Vorwissens der Lernenden eingesetzt werden, um eine unnötige kognitive Belastung und daraus resultierende Lernerfolgseinbußen zu vermeiden (C.-H. Chen & Bradshaw, 2007; Mayer & Moreno, 2003; Neubrand et al., 2016; Shute & Towle, 2003).

- Feedback

Eine weitere Lernhilfe stellt das Feedback dar, welches ein entsprechendes Verständnis für die Aufgabe bei der bzw. dem Lernenden fördern soll. Feedback stellt die Bereitstellung von Informationen dar, die den Unterschied ausmachen im Hinblick auf das, was die Person bereits versteht und was sie missversteht (Hattie et al., 2014). Die Art des Feedbacks kann anhand der bereitgestellten Informationen in einfaches und elaboriertes Feedback unterteilt werden. Während einfaches Feedback Auskunft darüber gibt, ob eine Antwort richtig oder falsch war (*knowledge of results*, KOR) und u. U. noch die richtige Antwort vorgibt (*knowledge of correct results*, KCR), umfasst das elaborierte Feedback zusätzliche Hinweise und Erklärungen, die Lernenden dabei helfen sollen, die richtige Antwort auch zu verstehen (vgl. Jacobs, 2002; Lipowsky, 2015).

Die Entscheidung zu einfachem oder elaboriertem Feedback hängt von der Komplexität der Aufgabe und der Inhalte ab, da sich gezeigt hat, dass sowohl zu simple Rückmeldungen als auch zu elaborierte Erklärungen keine Vorteile darstellen bzw. sogar zu Lerneinbußen führen können (Mory, 2004). Ein weiterer Faktor, der in der Praxis berücksichtigt werden muss, ist die verfügbare Zeit: Das Analysieren der Antworten und Geben von Feedback ist sehr zeitaufwändig, weshalb in manchen Studien (z. B. Hilbert & Renkl, 2009) davon abgesehen

wurde. Nichtsdestotrotz weisen empirische Befunde darauf hin, dass Feedback beim Lernen die kognitive Belastung verringert (Moreno, 2004), das Lernen verbessert (Butler et al., 2007; Butler & Roediger, 2008; Lyster et al., 2013) und insbesondere Personen hilft, die wenig Vertrauen in ihre Antworten haben (Butler et al., 2008). Dadurch ist Feedback laut Hattie und Clarke (2019) der wohl kritischste und wirkungsvollste Aspekt des Lehrens und Lernens (s. auch Hattie, 2008).

Zusammenfassend kann festgehalten werden, dass der Einsatz der vorgestellten Strategien Lernenden dabei helfen kann, ihre Lernprozesse zu fördern und das Lernen zu erleichtern. Um Lernende mit der Strategie vertraut zu machen und den lernwirksamen Strategieeinsatz zu fördern, kann ein Training durchgeführt werden. In einem solchen Training können zusätzliche Lernhilfen wie Scaffolds und Feedback eingesetzt werden, um Lernende bei der Anwendung der Strategie zu unterstützen. Die Wahl der spezifischen Lernmethode sowie Entscheidungen bezüglich ihrer Vermittlung und des Einsatzes von Lernhilfen sind jedoch u. a. von den Voraussetzungen der Lernenden abhängig und entsprechend von Lehrenden zu beachten. Eine der Lernmethoden, mit deren Hilfe das bedeutungsvolle Lernen (s. Kapitel 2.1.3) gefördert und erfasst werden soll, aber deren Einsatz hohe kognitive Anforderungen an die Nutzer:innen stellt (Ajaja, 2011), ist das Concept Mapping.

2.2 Untersuchungsgegenstand Concept Mapping

Die Methode des Concept Mappings wurde ursprünglich von Novak und Gowin als Diagnoseinstrument entwickelt, um Veränderungen der Wissensstrukturen unter dem Einfluss von Unterricht zu detektieren, da mittels Concept Maps alle sechs Level der Bloom'schen (1956) Lernziel-Taxonomie visualisiert werden können (Dahnke et al., 1998; Novak, 1990a, 1990b; Novak & Gowin, 1984). Die Forschung und Entwicklung zum Concept Mapping im naturwissenschaftlichen Unterricht verknüpft Novak (1998) sowohl mit lernpsychologischen konstruktivistischen Theorien als auch Ausubels (1963, 2000) Assimilationstheorie zum bedeutungsvollen Lernen (s. Kapitel 2.1.3), da während der Concept Map-Erstellung aktive kognitive Prozesse vollzogen werden, wie z. B. die mentale Organisation und Integration neuen Wissens in bereits vorhandene Wissensstrukturen (s. auch Mayer, 2002a).

2.2.1 Anwendungsgebiete von Concept Maps

Seit ihrer Entwicklung wurden die Vorteile von Concept Maps auch als Lehr- und Lernstrategie erkannt: Heute werden Concept Maps auf vielen Bildungsebenen eingesetzt, in der Schule und an Universitäten, sowohl für die Lehre als auch die Forschung (Adesope & Nesbit, 2009; Dahnke et al., 1998; Fischler & Peuckert, 2000; Kinchin, 2014; Lawless et al., 1998; Wadouh, 2007). Die Einsatzbereiche erstrecken sich über Naturwissenschaften, wie

Physik (z. B. Hegarty-Hazel & Prosser, 1991a; Karakuyu, 2010; Martínez et al., 2013; Peuckert, 2000; Prosser, 1987), Geowissenschaften (z. B. Quinn et al., 2003), Chemie (z. B. BouJaoude & Attieh, 2008; Grüß-Niehaus, 2010; Jack, 2013; Markow & Lonning, 1998; Stracke, 2004) und Biologie (z. B. Bramwell-Lalor & Rainford, 2014; Hegarty-Hazel & Prosser, 1991b; Heinze-Fry & Novak, 1990; Jegede et al., 1990; Kinchin, 2000b; Kinchin et al., 2005; Schmid & Telaro, 1990), bis hin zu Sozial- und Wirtschaftswissenschaften (z. B. Chiou, 2008; Lawless et al., 1998). Als Lernmethode dient Concept Mapping hierbei der Lernförderung, um bspw. das Textverständnis zu fördern (Chang et al., 2002; Parker et al., 1996; s. auch Kinchin, 2000a; Lawless et al., 1998; Mandl & Fischer, 2000; Stracke, 2004). Darüber hinaus kann Concept Mapping von Lehrkräften zur Curriculumsentwicklung und Unterrichtsplanung genutzt werden (Haugwitz & Sandmann, 2009; Hilbert & Renkl, 2008). Im Folgenden werden die Einsatzbereiche des Concept Mappings näher erläutert, die für dieses Projekt relevant sind: zum einen als Lern-, zum anderen als Diagnoseinstrument⁵.

2.2.1.1 Concept Maps als Lerninstrument

Concept Maps können auf unterschiedliche Weise als Lerninstrument eingesetzt werden, wobei zwischen zwei Ansätzen unterschieden werden kann: der Konstruktion und der Rezeption (Schroeder et al., 2018). Die Konstruktion beschreibt die Erstellung von Concept Maps, wohingegen bei der Rezeption den Lernenden eine fertige Concept Map vorgelegt wird, die es zu analysieren und deren Inhalte es zu lernen gilt. Das Ziel solcher fertigen Concept Maps ist es, die Lernenden mit weniger Informationen zu konfrontieren als sie bspw. in einem Text vorhanden sind. Durch eine Reduktion der Lerninhalte und die Präsentationsform sind die Informationen leichter zu verarbeiten. Auf diese Weise wird die extrinsische kognitive Belastung reduziert und das Lernen erleichtert. Ein weiterer Vorteil des Lernens mit fertigen Concept Maps ist, dass der Prozess des Erstellens vermieden wird, der für unerfahrene Concept Mapper:innen überfordernd sein kann (Allen & Tanner, 2003; Chang et al., 2001; Hilbert, Nückles, & Matzel, 2008). Eine der schwierigsten Aufgaben beim Erstellen von Concept Maps stellt das Treffen von Entscheidungen dar, die das Priorisieren der Konzepte und ihren Beziehungen zueinander in der Concept Map beinhaltet (Chevron, 2014; Novak, 2010; Novak & Cañas, 2006). Im Detail fällt es Lernenden zudem schwer, diese Beziehungen in der Concept

⁵ Im Speziellen sind für das hier vorgestellte Projekt die Einsatzbereiche des Concept Mappings als Lern- und Diagnoseinstrument relevant, in denen Teilnehmer:innen Concept Maps mit Papier und Stift in Einzelarbeit vervollständigen bzw. erstellen. Die Erstellung von Concept Maps in Gruppen sowie mit diversen Softwares an PCs ist ebenfalls in vielen Studien implementiert worden, doch da diese Bedingungen nicht in diesem Projekt vorgesehen waren, wird auf einen Vergleich von analog und digital erstellten Concept Maps verzichtet. Für einen Überblick zu kollaborativem Concept Mapping siehe z. B. Adesope und Nesbit (2009); zum Einsatz von Computer-basierten Concept Maps siehe z. B. Erdogan (2009) sowie Ifenthaler und Hanewald (2014).

Map genau zu definieren (Bentley et al., 2011; Brüning & Saum, 2017), was in unpräzisen Pfeilbeschriftungen resultieren kann (Kharatmal & Nagarjuna, 2010) und ebenfalls ein Problem für die Person darstellt, die für die Auswertung der Concept Map zuständig ist (s. Kapitel 2.2.1.2). Basierend auf diesen Argumenten für fertige Concept Maps wurden unterschiedliche Ansätze entwickelt, um Lernende an die Konstruktion eigener Concept Maps heranzuführen und ihrer kognitiven Überforderung vorzubeugen und/oder das Wissen bzw. die Wissensstrukturen der Lernenden zu erfassen.

Für die hier vorgestellte Arbeit sind drei Ansätze von Bedeutung, die in diversen Studien eingesetzt wurden und sich in der dargebotenen Unterstützung und daraus resultierenden Freiheitsgraden unterscheiden (Cañas et al., 2012): (1) *skeleton maps*, (2) Begriffssets und (3) das freie Concept Mapping.

- (1) Beim Einsatz sogenannter *skeleton maps* wird ein hohes Maß an Struktur und somit wenig Freiheit vorgegeben. *Skeleton maps* stellen vorgefertigte Concept Maps dar, in denen fehlende Pfeile, Beschriftungen oder Konzepte ergänzt werden sollen (Chang et al., 2001). Lernenden soll auf diese Weise ermöglicht werden, ihre Aufmerksamkeit auf spezifische Beziehungen zwischen Konzepten zu richten, was in einer Reihe von Studien zu gesteigerten Lernerfolgen führte (z. B. Chang et al., 2002; Ruiz-Primo, Schultz, et al., 2001). Autor:innen wie Novak und Cañas (2008) empfehlen sie daher und Feigenspan und Rayder (2017) betonen, dass sowohl *skeleton maps* als auch selbst erstellte Concept Maps vernetztes Denken fördern können. Wenn das Ziel jedoch darin besteht, Unterschiede in den Wissensstrukturen oder Fehlvorstellungen der Lernenden zu bewerten, ist ein wenig strukturierter Ansatz besser geeignet als ein hoch strukturierter mit *skeleton maps* (Ruiz-Primo, Schultz, et al., 2001), weshalb Forscher:innen (z. B. Hilbert, Nückles, & Matzel, 2008) dafür plädieren, dass Lernende langfristig darin geschult werden sollten, ihre eigenen Concept Maps für das Lernen zu konstruieren. Nichtsdestotrotz befürworten Brandstädter et al. (2012) die Einschränkung der Freiheitsgrade und bevorzugen ein hochgradig gerichtetes Mapping gegenüber einem ungerichteten Mapping, da der gerichtete Ansatz verlässlichere Ergebnisse ermöglicht. Als zusätzliche Problemvarianten schlagen Sie z. B. die Begrenzung der verfügbaren Pfeile oder die Vorgabe einer vorbereiteten Concept Map vor.

- (2) Einen ähnlichen Ansatz stellt die Bereitstellung von Begriffssets dar, mit denen Lernende ihre Concept Maps erstellen. Den Lernenden wird mehr Freiheit als in *skeleton maps* gewährt und sie haben die Möglichkeit, Concept Maps insofern selbst zu erstellen, da sie über die Anordnung und Verbindungen der Konzepte frei entscheiden können, aber nicht befürchten müssen, dass sie falsche Konzepte verwenden. Dieser Ansatz stellt ebenfalls eine allgemein anerkannte Methode dar (z. B. Cañas et al., 2003; Hauser et al., 2006; Novak & Cañas, 2008), allerdings ist die Arbeit mit vorgegebenen Konzepten weniger inhaltsreich und offen als die freie Konstruktion einer Concept Map (Ruiz-Primo, Shavelson, et al., 2001).
- (3) Ein solches Maximum an Freiheit und Minimum an Struktur wird den Lernenden beim freien Concept Mapping gegeben. Hier entscheiden die Ersteller:innen der Concept Maps selbst darüber, welche Begriffe sie nutzen und wie sie diese anordnen und vernetzen. Die Wirksamkeit von selbst erstellten Concept Maps und ihr positiver Einfluss auf den Lernerfolg wurden in zahlreichen Studien und Metaanalysen nachgewiesen (z. B. Horton et al., 1993; Jegede et al., 1990; Nesbit & Adesope, 2006; Roessger et al., 2018; Schroeder et al., 2018), auch wenn der Einfluss weiterer Faktoren wie verbale und kognitive Fähigkeiten sowie Vorwissen diskutiert wird (z. B. Ajaja, 2013; Hay et al., 2008; Nesbit & Adesope, 2006; Schmid & Telaro, 1990). Da die Syntax in Concept Maps auf eine Substantiv-Verb-Substantiv-Verbindung in den Propositionen reduziert wird, die einfacher und für schwache Leser:innen und Schreiber:innen zugänglicher ist als die typische Syntax in der Prosa, gehen manche Forscher:innen (z. B. Cañas et al., 2003; Haugwitz et al., 2010; Karakuyu, 2010; Nesbit & Adesope, 2013; O'Donnell et al., 2002; Parker et al., 1996) davon aus, dass Lernende mit weniger Vorkenntnissen und/oder sprachlichen Kenntnissen besonders vom Einsatz der Methode profitieren. Andere Forscher:innen dagegen verweisen auf ihre Befunde, dass Lernende mit einem vergleichsweise hohen Maß sprachlicher Fähigkeiten am meisten vom Concept Mapping profitierten (Chularut & DeBacker, 2004; Maker et al., 2015). Ajaja (2013) gibt zu bedenken, dass Concept Mapping hohe kognitive Fähigkeiten und eine sehr gute Beherrschung des Themas erfordert, so dass Lernende mit geringen kognitiven Fähigkeiten das Potenzial des Concept Mapping nicht voll ausschöpfen können. Diese Behauptungen stehen jedoch im Gegensatz zu den Ergebnissen von Schmid und Telaro (1990), die zeigen, dass Concept Mapping sowohl Lernenden mit geringen als auch mit hohen kognitiven Fähigkeiten hilft. Abgesehen von den verbalen und kognitiven Fähigkeiten verweisen auch Hay et al. (2008) in Bezug auf die von Ajaja (2013)

erwähnte Bedeutung des Vorwissens für den Lernerfolg. Sie gehen davon aus, dass das Vorwissen einen der wichtigsten Lernerfolgskriterien darstellt (s. auch Ausubel et al., 1978). Die Qualität des Lernens wird somit zu einem großen Teil durch die Ausgangsbedingungen, d. h. das Vorwissen der Lernenden, bestimmt⁶. Folglich sind Lernende mit einem guten Verständnis der Inhalte besser in der Lage, die Lerninhalte sinnvoll einzuordnen, selbst wenn sie eine unvertraute Lernstrategie wie das Concept Mapping einsetzen, was eine Erfassung und Berücksichtigung ihres Vorwissens unabdingbar macht (vgl. Kapitel 4).

Im Fokus des hier vorgestellten Projektes steht die Konstruktion von Concept Maps, da sie sich sowohl in bisherigen Studien als lernwirksamer als die Rezeption herausgestellt hat (Schroeder et al., 2018) als auch mehr Aufschluss über kognitive Prozesse und Lernergebnisse zulässt (Hilbert, Nückles, & Matzel, 2008; Ruiz-Primo, Schultz, et al., 2001). Aufgrund der Tatsache, dass Concept Mapping, wie oben erwähnt, hohe kognitive Ansprüche an die Ersteller:innen stellt, wurden für das Üben der Konstruktion unterschiedlich intensive, direkte Förderansätze (vgl. Kapitel 2.4 Lernstrategietraining; Kapitel 5 und 6) für das Concept Mapping genutzt, in denen z. T. *skeleton maps*, Begriffssets und Aufgaben zum freien Erstellen verwendet wurden, um eine schrittweise Heranführung zu gewährleisten, die der kognitiven Überforderung entgegenwirken sollte. Anhand einer kriteriengeleiteten Zusammenfassung werden einige der in früheren Studien verwendeten unterschiedlichen Trainingsansätze in Kapitel 2.3.1 erläutert. Vorher sollen die Verwendung von Concept Mapping als Diagnoseinstrument (Kapitel 2.2.1.2) und die Hintergründe der für das Concept Mapping reklamierten und empirisch belegten lernförderlichen Effekte (Kapitel 2.2.1.3; 2.2.2) vorgestellt werden.

2.2.1.2 Concept Maps als Diagnoseinstrument

Als Diagnoseinstrument eignen sich Concept Maps, um Wissen, Verständnis und Lernzuwachs, aber auch Fehlkonzepte und Missverständnisse der Lernenden zu ermitteln (Gouli et al., 2003; Novak, 2003; Romero et al., 2017; Ruiz-Primo & Shavelson, 1997; Watson et al., 2016), weil sie viele Informationen über die Wissensstrukturen von Lernenden liefern: Die Anzahl der Konzepte und ihrer Beziehungen kann als Indikator für die Qualität des Verständnisses der Lernenden gesehen werden, während die Art und Weise, wie die Concept Map organisiert ist, Aufschluss darüber gibt, wie das Verständnis strukturiert ist, wobei Tiefe,

⁶ Dies kann auch im Zusammenhang mit den vorgestellten Befunden zur Bedeutung von bestehenden Wissensstrukturen für die Elaboration gesehen werden (s. Kapitel 2.1).

Breite und Vernetzung anhand der Concept Map erkennbar werden (Novak & Gowin, 1984; Pruett & Weigel, 2020; Turns et al., 2000).

Durch die Menge an Informationen sind entsprechend große Herausforderungen bei der Auswertung von Concept Maps zu bewältigen (Allen & Tanner, 2003; Liu & Hinchey, 1996; Ruiz-Primo & Shavelson, 1996; Watson et al., 2016). Trotzdem betonten schon Ruiz-Primo und Shavelson (1997), dass Concept Maps trotz ihrer Komplexität reliabel bewertet werden können. Bei der Bewertung von Concept Maps wird hierbei häufig ein quantitativer Ansatz bevorzugt (Kinchin et al., 2000), der auf dem von Novak und Gowin (1984) entwickelten Scoring-Protokoll basiert. In der Vergangenheit wurden viele unterschiedliche Evaluationsmethoden mit unterschiedlichen Schwerpunkten entworfen (Besterfield-Sacre et al., 2004), zu denen es laut Dowd et al. (2015) keinen Konsens darüber gibt, welcher der beste Ansatz sei. Zu den drei Ansatzkategorien gehören:

- (1) Bei der Auswertung von Concept Maps wird ein Fokus auf Komponenten wie Propositionen und Hierarchieebenen der Concept Map gelegt (z. B. mit einem Punktesystem; Bayram, 1995; McClure et al., 1999).
- (2) Die Concept Map der bzw. des Lernenden wird mit einer anderen verglichen (z. B. mit einer Expert:innen-Concept-Map; Hegarty-Hazel & Prosser, 1991a, 1991b; Ruiz-Primo & Shavelson, 1996; Rye & Rubba, 2002; Schreiber & Abegg, 1991).
- (3) Es wird eine Kombination beider Strategien genutzt. Dieser Ansatz wird z. B. von Novak und Gowin (1984) vorgeschlagen.

Bei einem Fokus auf den Komponenten einer Map soll i. A. die Qualität der Concept Map quantifiziert werden, indem ein Punktesystem genutzt wird. Bayram (1995) nutzte ein solches System, das auf den hierarchischen Ebenen der Karte, korrekten Propositionen und Verzweigungen basiert. Ein Punkt wird für jede hierarchische Ebene, jede Beziehung zwischen Konzepten, die durch eine korrekte Proposition hergestellt wurde, und jede Querverbindung vergeben. Außerdem wird ein Punkt für jede Verzweigung der ersten Ebene vergeben, zwei Punkte für jede Verzweigung der zweiten Ebene, drei Punkte für jede Verzweigung der dritten Ebene und so weiter. Die Punktzahl einer Concept Map ergibt sich dann aus der Summe dieser Punkte. Die Berechnung der Gesamtpunktzahl aus den Einzelwerten wurde auch von McClure et al. (1999) gewählt. Allerdings konzentrierten sie sich auf die Qualität der Propositionen und entwickelten eines der einfachsten und häufig genutzten Bewertungsschemata, in dem jeder Proposition ein Wert von 0 bis 3 Punkten zugeordnet wird: Wenn keine Beziehung zwischen

den Konzepten besteht, erhält die Proposition einen Wert von 0; ein Punkt wird vergeben, wenn es eine Beziehung zwischen den Konzepten gibt, aber die Pfeilbeschriftung keinen Sinn ergibt; zwei Punkte erhält die Proposition, wenn die Pfeilbeschriftung Sinn ergibt, aber die Pfeilrichtung nicht, und drei Punkte, wenn die Proposition korrekt und sinnvoll ist (vgl. Kapitel 4, 2.3; 5, 2.3).

Im zweiten Ansatz wird davon ausgegangen, dass es für die Struktur eines Inhaltsbereichs eine ideale Organisation gibt, die das Thema am besten widerspiegelt (Ruiz-Primo & Shavelson, 1996). Daher wird z. B. eine Expert:innen-Map als „Musterlösung“ genutzt, um die Wissensstände der Expert:innen mit denen der Lernenden zu vergleichen (z. B. Hegarty-Hazel & Prosser, 1991a, 1991b; Rye & Rubba, 2002; Schreiber & Abegg, 1991). Eine Expert:innen-Map entsteht entweder ohne das Lernmaterial, indem die Urteile zu einem fachlichen Sachverhalt von mehreren Expert:innen aus dem spezifischen Bereich zusammengefasst werden und das Endergebnis eine Referenz-Struktur bietet (z. B. Hegarty-Hazel & Prosser, 1991a, 1991b; Rye & Rubba, 2002) oder mit dem Lernmaterial, wenn Expert:innen eine Concept Map z. B. mit einem Text erstellen, der auch den Proband:innen vorgelegt wird (z. B. Stasz et al., 1976; s. auch Shavelson, 1972).

Andere Forscher:innen (z. B. Novak & Gowin, 1984) plädieren für einen kombinierten Ansatz, da sie der Meinung sind, dass bei der Bewertung einer Concept Map sowohl auf die Qualität und Quantität der Concept Map-Komponenten als auch die Qualität im Vergleich zu einer Expert:innenlösung geachtet werden sollte. Novak und Gowin (1984) nutzten zusätzlich zu ihrem Punktesystem das Vergleichs-Bewertungskriterium. Die Expert:innen-Map wird ebenso wie die Concept Map der bzw. des Lernenden nach dem Punktesystem bewertet und dann die Punktzahl der Concept Map der bzw. des Lernenden durch die Punktzahl der Expert:innen-Map dividiert, um einen Prozentsatz für den Vergleich zu erhalten. Dadurch ist es möglich, dass Lernende besser abschneiden können als die Expert:innen und auf dieser Bewertungsgrundlage mehr als 100 % erhalten.

Die Bewertungsschemata sind jedoch immer abhängig von der Qualität der Propositionen und dadurch von den Concept Mapping-Fertigkeiten der Lernenden: Wenn eine Beziehung nicht klar definiert wurde, kann u. U. nicht erkannt werden, was die Proposition aussagen soll und ob dies auf einen Wissensmangel oder fehlende Concept Mapping-Fertigkeiten zurückzuführen ist. Neben der Ebene der einzelnen Propositionen berichtet Chevron (2014) von Problemen, die sich auf die gesamte Concept Map beziehen: Lernende bemühen sich zwar, den zu erforschenden Bereich abzugrenzen, neigen aber häufig dazu, von dem inhaltlichen Fokus, der in der Map dargestellt werden soll, abzuweichen und Maps zu konstruieren, die sich auf

einen viel größeren Wissensbereich beziehen als von der Lehrkraft beabsichtigt. Dies führt wiederum zu einem erhöhten Arbeitsaufwand für die Lehrkraft. Wenn also bei Lernenden diese Probleme auftreten und sie in ihren Concept Maps nicht das darstellen können, was sie wissen, oder unnötig viele Informationen in Propositionen übertragen werden, leidet die Verlässlichkeit der Schlüsse, die anhand der Concept Maps gezogen werden können. Daher sind die Concept Mapping-Fertigkeiten der Lernenden essenziell für die Reliabilität und Validität der Interpretation von Concept Maps (s. auch Brandstädter et al., 2012; Ruiz-Primo, Schultz, et al., 2001). Dies spricht einerseits für den Einsatz hochstrukturierter Concept Mapping-Aufgaben, wie sie von Brandstädter et al. (2012) vorgeschlagen werden, und andererseits für den Vorschlag von Ruiz-Primo und Shavelson (1997), Concept Maps nicht als einzige Informationsquelle über das Wissen eines Lernenden zu nutzen, sondern als eine der möglichen Quellen. Doch während diese beiden Ansätze Concept Mapping als Diagnoseinstrument vorsehen, aber davon ausgehen, dass Lernende Probleme bei der korrekten Erstellung ihrer Concept Maps haben, könnte ein präventiver Ansatz darin bestehen, den Lernenden das Concept Mapping so zu vermitteln, dass sie imstande sind, ihr Wissen in einer Concept Map darzustellen (s. auch Hilbert, Nückles, & Matzel, 2008). Dies würde eine Reduktion der Freiheitsgrade, die Lernende beim Concept Mapping einschränken, obsolet machen und gleichzeitig mehr Erkenntnisse zu den Wissensstrukturen der Ersteller:innen ermöglichen (Ruiz-Primo, Shavelson, et al., 2001).

Um Concept Maps als verlässliches Diagnoseinstrument einzusetzen ist es daher essenziell, dass das Wissen der Lernenden tatsächlich in der Concept Map repräsentiert wurde, was wiederum durch die Vertrautheit mit der Strategie als solcher und der Erstellung von Propositionen im Speziellen bedingt wird (s. auch Brandstädter et al., 2012). Wie eine solche Vermittlung des Concept Mapping in bisherigen Studien implementiert wurde und welche Vor- und Nachteile mit den unterschiedlichen Ansätzen verbunden sind, wird in Kapitel 2.3.1 dargelegt.

2.2.1.3 Concept Mapping als metakognitives Werkzeug

Viele Lernende tendieren dazu, ihre Lernfortschritte und ihr Wissen zu überschätzen oder ihrer Lerndefizite nicht bewusst zu sein (Dunlosky & Rawson, 2012; Ellis et al., 2004; Isaacson & Fujita, 2006). Dies führt zu einer Illusion des Wissens (*illusion of knowing*; Bjork, 1999; Eshuis et al., 2021; Isaacson & Fujita, 2006) und wird zum Problem, wenn Lernende sich selbst nicht korrekt einschätzen und dementsprechend darauf verzichten, ihre Wissenslücken zu schließen. Es bedarf daher einer Externalisierung des Wissens, so dass Lernenden ihre Wissenslücken aufgezeigt werden (Eshuis et al., 2021). In diesem Zusammenhang hat sich

gezeigt, dass Concept Mapping Lernenden auf der metakognitiven Ebene dabei hilft, ihr Wissen zu visualisieren und in diesem Zuge eigene Wissenslücken oder Verständnisschwierigkeiten zu erkennen sowie Vorstellungen zu reflektieren (Chularut & DeBacker, 2004; Friedrich & Mandl, 2006; Grüß-Niehaus, 2010; Hilbert & Renkl, 2008; Hundertmark, 2012; Lumer & Winter, 2019; Nesbit & Adesope, 2006; Novak, 1990a; Novak & Cañas, 2008; Roth & Roychoudhury, 1993; s. auch Kapitel 4, 1.1). Dies soll im Zuge der Erstellung der Concept Map geschehen, wenn Lernende erkennen, dass sie z. B. nicht die elementaren Konzepte aus dem Text extrahieren oder sie sinnvoll miteinander verbinden können. Bemerken Lernende, z. B. bei der Arbeit mit einem Text, Differenzen zwischen Textinhalten und ihrer Concept Map oder fehlende Konzepte, die integriert werden müssten, „zeugt dies von einer guten metakognitiven Kontrolle, die dem Verstehen dient“ (Lumer & Winter, 2019, S. 66). Lernende können im Anschluss diese Erkenntnisse nutzen, um ihre Wissenslücken zu schließen und ihren Lernerfolg zu steigern (Hilbert & Renkl, 2008; s. Kapitel 2.1.4).

Die Förderung metakognitiver Prozesse hat dazu geführt, dass Concept Mapping von Vielen als effektives metakognitives Werkzeug gesehen wird und die metakognitive Förderung als ein Hauptargument für den Einsatz dieser Lernmethode gilt (Bramwell-Lalor & Rainford, 2014; Kinchin, 2011; Mintzes et al., 1997; Ritchhart et al., 2009; Salmon & Kelly, 2014). Doch auch wenn Concept Mapping als effektives metakognitives Werkzeug gesehen wird und metakognitive Prozesse wie das Überwachen und die Evaluation eine Schlüsselrolle beim selbstregulierten Lernen einnehmen, da sie die Grundlage für weitere Kontrollprozesse wie Bewertung, Planung und Anpassung des eigenen Handelns bilden (Jersakova et al., 2017; Knoll et al., 2017; Metcalfe, 2009; Stark et al., 2008), wurde der explizite Einfluss des Concept Mapping auf die metakognitiven Prozesse des Überwachens und der Evaluation bisher nur in wenigen Studien untersucht. Ausnahmen bilden die Studien von Karpicke und Blunt (2011b), Blunt und Karpicke (2014) sowie O’Day und Karpicke (2020), in denen der Einfluss von unterschiedlichen Lernstrategien und der Verfügbarkeit des Lernmaterials (*wiederholtes Lesen des Lernmaterials vs. Concept Mapping mit Lernmaterial vs. free recall ohne Lernmaterial vs. Kombination aus free recall und Concept Mapping*) auf das metakognitive Bewusstsein im Sinne der Evaluation untersucht wurde. Die Forscher:innen baten die Proband:innen nach der Lernphase zu beurteilen, an wie viel Prozent des Lerninhalts (0 bis 100 %) sie sich eine Woche später noch erinnern würden. Eine solche metakognitive Vorhersage wird als *judgment of learning* (JOL) bezeichnet. Im Allgemeinen können diese Urteile die Grundlage für zukünftige Lernsituationen bilden (Nelson & Dunlosky, 1991). Je nachdem, wie erfolgreich die aktuelle Sitzung war, muss Zeit für die Wiederholung des Materials eingeplant werden, weshalb eine

akkurate Selbsteinschätzung den Lernenden helfen kann, ihr Zeitmanagement und damit das Lernen selbst zu optimieren (Knoll et al., 2017; Metcalfe, 2009; Thiede, 1999; Thiede et al., 2003). Folglich ist die Genauigkeit eines JOL entscheidend für die Lerneffizienz, da jede Diskrepanz zwischen einem JOL und der tatsächlichen Lernleistung zu einem weniger effizienten Lernen führen kann (Nelson & Dunlosky, 1991; Tauber et al., 2013; Yang et al., 2017; Yang et al., 2018).

Allerdings zeigen fast alle Studien, die JOLs untersucht haben, dass die Urteile der Lernenden ungenau sind (z. B. Nelson & Dunlosky, 1991), was auf übermäßiges oder mangelndes Vertrauen zurückzuführen ist (Hanczakowski et al., 2013). Übermäßiges Vertrauen beschreibt das Phänomen, wenn das JOL höher ist als die tatsächliche Lernleistung, während mangelndes Vertrauen ein JOL beschreibt, das niedriger ist als die Lernleistung (Cauvin et al., 2019). Derartige Diskrepanzen zeigen sich auch in den Studien zum Concept Mapping von Karpicke und Blunt (2011b), Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) sowie in Studien ohne Concept Mapping-Bezug (z. B. Karpicke, 2012; Roediger & Karpicke, 2006b). Allerdings deuten Vergleiche zwischen JOLs und Lernergebnissen darauf hin, dass die Verfügbarkeit des Lernmaterials einen Einfluss auf die JOLs hat: Die Ergebnisse von Karpicke und Blunt (2011b) sowie Blunt und Karpicke (2014) zeigen, dass Lernende, denen das Lernmaterial zur Verfügung steht (*Elaborationssetting*), höhere JOLs geben als die Gruppen, die während der Lernphase aufgefordert worden werden, nach dem Lesen des Textes ihre Aufgabe (Notizen erstellen im Sinne des *free recall*) ohne das Lernmaterial zu lösen (*Retrievalsetting*). In den anschließenden Tests zeigte sich in den beiden Studien jedoch, dass die Gruppen im Retrievalsetting besser abschnitten als jene im Elaborationssetting.

Als Grund für die falschen Prognosen bezüglich der eigenen Lernleistung bei verfügbarem Lernmaterial wird eine falsche Urteilsbasis der JOLs diskutiert (vgl. Kapitel 5, 1.2): Wenn Lernende das Material zur Verfügung haben, wie es beim wiederholten Lesen der Fall ist, verläuft die Bearbeitung des Inhalts reibungslos und leicht. In solchen Fällen sind die JOLs der Lernenden nach der Lernphase unrealistisch hoch im Vergleich zu ihren tatsächlichen Leistungen in späteren Testsituationen. Im Gegensatz dazu verändert das aktive Abrufen der Inhalte aus dem Gedächtnis die Informationsbasis, die die Lernenden zur Beurteilung ihres Lernens verwenden. Anstatt zu beurteilen, wie leicht die Informationen des Textes gelesen und/oder verstanden wurden, basiert die Beurteilung nun auf der Leichtigkeit oder Schwierigkeit, mit der die Informationen abgerufen werden können, was in früheren Studien zu genaueren Urteilen geführt hat (z. B. Kelemen, 2000; King et al., 1980; Koriat, 1997). So gaben King et al. (1980) den Teilnehmer:innen ihrer Versuchsgruppen die Aufgabe,

ihr Wissen nach einer Lernphase zu testen. Dabei stellten sie fest, dass die Teilnehmer:innen genauere Urteile abgaben, wenn sie vor dem JOL ihr Wissen getestet hatten. Die Genauigkeit von JOLs kann also verbessert werden, indem man den Lernenden die Möglichkeit gibt, ihre tatsächlichen Lernleistungen zu überprüfen. Diese Abrufversuche können Informationen über die Unterscheidung zwischen bekanntem und unbekanntem Material liefern und zu genaueren JOLs führen. Basierend auf den Berichten von Blunt und Karpicke (2014) sowie Karpicke und Blunt (2011b) ist daher anzunehmen, dass auch beim Concept Mapping die JOL-Genauigkeit von der Verfügbarkeit des Lernmaterials beeinflusst wird. Ob die JOL-Genauigkeit auch durch ein vor der Lernphase erhaltenes Concept Mapping-Training beeinflusst wird, ist noch nicht geklärt. Es ist jedoch möglich, dass die metakognitiven Vorteile u. a. von der adäquaten Einbettung metakognitiver Aufgaben schon während des Unterrichts und einem längeren Training abhängen: Diesbezüglich merkt Chevron (2014) an, dass eine solche Einbettung v. a. am Anfang wichtig sei, da solche Prozesse für Lernende mit einem Mehraufwand verbunden sind, später aber zu einer reibungslosen und dauerhaften Anwendung metakognitiver Aktivitäten führen.

2.2.2 Wissenserwerb mithilfe von Concept Mapping

Neben der Förderung metakognitiver Prozesse ist die Förderung des bedeutungsvollen Lernens ein weiteres Hauptargument für Concept Mapping, wie auch in Metaanalysen berichtet wird (Nesbit & Adesope, 2006; Schroeder et al., 2018), auch wenn noch unklar ist, welche spezifischen Concept Mapping-Charakteristika für die elaborationsfördernden Effekte verantwortlich sind (Schroeder et al., 2018). Gründe für die lernunterstützenden Effekte werden in der Kognitionspsychologie zum einen in den Parallelen des Concept Mappings zu theoretischen Modellen der Gedächtnisorganisation (Collins & Quillian, 1969; Quillian, 1968; s. Kapitel 2.1.1) und zum anderen im Rahmen der Theorie der Verarbeitungstiefe (Craik & Lockhart, 1972; s. Kapitel 2.1.4) gesehen.

Bezüglich der Gedächtnisorganisation erfordert die Erstellung von Concept Maps von den Lernenden, dass sie die zentralen Begriffe bzw. Konzepte des Themas und deren Beziehungen zueinander (z. B. aus einem Text) extrahieren und anschließend diese semantischen Beziehungen graphisch in der Concept Map darstellen (Hilbert, Nückles, Renkl, et al., 2008; Nesbit & Adesope, 2013). Concept Maps sollen daher sowohl die Inhalte selbst als auch die diesbezügliche Wissensorganisation und kognitiven Strukturen einer Person sichtbar machen, weshalb sie als Analogie zu semantischen Netzwerkmodellen des Gedächtnisses gelten (den Elzen-Rump & Leutner, 2007; Jüngst, 1995; Jüngst & Strittmatter, 1995; Novak & Cañas, 2008; Stracke, 2004; vgl. Kapitel 4, 1.1. und 5, 1.1). Des Weiteren wird davon

ausgegangen, dass mithilfe von Concept Maps durch die simultane Repräsentation von Inhalten in verbaler und piktoraler Form (*Doppelkodierungstheorie*; Paivio, 1986) neue Informationen leichter im Gedächtnis verarbeitet werden können. Im Vergleich zu einem reinen Text können Concept Maps dabei helfen, die kognitive Belastung auf die visuellen und verbalen Kanäle des Arbeitsgedächtnisses zu verteilen und somit eine Überlastung des verbalen Arbeitsgedächtnisses zu vermeiden (s. Kapitel 2.1.5.2); diese duale Kodierung soll zudem den Abruf dieser Informationen erleichtern (Nesbit & Adesope, 2013).

Die Entlastung des Arbeitsgedächtnisses wird bei der Arbeit mit Texten ebenfalls dadurch begünstigt, dass die Concept Map-gebundenen Informationen nicht ständig im Arbeitsgedächtnis präsent gehalten werden müssen, da eine Externalisierung (Cox, 1999) der angenommenen kognitiven Prozesse sowie Repräsentationsstrukturen durch das graphische Festhalten in der Concept Map stattfindet (Baddeley, 2012; Hasselhorn & Gold, 2017; Novak & Cañas, 2008). Dies soll Lernenden zum einen dabei helfen, kohärente mentale Wissensstrukturen aufzubauen; zum anderen werden die Textinformationen auf eine verarbeitungsfreundliche Art präsentiert, so dass Lernende zum einen die Verbindungen zwischen Konzepten des Textes besser identifizieren können und zum anderen wichtige Information schnell ablesen und abrufen können (Hilbert, Nückles, & Matzel, 2008; Lumer & Winter, 2019; Nesbit & Adesope, 2006; Weinstein & Mayer, 1986; vgl. Kapitel 4, 1.1).

Bezogen auf die Verarbeitungstiefe erfordern die methodischen Schritte des Concept Mappings insofern eine vertiefte Analyse des Lernmaterials, da eine Reduktion der Textinformationen auf das Wesentliche sowie eine Umstrukturierung der Inhalte gemäß der semantischen Nähe der Konzepte stattfindet, was eine Zuordnung des Concept Mappings zu den Tiefenstrategien (Craik & Lockhart, 1972) erklärt, aber mit einer entsprechenden kognitiven Anstrengung verbunden ist (Novak & Gowin, 1984). Der Einsatz solcher Strategien führt jedoch in der Regel zu einer verbesserten Behaltensleistung bei den Lernenden (den Elzen-Rump & Leutner, 2007; O'Donnell et al., 2002), die auch im Zusammenhang mit Concept Mapping berichtet wird (z. B. Cadorin et al., 2014; Cañas & Novak, 2014; Hattie, 2008; Haugwitz, 2009; Hilbert & Renkl, 2008; Nesbit & Adesope, 2006; s. auch Kapitel 4, 1.1).

Im Zusammenhang mit bedeutungsvollem Lernen kann daher festgehalten werden, dass das Concept Mapping von Lernenden kognitive Prozesse der Organisation erfordert, da sie die Beziehungen zwischen den Konzepten ermitteln und visuell darstellen. Das Kriterium der Wissensorganisation wäre in diesem Fall erfüllt. Das Kriterium der aktiven Auseinandersetzung mit dem Lerninhalt wird zudem erfüllt, wenn die Lernenden selbstständig Beziehungen mithilfe der Pfeilverbindungen formulieren und die Inhalte dekontextualisieren, indem die Konzepte

ihrer semantischen Nähe gemäß geordnet werden (s. auch Heinze-Fry & Novak, 1990). Im Zuge der Concept Map Erstellung werden folglich die Informationen nicht isoliert abgespeichert, da neue Konzepte ihre Bedeutung erst durch Beziehungen zu bereits vorhandenem Wissen bzw. Konzepten erhalten (Boo & Hoh, 2001). Dies wird als Indikator für bedeutungsvolles Lernen gesehen (Mayer, 2002). Als ein weiterer Indikator wird die Erstellung von Querverbindungen gesehen, da es sich oft neue und somit kreative Verbindungen der Concept Mapping-Ersteller:innen handelt und auf das Erkennen komplexer Interaktionen auf unterschiedlichen konzeptuellen Ebenen hindeuten (Chevron, 2014; s. Kapitel 2.2.2.3). Darüber hinaus kann – im konkreten Fall des Concept Mappings mithilfe eines Textes – sowohl während als auch nach der Extraktion von Informationen aus dem Lernmaterial das Vorwissen in die Concept Map durch ergänzende Begriffe und Verbindungen integriert werden, was die kognitiven Prozesse der Elaboration (s. Kapitel 2.1.3) erfordert und durch das Concept Mapping gefördert werden kann (DiCarlo, 2006; Hilbert, Nückles, Renkl, et al., 2008; Sumfleth et al., 2010). Eine derartige Organisation von Informationen sowie deren Integration in bestehende Wissensstrukturen durch Lernende soll sowohl ihr Denken in Zusammenhängen sowie eine effektive Integration von komplex strukturierten Informationen beim Lernen fördern (den Elzen-Rump & Leutner, 2007; Lumer & Winter, 2019) als auch das Behalten von Wissen verbessern (Heinze-Fry & Novak, 1990).

Anhand empirischer Befunde (z. B. Haugwitz & Sandmann, 2009; Karpicke & Blunt, 2011b; O'Day & Karpicke, 2020; Reader & Hammond, 1994) ist jedoch davon auszugehen, dass der Einsatz von Concept Maps als Lerninstrument einen differenziellen Einfluss auf den Wissenserwerb in unterschiedlichen Wissensdomänen (bzw. Wissenshauptkategorien nach Krathwohl, 2002) hat. Die Befundlage zur Förderung des Wissenserwerbs in den Domänen des deklarativen, strukturellen und konzeptuellen Wissens ist jedoch uneinheitlich, was auch auf die Nutzung unterschiedlicher Messinstrumente zurückzuführen sein könnte (Haugwitz & Sandmann, 2009; Markow & Lonning, 1998). Zur Erhebung des Wissens von Lernenden in diesen Domänen wurden in der Vergangenheit Messinstrumente wie Multiple-Choice-Tests, Tests mit offenem oder halboffenem Antwortformat und Similarity Judgments Tests (SJT) eingesetzt (Goldsmith et al., 1991; Großschedl, 2010; Großschedl & Harms, 2013a; Haugwitz & Sandmann, 2009). Ebenso wurden Concept Maps selbst eingesetzt, die mithilfe unterschiedlicher Bewertungsschemata ausgewertet wurden (Gouli et al., 2003; Horton et al., 1993; Nesbit & Adesope, 2006; Schroeder et al., 2018; s. auch Kapitel 2.2.1.2).

2.2.2.1 Erwerb deklarativen Wissens

Ausgehend von der Annahme, dass Concept Mapping aufgrund der damit verbundenen Prozesse bedeutungsvolles Lernen fördert (s. Kapitel 2.1.3 und 2.2.2), sollte dadurch auch der Erwerb von deklarativem Wissen unterstützt werden. In diesem Zusammenhang konnten Studien zeigen, dass Faktenwissen, das durch bedeutungsvolles Lernen erworben wurde, länger behalten wird (z. B. Bransford et al., 1999; Reader & Hammond, 1994). Diese Wirkung des Concept Mappings wurde ebenfalls in der Metaanalyse von Nesbit und Adesope (2006) gefunden, in der von einer hohen Effektstärke für die reine Behaltensleistung berichtet wird.

In der Studie von Reader und Hammond (1994) zeigte sich, dass die Studierenden der Concept Mapping-Gruppen die Informationen besser behalten konnten und dies sowohl für deklaratives als auch strukturelles Wissen galt. Die Autoren schlossen daraus, dass Concept Mapping den Wissenserwerb in beiden Domänen fördert, und die Wirkung nicht auf den Erwerb von strukturellem Wissen beschränkt ist. Auch in anderen Studien konnte gezeigt werden, dass Concept Mapping zu besseren Lernleistungen führt im Vergleich zu anderen Lernstrategien wie z. B. dem wiederholten Lesen des Materials (z. B. Bramwell-Lalor & Rainford, 2014; Karpicke & Blunt, 2011b; O'Day & Karpicke, 2020). Andere Forscher:innen hingegen berichten, dass Lernende die Vorteile des Concept Mappings nicht ausnutzen konnten, da ihre Concept Maps eine inadäquate Struktur aufwiesen (Reader & Hammond, 1994) oder Concept Mapping sich als weniger effektiv herausstellte im Vergleich zu anderen Lernstrategien wie dem Notizen Erstellen ohne verfügbares Lernmaterial (z. B. Karpicke & Blunt, 2011b; Markow & Lonning, 1998; O'Day & Karpicke, 2020).

Auch Markow und Lonning (1998) fanden in ihrer Studie zum Vergleich der Lerneffektivität von Concept Mapping mit dem Schreiben von Zusammenfassungen keine positiven Effekte des Concept Mapping. Die Autoren erwägen jedoch, dass die Nachteile einerseits auf eine mangelnde Vertrautheit der Lernenden mit dem Concept Mapping und andererseits auf die Nutzung unterschiedlicher Lern- und Testinstrumente zurückführbar sein könnten. In ihrer Studie erhielten die Studierenden eine Einführung in das Concept Mapping im Rahmen einer einzigen Sitzung und setzten es anschließend als Lernstrategie ein. Ihr Wissen wurde aber später nicht im selben, sondern einem anderen Format (Multiple-Choice-Test) erhoben. Die Autoren empfehlen daher die Nutzung weiterer Instrumente, um das Wissen von Lernenden zu erfassen.

Die Eignung von Multiple Choice Tests zur Wissenserhebung nach dem Lernen wurde auch von Haugwitz und Sandmann (2009) untersucht, die kooperatives Concept Mapping dem Schreiben von Zusammenfassungen gegenüberstellten. Im Fokus stand der Effekt der Lernstrategien auf den Wissenserwerb und die Behaltensleistung von gymnasialen Achtklässler:innen. Eine Woche nach der Interventionsphase wurde die Lernleistung mit einem Multiple Choice Test (zum inhaltlichen und anwendungsbezogenen Wissen) und einem halboffenen Antwortformatstest (nur zum inhaltlichen Wissen) erhoben. Dieselben Tests wurden sechs Monate nach der Interventionsphase erneut eingesetzt. Die Ergebnisse zeigen, dass Concept Mapping im Gegensatz zum Schreiben von Zusammenfassungen einen positiven Einfluss auf den Wissenserwerb, das Behalten des inhaltlichen Wissens und das langfristige Anwendungswissen hatte. Allerdings zeigte sich auch, dass Concept Mapping den Erwerb von Inhaltswissen besser unterstützte als das Lösen von anwendungsorientierten Aufgaben. Die Forscherinnen gaben als mögliche Erklärung, dass die Struktur der Concept Map einem Text überlegen ist und dies einen positiven Einfluss auf die langfristige Behaltensleistung zu haben scheint. Es stellte sich jedoch in einer Analyse der Effektgrößen heraus, dass Fragen im halboffenen Antwortformat im Vergleich mit Multiple-Choice-Fragen nicht besser geeignet waren, den Effekt von Concept Mapping auf das inhaltliche Wissen zu erheben. Die Frage nach geeigneten Messinstrumenten ist jedoch kein Concept Mapping-spezifisches Problem, da auch Forscher:innen aus anderen Bereichen von ungeeigneten Instrumenten für Lernerfolgsmessungen berichten und dafür plädieren, dass erworbenes Wissen auch mit adäquaten Instrumenten gemessen wird (z. B. Acton et al., 1994; McGaghie et al., 2004). Dies gilt auch für die Erfassung strukturellen und konzeptuellen Wissens, die in den folgenden beiden Abschnitten thematisiert werden.

2.2.2.2 Erwerb strukturellen Wissens

Der Einsatz von Concept Maps zur Erfassung des strukturellen Wissens wird von vielen weiteren Forscher:innen empfohlen und/oder umgesetzt, da Concept Maps genutzt werden können, um wichtige Aspekte der Struktur des deklarativen Wissens zu erfassen (Anohina-Naumeca, 2015; Jonassen, 2000; Jonassen & Marra, 1994; Ruiz-Primo & Shavelson, 1997; Walker & King, 2003). Anohina-Naumeca (2015) geht jedoch davon aus, dass Concept Mapping auch als Lernmethode den Lernenden beim Erwerb strukturellen Wissens helfen kann: Zum einen hilft Concept Mapping Lernenden, ihr strukturelles Wissen zu externalisieren. Dies ist zwar auch das Ziel beim Einsatz des Concept Mapping als Diagnoseinstrument, aber mit der Absicht, das Ergebnis der Externalisierung zu bewerten. Zum anderen soll Concept Mapping Lernende dabei unterstützen, ihr Verständnis des Wissensbereichs neu zu konzeptualisieren und

ihr strukturelles Wissen durch das Hinzufügen neuer Propositionen zu erweitern (Anohina-Naumecca, 2015; s. auch Ryssel et al., 2008). Für diese Annahmen sprächen Erkenntnisse aus der Psychologie, die bestätigen, dass eine gut entwickelte Wissensstruktur die Aufnahme neuer Informationen leichter ermöglicht als eine schlecht entwickelte (R. C. Anderson, 1984; H. Gruber & Stamouli, 2020; Renkl, 2020). Ferner ist aus der Forschung zum Concept Mapping bekannt, dass die Organisation von neuen Informationen und deren Verknüpfung mit dem Vorwissen durch die Lernmethode erleichtert wird (z. B. Novak & Cañas, 2008; Reader & Hammond, 1994). Forscher:innen wie Jonassen und Marra (1994) bezeichnen Concept Maps daher auch als Formalismus zur Darstellung von strukturellem Wissen. Reader und Hammond (1994) verweisen zudem darauf, dass sich Concept Mapping zwar auf den Erwerb von strukturellem Wissen auswirkt, dies jedoch einen Dominoeffekt hat, der den Erwerb von Faktenwissen und anderen Arten von Wissen fördert. Ein solcher Effekt sei jedoch auch wieder vom Vorwissen abhängig, da Faktenwissen einfacher integriert werden kann, wenn bereits gut strukturierte Schemata im relevanten Wissensbereich existieren.

Allerdings sind Studien rar, in denen Concept Mapping als Lerninstrument eingesetzt wird, um explizit den Erwerb strukturellen Wissens zu fördern. Einzelne Ausnahmen (z. B. Großschedl & Harms, 2013b), die den Einfluss des Concept Mappings auf die Struktur des Wissens untersuchten, berichten, dass Concept Mapping ebenso lernwirksam war wie das Erstellen von Notizen unter der Bedingung, dass die Concept Maps mit Unterstützung durch metakognitive Prompts erstellt werden konnten. Waren die Prompts nicht verfügbar, zeigte sich ein signifikanter Vorteil des Notizen Erstellens. Andere Autor:innen wie z. B. Leaby et al. (2010) berichten, dass Concept Mapping im Gegensatz zu traditionellen Lehrmethoden keine Vorteile im anschließenden Test brachte, allerdings erzielten Lernende, die bessere Concept Maps erstellten und den Nutzen der Lernstrategie erkannten, höhere Lernerfolge. Daraus ließe sich ableiten, dass das Erkennen des Nutzens von Concept Mapping und höhere Lernerfolge von der Qualität der Concept Maps abhängen, doch die Forscher:innen merken selbst an, dass die Erstellung von drei Concept Maps über ein Semester hinweg womöglich nicht ausgereicht haben könnte, um einen maßgeblichen Einfluss zu haben.

2.2.2.3 Erwerb konzeptuellen Wissens

Konzeptuelles Wissen gilt häufig als Lernziel, da es das Verständnis einer Person widerspiegelt und als Voraussetzung für die kritische Analyse von Problemen und deren Lösungen ist (Montfort et al., 2009; Watson et al., 2016). Concept Maps werden (im weiteren Sinne von *nomen est omen*) eingesetzt, um konzeptuelles Wissen sowohl zu fördern als auch zu erfassen (Besterfield-Sacre et al., 2004; Brandstädter et al., 2012; Cassata et al., 2004;

Hegarty-Hazel & Prosser, 1991a, 1991b; Watson et al., 2016). Bezüglich der Erfassung konzeptuellen Wissens mithilfe von Concept Maps sieht Chevron (2014) v. a. in dem Vorhandensein von Querverbindungen (*cross-links*) in einer Concept Map ein Indiz für ein komplexes und integriertes Wissen (s. Kapitel 2.2.2). Diese Verbindungen stellen ihres Erachtens die Beziehungen zwischen entfernten Konzepten in verschiedenen Bereichen der Concept Map dar und spiegeln komplexe Interaktionen auf unterschiedlichen konzeptuellen Ebenen wider. Als solche sind sie oft neue und somit kreative Verbindungen der Concept Mapping-Ersteller:innen und deuten auf bedeutungsvolles Lernen hin. Romero et al. (2017) sehen in solchen Verbindungen einen weiteren Vorteil des Concept Mappings: Die Autor:innen nehmen an, dass durch die Herstellung neuer Beziehungen zwischen den Begriffen die Kreativität gefördert wird, wobei ebenso ein positiver Effekt auf das Selbstbewusstsein und die Motivation festgestellt werden kann.

Bezüglich des Einflusses von Concept Mapping auf den Erwerb konzeptuellen Wissens berichten Nesbit und Adesope (2006) in ihrer Metaanalyse geringe Effektstärken für das Behalten und gleichzeitige Anwenden von Wissen. Dies könnte insofern positiv gedeutet werden, als Concept Mapping sowohl das Erlernen und Behalten von Inhalten als auch die Anwendung von Wissen fördert. Verantwortlich für die schwächeren Effekte sehen Haugwitz und Sandmann (2009) den kognitiven Anspruch von Anwendungsaufgaben und postulieren, dass Concept Mapping sich eher für den Wissenserwerb ohne Anwendungsbezug eignet. Andere Studien belegen jedoch, dass deklaratives Wissen, welches durch bedeutungsvolles Lernen erworben wurde, nicht nur länger behalten wird, sondern dass dieses Wissen auch viel erfolgreicher zur Lösung neuer Probleme genutzt werden kann (z. B. Bransford et al., 1999).

Solche Vorteile des Concept Mappings konnten allerdings in einigen Studien nicht festgestellt werden: Karpicke und Blunt (2011b) nutzten Concept Maps als Lernmethode und erhoben konzeptuelles Wissen mithilfe von Fragen, die sich auf wortwörtliche Inhalte bezogen (*verbatim questions*), und Transferfragen (*inference questions*). Sie fanden heraus, dass Lernende, die Concept Maps erstellten, schlechtere Leistungen erzielten als diejenigen, die ohne das verfügbare Lernmaterial Notizen erstellten. Andere Autor:innen (z. B. Markow & Lonning, 1998) berichten ebenfalls von nicht nachweisbaren Vorteilen durch Concept Mapping für den Erwerb konzeptuellen Wissens. Allerdings merken Karpicke und Blunt (2011b) bezüglich ihrer verwendeten Messinstrumente an, dass sich die Lern- und Diagnoseinstrumente unterschieden: Während das Concept Mapping zum Lernen eingesetzt wurde, fand die Messung des konzeptuellen Wissens und des Faktenwissens mit anderen Instrumenten statt (s. auch Kapitel 2.2.2.1). Die Autor:innen nutzten darum in einem zweiten Experiment Concept Maps

als Diagnoseinstrument, um die Instrumente anzugleichen. Die Concept Mapping-Gruppen hatten jedoch auch in diesem Format keinen Vorteil gegenüber der Gruppe, die in der Lernphase ohne das verfügbare Lernmaterial Notizen erstellte. Aus diversen Gründen, die im folgenden Kapitel thematisiert werden, wird die Verlässlichkeit der Ergebnisse von Karpicke und Blunt (2011b) jedoch angezweifelt (Mintzes et al., 2011), weshalb die Befunde von Nesbit und Adesope (2006) nur eingeschränkt als widerlegt betrachtet werden können.

Es lässt sich daher festhalten, dass die Lernwirksamkeit des Concept Mapping für alle drei Wissensdomänen postuliert wird; die Ergebnisse der Studien verdeutlichen aber, dass zur Lernwirksamkeit des Concept Mapping in allen drei Wissensdomänen uneinheitliche Befunde vorliegen und daraus unterschiedliche Empfehlungen abgeleitet werden können. Während sich Concept Mapping als lernwirksam erwiesen hat, war der Lernvorteil des Concept Mappings in manchen Wissensdomänen niedriger im Vergleich zu anderen Lernstrategien. Mögliche Gründe können einerseits in den Unterschieden der Erhebungsmethoden liegen, weshalb manche Forscher:innen (z. B. Markow & Lonning, 1998) empfehlen, nicht nur ein Instrument zur Wissenserhebung zu nutzen. Andererseits können die Ergebnisdifferenzen auf einen weiteren Faktor zurückzuführen sein, der zwischen den Studien variierte und als ein zentraler Einflussfaktor für die Effektivität des Concept Mappings als Lernmethode gilt: das vorgeschaltete Training im Concept Mapping.

2.3 Trainingsansätze und Einsatzmöglichkeiten von Concept Mapping

Um Lernende mit dem Concept Mapping vertraut zu machen, wurden in der Vergangenheit unterschiedliche Ansätze entwickelt. Einige spezifische Maßnahmen wie der Einsatz von *skeleton maps* oder Begriffssets wurden bereits in Kapitel 2.2.1.1 vorgestellt. Daran anknüpfend werden im Folgenden Trainingsansätze vorgestellt, von denen manche *skeleton maps* oder andere Lernhilfen wie Feedback (s. Kapitel 2.1.5.2) beinhalteten. Im Anschluss an die Vorstellung unterschiedlicher Trainingsansätze und ihrer Bedeutung für die Lernleistung werden die in der Forschung diskutierten Einsatzmöglichkeiten des Concept Mappings unter dem Aspekt der Textverfügbarkeit thematisiert (s. Kapitel 2.3.2).

2.3.1 Trainingsansätze im Concept Mapping

Bezüglich des Einsatzes von Concept Mapping als Lernmethode weisen diverse Befunde darauf hin, dass Lernende, die mit dem Concept Mapping nicht vertraut sind, Probleme mit der Ausführung und daraus resultierende Lerneinbußen haben (Cañas et al., 2003; Cimolino et al., 2003; Eshuis et al., 2021; Jüngst & Strittmatter, 1995; Michalak & Müller, 2017; Mintzes et al., 2011; Novak & Cañas, 2008). Symptomatisch äußern sich diese Probleme z. B. darin, dass Lernende Schwierigkeiten haben, zwischen Konzepten und Verknüpfungen zu

unterscheiden und die richtigen Verknüpfungsbezeichnungen zu formulieren, um korrekte, sinnvolle Propositionen bei der Erstellung einer Concept Map zu erstellen (Allen & Tanner, 2003; Cimolino et al., 2003; Eshuis et al., 2021; Novak & Cañas, 2008). Ferner berichten Michalak und Müller (2017), dass Personen ohne Concept Mapping-Erfahrung die ganzheitliche Betrachtung einer Grafik fremd war und dies in einer linearen Herangehensweise an die Diagramme und einer Orientierung an Leitfragen resultierte (vgl. Kapitel 1 und 4, 1.3). Cañas et al. (2003) gehen davon aus, dass die Lerneinbußen auf eine zusätzliche kognitive Belastung zurückzuführen sind, die durch unvorbereitetes Concept Mapping entsteht.

Um diesen Problemen vorzubeugen, empfehlen Michalak und Müller (2017) sowie Cañas et al. (2003) eine methodische Anleitung bzw. ein Training, um die kognitive Belastung der Lernenden beim Anwenden der Methode zu reduzieren und höhere Lernerfolge zu erzielen. Diese Überlegungen werden durch die Ergebnisse zahlreicher Studien gestützt, die belegen, dass Lernende durch Trainingsmaßnahmen höhere Lernerfolge erzielen (Nesbit & Adesope, 2006; Schroeder et al., 2018). Darüber hinaus verweisen Hilbert und Renkl (2008) auf die Bedeutung der Erfahrung im Concept Mapping für metakognitive Prozesse und dass Lernende mit mehr Erfahrung im Concept Mapping eine metakognitiv stärker geprägte Herangehensweise aufweisen. Diese äußert sich in einer effektiveren Planung, einer stärkeren Konzentration auf die Verbindungen zwischen Konzepten und einer ausgeprägteren Überwachung ihres Lernfortschritts. Ein Training im Concept Mapping ist laut Mintzes et al. (2011) nicht nur für leistungsschwache, sondern auch für normalerweise leistungsstarke Lernende notwendig, die ebenfalls einen längeren Zeitraum benötigen, um das Concept Mapping zu beherrschen.

Schwendimann (2015) nennt daher das Training als einen der Faktoren, von denen die Effektivität der Strategie abhängt und viele weitere Forscher:innen erwähnen die Bedeutung eines Trainings für den souveränen und lernförderlichen Einsatz von Concept Mapping (Haugwitz & Sandmann, 2009; Hilbert & Renkl, 2005, 2008; Hilbert, Renkl, et al., 2008; Jegede et al., 1990; Jüngst & Strittmatter, 1995; McCagg & Dansereau, 1991; Morse & Juras, 2008; Renkl & Nückles, 2006; Roessger et al., 2018; Romero et al., 2017; Sumfleth et al., 2010; vgl. Kapitel 4, 1.3). Die Interpretation und Umsetzung solcher Empfehlungen ist jedoch uneinheitlich in Forschung und Praxis, zumal der Einfluss von Trainingsmaßnahmen selten explizit untersucht wurde, da in solchen Fällen meist nur die Effektivität der Methode als solche im Studienfokus stand (Arnaudin & Mintzes, 1985; Hilbert & Renkl, 2009; Mintzes et al., 2001; Quinn et al., 2003; vgl. Kapitel 4, 1.3). Daher wurde in der Vergangenheit von diversen Forscher:innen (z. B. Hilbert & Renkl, 2008, 2009; Karpicke & Blunt, 2011b) der Mangel an

klaren Handlungsempfehlungen bezüglich der Inhalte und Dauer eines geeigneten Trainings moniert. Dieses Desiderat wird im hier vorgestellten Projekt erforscht. Im Fokus steht die Untersuchung unterschiedlich intensiver, direkter Trainingsansätze für das Concept Mapping, die eingesetzt werden, um die Vertrautheit der Lernenden mit der Methode zu fördern. Diese Vertrautheit mit Concept Mapping ist, wie bereits betont, nicht nur für dessen Nutzung als Lerninstrument, sondern auch als Diagnoseinstrument entscheidend (s. Kapitel 2.2.1.1 und 2.2.1.2).

Für die Vermittlung von Concept Mapping gab es diverse Versuche, die häufig eine Anleitung mit sukzessiv auszuführenden Schritten einschlossen. Die unterschiedlichen Concept Mapping-Trainings haben gemeinsam, dass den Lernenden die erforderlichen Schritte beigebracht werden sollen: wichtige Begriffe identifizieren, hierarchische Beziehungen zwischen Begriffen bestimmen und sinnvolle Propositionen bilden (vgl. Kapitel 5, 1.3), wobei diese Schritte durch metakognitive Prozesse ergänzt werden können bzw. sollen. Diese für das selbstständige und lerneffektive Concept Mapping wichtigen Schritte werden von Forscher:innen unterschiedlich gewichtet und erneut untergliedert, doch die meisten Trainings, die in der wissenschaftlichen Literatur über Concept Mapping beschrieben werden, sind an die Empfehlungen von Novak und Cañas (2006, 2008) angelehnt (vgl. Kapitel 4, 1.3). Novak und Cañas (2006, 2008) empfehlen, dass eine Fokusfrage zu Beginn gewählt wird, die in der Concept Map gelöst werden soll. Danach werden die wichtigsten 15-25 Konzepte identifiziert, in einer vorläufigen Concept Map angeordnet und miteinander verbunden. Sobald die erste Version erstellt wurde, soll nach möglichen Querverbindungen gesucht werden, die wiederum das Verständnis der Lernenden bezüglich der Verbindungen zwischen unterschiedlichen Bereichen der Concept Map repräsentieren.

Im Gegensatz dazu empfehlen Quinn et al. (2003) das Identifizieren der lediglich 10-15 wichtigsten Konzepte, deren Arrangieren auf Papier und das Erstellen von Querverbindungen. Jonassen et al. (1993) schlagen zehn Schritte zur Konstruktion von Concept Maps vor, darunter das Sammeln, Selektieren und Ordnen von Konzepten, das Erstellen der Verbindungen sowie das Evaluieren der Concept Map. Während sich diese Schritte vornehmlich auf die Erstellung von Propositionen beziehen, konzentrieren sich Hilbert und Renkl (2008, 2009) auf die metakognitive Ebene und beschreiben erfolgreiches Concept Mapping als Abfolge von drei sich immer wiederholenden Schritten: 1. Planung der Concept Map, 2. Achtgeben auf die Beziehungen zwischen den Konzepten bei der eigentlichen Erstellung der Concept Map und 3. ständige Kontrolle auf Richtigkeit und Vollständigkeit. Diese metakognitiven Schritte gewinnen vor dem Hintergrund der Erkenntnisse von Chiu (2004) an Bedeutung: Chiu (2004)

stellte fest, dass Ergänzungen oder Änderungen an der Concept Map für Lernende sehr frustrierend sein können, was Lernende davon ablenkt, sich auf ihr Wissen zu konzentrieren, und den Lerneffekt schmälert. Dementsprechend wäre die Fokussierung auf die Vermittlung metakognitiver Prozesse der Planung hilfreich, wie Hilbert und Renkl (2008, 2009) sie empfehlen, um zu verhindern, dass nachträglich notwendige Veränderungen in der Concept Map in einem lernhinderlichen Maß vorgenommen werden müssen.

Ungeachtet der Anzahl der Schritte findet das Training in der Regel unter persönlicher Anleitung durch eine Lehrkraft statt, die den Lernenden sowohl theoretisches Input als auch die Möglichkeit gibt, vor der kritischen Lernphase praktische Erfahrungen zu sammeln, auch wenn es Ausnahmen gibt (z. B. Pruetz & Weigel, 2020). Pruetz und Weigel (2020) gaben ihren Teilnehmer:innen detaillierte Anleitungen zum Erstellen von Concept Maps und stellten Online-Materialien zur Verfügung, nutzten jedoch keine gemeinsamen Trainingssitzungen, um die Proband:innen auf die Erstellung von digitalen Concept Maps vorzubereiten. Die Autoren merken selbstkritisch an, dass sich diese Herangehensweise negativ auf den wahrgenommenen Nutzen sowie die Qualität der Concept Maps ausgewirkt haben könnte. Abgesehen von dieser Ausnahme werden basierend auf bisherigen Forschungsberichten zwei Kategorien des Concept Mapping-Trainings im Folgenden vorgestellt: die Kurzeinführung und das extensive Training.

2.3.1.1 Kurzeinführungen

Kurzeinführungen basieren auf der Annahme, dass die Charakteristika einer Concept Map in relativ kurzer Zeit erklärt werden können und Concept Mapping relativ leicht zu erlernen ist, wie Forscher:innen (z. B. Jonassen et al., 1993) in der Vergangenheit betont haben. Hay et al. (2008, vgl. S. 302) postulieren, dass Concept Mapping in einer einzigen Sitzung innerhalb von 10-20 Minuten gelehrt werden kann und dass die meisten Lernenden weitere 20-30 Minuten für ausreichend halten, um eine vernünftige Concept Map zu erstellen. Diese Angaben veranschaulichen, was unter einer Kurzeinführung i. A. verstanden werden kann: Eine einmalige Sitzung mit einer theoretischen Einführung inklusive Vorstellung der Charakteristika einer Concept Map zur Vermittlung deklarativen Wissens zur Lernstrategie. Dieser theoretische Part kann ggf. in derselben Sitzung um ein zusätzliches Zeitfenster für das praktische Ausprobieren der Strategie ergänzt werden, um prozedurales Wissen (s. Kapitel 2.1.2) zu vermitteln. Solche Kurzeinführungen wurden, basierend auf den oben erwähnten Annahmen und Empfehlungen, in mehreren Studien durchgeführt (z. B. Blunt & Karpicke, 2014; Eshuis et al., 2021; Großschedl & Harms, 2013b; Hilbert & Renkl, 2009; Karpicke & Blunt, 2011b; O'Day & Karpicke, 2020).

Die vorgeschlagenen Trainingsparameter von Hay et al. (2008) wurden in vergleichbarer Form in der Studie von Großschedl und Harms (2013b) implementiert, allerdings verfolgten sie einen Trainingsansatz mit theoretischen und praktischen Trainingselementen, die jeweils 15 Minuten lang waren. Eine weitere Studie, die eine ähnlich lange Einführung in das Concept Mapping verwendete, wurde von Brandstädter et al. (2012) durchgeführt, deren theoretische Einführung 15 Minuten und der praktische Part 20 Minuten umfasste. O'Day und Karpicke (2020) basierten ihre Kurzeinführung in das Concept Mapping auf Trainings früherer Studien (Blunt & Karpicke, 2014; Karpicke & Blunt, 2011b) und stellten den Teilnehmer:innen zusätzlich zur persönlichen Einführung ein Handout mit einer Concept Map als Beispiel zur Verfügung, die sie während der Concept Mapping-Lernaufgabe zu Rate ziehen konnten. Die Ergebnisse von O'Day und Karpicke (2020) zeigen, dass Studierende, die Concept Maps erstellten, in einem unmittelbar folgenden Test besser abschnitten als diejenigen, die nur den Text lasen, und solche, die die Inhalte des zuvor gelesenen Texts in Form von Notizen schriftlich festhielten (*retrieval practice*). Dieser Vorteil der Concept Mapping-Gruppe verschwand jedoch innerhalb einer Woche wieder und spiegelt die Nachteile des Concept Mappings gegenüber dem Abruf aus dem Gedächtnis (*retrieval practice*) von z. B. Karpicke und Blunt (2011b) wider, auf dessen Trainingsansatz O'Day und Karpicke (2020) den ihren basierten hatten. Das Training von Karpicke und Blunt (2011b) wurde jedoch schon seinerzeit von Mintzes et al. (2011) kritisiert, die vor dem Hintergrund ihrer eigenen Erkenntnisse die Bedeutung einer ausreichenden Vertrautheit der Lernenden mit dem Concept Mapping betont hatten. Mintzes et al. (2011) vermuteten, dass eine Kurzeinführung nicht ausreicht, um die anspruchsvolle Lernmethode adäquat einzusetzen, wohingegen sich die Leistungen der Proband:innen deutlich verbessern, wenn Trainings über mehrere Wochen und mit Feedback implementiert werden. Die Nachteile der Concept Mapping-Methode seien folglich auf mangelnde Übung und Vertrautheit der Teilnehmer:innen zurückzuführen. Während Karpicke und Blunt (2011a) ihr Vorgehen mit Verweisen auf andere Studien begründen, in denen ebenfalls derartige Kurzeinführungen genutzt wurden, relativierten andere Forscher:innen wie Jüngst und Strittmatter (1995) ihre eigenen Aussagen: Trotz ihrer Annahmen zur vermeintlichen Leichtigkeit der Concept Mapping-Erlernbarkeit merken sie an, dass die Lernenden erst nach 10 bis 15 selbst erstellten Concept Maps ausreichend mit Concept Mapping als Lerninstrument vertraut sind.

2.3.1.2 Extensive Trainings

Derartige Übungsmöglichkeiten werden Lernenden in einem extensiven Training gegeben, das über mehrere Sitzungen hinweg stattfinden und zusätzliche Elemente wie Prompting und Feedbackmaßnahmen beinhalten kann (Miller et al., 2008; Mintzes et al., 2011; Pearsall et al., 1997; Schwendimann, 2015; s. auch B. L. Martin et al., 2000; Quinn et al., 2003). Doch auch in extensiven Trainings variieren Parameter wie die Gesamttrainingsdauer und die Anzahl der Sitzungen zwischen den Trainings: Während in der Studie von Pearsall et al. (1997) eine einzige Trainingssitzung von drei Stunden durchgeführt und über ein gesamtes Semester hinweg Feedback zu den erstellten Concept Maps gegeben wurde, nutzten andere Forscher:innen Trainingseinheiten à 60 Minuten für zwei Sitzungen (Romero et al., 2017) oder drei Sitzungen (Sumfleth et al., 2010).

Ein Training kann jedoch auch über die Sitzungen vor Ort hinausgehen, wie Quinn et al. (2003) zeigen: Sie gaben ihren Studienteilnehmer:innen eine schriftliche Anleitung zum Concept Mapping und die Anweisung, eine Concept Map zu einem vertrauten Thema (z. B. Lebensmittel) als Übung zu erstellen. Die Teilnehmer:innen erstellten Concept Maps in der Sitzung, indem sie in Dreier- oder Vierergruppen arbeiteten, und erhielten zusätzlich Concept Map-Hausaufgaben, um weiter zu üben. Basierend auf ihren Erkenntnissen aus dieser Studie und anderen Erfahrungen kommen sie zu dem Schluss, dass praktisch alle Studierenden innerhalb einer 50-minütigen Sitzung das Erstellen einer einfachen Concept Map lernen können und dabei Übung und Sicherheit im Umgang mit der Methode erlangen (vgl. Quinn et al., 2003, S. 12).

Eine noch längere Trainingsphase, die jedoch ausschließlich unter Anleitung stattfand, wurde von Jegede et al. (1990) genutzt, die Schüler:innen der zehnten Klasse über einen Zeitraum von drei Wochen im Concept Mapping trainieren konnten, einschließlich theoretischer und praktischer Phasen. Ihre Ergebnisse zeigen, dass ihr Concept Mapping-Training und der anschließende Einsatz der Methode zu signifikant besseren Testergebnissen in Biologie führte als eine traditionelle Unterrichtsstrategie. Ein höherer Lernerfolg bzw. eine höhere Concept Map Qualität durch ein extensives Training zeigte sich auch in weiteren Studien (z. B. Pearsall et al., 1997; Sumfleth et al., 2010). Die Bedeutung einer ausreichenden Trainingsdauer fassen Hilbert und Renkl (2009) in dem Prinzip „je mehr, desto besser“ zusammen: „It is important to note that the more training learners have in using concept mapping for learning, the more successful they are in applying this learning technique“ (Hilbert & Renkl, 2009, S. 269).

Ergebnisse anderer Studien legen jedoch nahe, dass Concept Mapping ausreichend, aber nicht unnötig lang trainiert bzw. eingesetzt werden sollte: Nesbit und Adesope (2006) zeigen in ihrer Metaanalyse, dass Concept Mapping effektiver als andere Lernmethoden war, wenn Lernende über einen Zeitraum von bis zu fünf Wochen Concept Maps erstellten im Vergleich zu einem längeren Zeitraum. Schroeder et al. (2018) vermuten, dass mit der Neueinführung der Methode ein *novelty effect* (Cook & King, 1968) verbunden ist, der dazu führt, dass Lernende aufgrund der Neuheit der Lernmethode bis zu einem gewissen Zeitpunkt besser lernen, bis dieser Effekt verschwindet. Aufgrund ihrer eigenen Ergebnisse kommen Schroeder et al. (2018) zu einer anderen Einschätzung, die wiederum der Aussage von Hilbert und Renkl (2009) Gewicht verleiht: Die Analysen von Schroeder et al. (2018) zeigen, dass Studien, die weniger als einer Woche dauerten, wesentlich weniger effektiv waren als Studien mit einer Dauer von einer bis vier Wochen und Studien mit einer Dauer von mehr als vier Wochen. Die Autor:innen schließen daraus, dass mit zunehmender Erfahrung der Lernenden im Concept Mapping die kognitive Belastung, die mit dem Ausführen der methodischen Schritte selbst verbunden ist, abnimmt. Folglich können die Lernenden ihre kognitiven Ressourcen dem Lernmaterial widmen und müssen sich nicht auf die Anwendung der Methode konzentrieren.

Daraus lässt sich ableiten, dass eine Grundvertrautheit mit der Lernmethode wichtig für einen lernwirksamen Einsatz ist, aber ein Training darüber hinaus unbeabsichtigte negative Effekte haben könnte. Insbesondere die Ergebnisse von Nesbit und Adesope (2006) legen nahe, dass sich sowohl eine zu kurze als auch eine zu lange Implementierung negativ auf die Wirksamkeit auswirken kann. Dies verdeutlicht, wie wichtig eine Anpassung in beiden Fällen an die Voraussetzungen der Lernenden und Umstände ist. Wie eine solche Anpassung vollzogen werden kann, zeigt sich an zwei Unterkategorien der extensiven Trainings: dem extensiven Training ohne Lernhilfen und dem Training mit Lernhilfen.

Generell erfordert ein Training ohne Lernhilfen ein selbstständigeres Arbeiten, stellt aber die weniger aufwändigere Variante dar, die sich ebenfalls als lernwirksam erwiesen hat (Hauser et al., 2006) und das Vorgehen in Schulklassen und Hochschullehrveranstaltungen abbildet (Hilbert & Renkl, 2009; Serbessa, 2006; Tsai, 2002; s. auch Kapitel 6.1.4). Hauser et al. (2006) analysierten die Auswirkungen verschiedener Grade von Scaffolding für Concept Mapping beim Lernen aus Texten. Die Studierenden wurden in eine von fünf Gruppen geteilt. Drei Gruppen erstellten Concept Maps (*mit einem Begriffsset vs. räumlich angeordnete Konzepte, die nur verknüpft werden vs. Konstruktion einer Concept Map von Grund auf*), eine Gruppe studierte eine ausgearbeitete Expert:innen-Concept-Map und eine Kontrollgruppe führte keine Aktivitäten nach dem Lesen des Textes durch. Die Konstruktion einer Concept

Map von Grund auf erwies sich als fast ebenso hilfreich wie die Bereitstellung einer ausgearbeiteten Map, was das Lernen im Vergleich zu den anderen Bedingungen am effektivsten unterstützte. Allerdings lasen die Studierenden der Concept Mapping-Gruppen nur einen Einführungstext zum Concept Mapping, was eine Kurzeinführung mit den damit verbundenen potenziellen Lernnachteilen darstellt. Dagegen berücksichtigt ein extensives Training ohne Lernhilfen einerseits, dass Concept Mapping mehr Übung verlangt als den Lernenden im Rahmen einer Kurzeinführung gewährt wird. Andererseits kann dem Umstand Rechnung getragen werden, dass Lernhilfen wie Promptingmaßnahmen und Feedback einen zusätzlichen Aufwand an Zeit und Energie für die Lehrenden bedeuten und dieser Aufwand in der Praxis aufgrund limitierter Ressourcen nicht immer geleistet werden kann (Hilbert & Renkl, 2009; s. Kapitel 2.1.5.2). Nach der Einführung wird den Lernenden daher wiederholt Zeit zum Üben gegeben, doch der Lernprozess wird nicht weiter durch die Lehrkraft begleitet.

Das Training mit Lernhilfen stellt demnach zwar das aufwändigere dar, berücksichtigt aber, dass der Lernstrategieeinsatz ohne Lernhilfen in einem defizitären Nutzen der Strategie resultieren kann, da die selbstständige Erstellung von Concept Maps überfordernd sein kann (Allen & Tanner, 2003; s. auch Kapitel 2.2.1.1). Die Überforderung zeigt sich beim Concept Mapping z. B. darin, dass einfache Kausalkettenstrukturen oder lineare Ketten erstellt werden, die jedoch ein Hindernis für das Lernen komplexer Inhalte und Prozesse darstellen, wie in diversen Studien (z. B. Perkins & Grotzer, 2005; Raia, 2005; Webb & Boltt, 1990) gezeigt werden konnte. Wie die Umsetzung von Lernhilfen im Zusammenhang mit Concept Mapping aussehen kann, wird im Folgenden vorgestellt, ergänzend zu den Ausführungen in den Kapiteln bezüglich des Einsatzes von Lernhilfen im Allgemeinen (s. Kapitel 2.1.5.2) und Scaffolds im Speziellen wie *skeleton maps* (s. Kapitel 2.2.1.1). Ein Fokus wird hierbei auf Promptingmaßnahmen und Feedback gelegt, die sich unabhängig voneinander in diversen Studien als lernwirksam erwiesen haben (z. B. Großschedl & Harms, 2013b; Hilbert, Nückles, Renkl, et al., 2008; Roessger et al., 2018) und auch im hier vorgestellten Projekt eingesetzt wurden (s. Kapitel 4 und 5).

2.3.1.3 Lernhilfen im Concept Mapping-Training

Neben Handouts, die eher einer generellen Erklärung von Concept Map-Charakteristika dienen und die theoretischen Ausführungshinweise auch in der Lernaufgabe verfügbar machen sollen (z. B. Chiu, 2004; Chularut & DeBacker, 2004; O'Day & Karpicke, 2020), wurden sowohl in Kurzeinführungen als auch Trainings mit mehreren Sitzungen Promptingmaßnahmen (z. B. Eshuis et al., 2021; Großschedl & Harms, 2013b; Hilbert, Nückles, Renkl, et al., 2008; Hilbert & Renkl, 2009; Nückles et al., 2010) und Feedback (z. B. Pearsall et al., 1997; Roessger

et al., 2018) eingesetzt, um Lernenden bei der Ausführung von Concept Mapping zu helfen und ihr Lernen zu verbessern.

- Promptingmaßnahmen

Zusätzlich zu ihrer Kurzeinführung in das Concept Mapping von 30 Minuten nutzten Hilbert und Renkl (2009) Prompts, die Lernenden dabei helfen sollten, sich heuristische Beispiele selbst zu erklären. Ihre Ergebnisse zeigen, dass diese Kombination Lernenden half, Concept Mapping für das Lernen aus Texten effektiv zu nutzen. Ein weiterer Ansatz, der in einigen Studien (z. B. Großschedl & Harms, 2013b; Hilbert, Nückles, Renkl, et al., 2008) eingesetzt wurde, um Lernenden beim Concept Mapping zu helfen und ihre metakognitiven Fähigkeiten zu unterstützen, stellt der Einsatz metakognitiver Prompts dar. Die Ergebnisse von Hilbert, Nückles, Renkl et al. (2008) deuten darauf hin, dass Lernende, die metakognitive Prompts beim Concept Mapping nutzten, deutlich bessere Lernergebnisse erzielten als die Kontrollgruppe ohne Prompting. Dies zeigt, dass Lernende, die mit Concept Mapping nicht vertraut sind, Probleme bei der Kontrolle ihres Mapping-Prozesses haben und der Unterstützung bedürfen. Zu ähnlichen Erkenntnissen kamen Großschedl und Harms (2013b): Sie untersuchten die Effekte der Faktoren *Lernstrategie* (Concept Mapping vs. Notizen Erstellen) und *Verfügbarkeit von metakognitiven Prompts* (Prompting vs. kein Prompting) auf den Erwerb konzeptuellen Wissens von Biologiestudierenden. Ihre Ergebnisse zeigen, dass Lernende, die Concept Maps erstellen, mehr von den Prompts für ihren Wissenserwerb profitierten als solche, die Notizen erstellen. Aus diesen Befunden leiteten die Forscher:innen ab, dass der Nutzen von Lernhilfen (in diesem Fall Prompts) auch vom Vertrautheitsgrad der Lernstrategie abhängt und Lernende, die mit dem Concept Mapping unvertraut sind, in ihrer Anwendung metakognitiver Lernstrategien z. B. durch entsprechende Prompts unterstützt werden sollten.

Bei der Verwendung von Prompts sollte jedoch beachtet werden, dass sie so häufig wie nötig zur Verfügung gestellt werden sollten, damit Lernende die Prompt-induzierte Herangehensweise fest in ihr Handeln integrieren, wie Hilbert, Nückles, Renkl et al. (2008) erkannten. In ihrer Studie stellte sich heraus, dass die einmalige Bereitstellung von Prompts nicht ausreichte, um Lernende in die Lage zu versetzen, die Lernstrategie auch in einem Follow-up-Test selbstständig erfolgreich anzuwenden, weshalb sie einen mehrmaligen Einsatz von Prompts empfehlen. Die Unterstützung durch Prompts sollte jedoch auch nicht zu lange stattfinden, da eine potenzielle Gefahrenquelle darin liegt, dass der Zweck der strategischen Selbstständigkeit von einigen Lernenden nicht erkannt wird, wenn sie an strategische Anweisungen von außen gewöhnt sind. Ähnliche Überlegungen stellten auch Nückles et al. (2010) an und verzichteten auf eine durchgängige Bereitstellung derselben Prompts. Stattdessen wählten sie den Ansatz des

fading, in dem die Prompts mit zunehmender Vertrautheit der Lernenden mit der Strategie reduziert wurden. Dies führte zu einer geringeren extrinsischen Belastung, da auf eine Redundanz und zusätzliche Belastung durch bekannte und somit obsolete Prompts verzichtet wurde, die das Lernen hätten behindern können. Die Ergebnisse von Nückles et al. (2010) belegen, dass bei einem Training zum Lernprotokoll-Schreiben durch die Anleitung von Prompts kognitive Lernstrategien vermehrt genutzt wurden und zu besseren Lernergebnissen führten.

- Feedback-Maßnahmen

Neben Promptingmaßnahmen kann Feedback eingesetzt werden, dessen unterschiedliche Einsatzmöglichkeiten bei der Lernstrategienutzung und deren Bedeutung für den Lernerfolg in Kapitel 2.1.5 vorgestellt wurden. Feedback bietet Lehrkräften die Möglichkeit, auf Wissenslücken und Missverständnisse der Lernenden zu reagieren, die bestehen bleiben können, selbst wenn die Lernenden Concept Maps erstellen und Wissenslücken selbstständig erkennen sollten (Cimolino et al., 2003; Novak, 2002). Forscher:innen (z. B. Chang et al., 2002; Eshuis et al., 2021; Kinchin, 2014; Morse & Jutras, 2008) gehen davon aus, dass das generelle Fehlen von angemessenem Feedback bei der Erstellung von Concept Maps eine mögliche Ursache für fortbestehende Wissenslücken sein könnte und empfehlen dessen Einsatz. Trotz des erhöhten Zeitaufwands, der mit Feedback verbunden ist, wurde ein mehrwöchiges Training mit wiederholten Übungsmöglichkeiten und Feedbackmaßnahmen aufgrund der trainingsassoziierten Lernerfolge in mehreren Studien zum Concept Mapping eingesetzt (z. B. Pearsall et al., 1997; Roessger et al., 2018; s. auch B. L. Martin et al., 2000; Mintzes et al., 2011; Morse & Jutras, 2008; Nesbit & Adesope, 2006; Quinn et al., 2003; Schroeder et al., 2018; Schwendimann, 2021). Einige Forscher:innen wie Roessger et al. (2018) integrierten in ihrem Concept Mapping-Training Feedback und erkannten, dass sich die Qualität der strukturellen Concept Maps der Studierenden durch wiederholtes Üben und Feedback verbesserte. Darüber hinaus verbesserte sich die Anzahl der Elaborationen in den Concept Maps, was den Lernerfolg erhöhte. Auch in der Studie von Pearsall et al. (1997) zeigte sich, dass eine Kombination aus extensiver Concept Mapping-Einführung (drei Stunden) und Feedback (über ein Semester hinweg) dazu führte, dass die Concept Maps zunehmend an Qualität gewannen: Im Laufe der Zeit wiesen die Concept Maps der Lernenden zunehmend elaborierte sowie differenzierte Strukturen auf, was darauf schließen lässt, dass die Lernenden offensichtlich das kontinuierliche Feedback nutzen konnten und sich auch ihre Wissensstrukturen verbesserten (vgl. Kapitel 4, 1.3).

Eine deutlich kürzere Einführung erhielten die Teilnehmer:innen in den Studien von Eshuis et al. (2021) und Chularut und DeBacker (2004). Letztere nutzten ein Training für die Concept Mapping-Gruppe, das aus einer 30-minütigen Einführung in Concept Mapping bestand, allerdings erhielten die Teilnehmer:innen zusätzliche Handouts und Feedback in mehreren Sitzungen. Die Handouts enthielten eine Einführung in Concept Mapping, eine Liste von Concept Map-Merkmalen sowie Beispiele von Concept Maps. Das Feedback erfolgte, indem die Concept Maps am Ende der Sitzung eingesammelt und in der folgenden Sitzung ausgeteilt wurden. Die Ergebnisse der Studie zeigen, dass sich die Concept Mapping-Gruppe zwischen Prä- und Posttest signifikant stärker verbesserte als die Lerngruppe, die ihre selbstgewählten Lernstrategien ohne Anleitung praktizierten. Die Kurzeinführung von Eshuis et al. (2021) beinhaltete, dass die Lernenden kurz darüber informiert wurden, was eine Concept Map ist und welche Unterschiede zwischen Konzepten und Verknüpfungen bestehen. Dies wurde anhand eines Concept Map-Beispiels aus einem anderen Bereich veranschaulicht. Darüber hinaus wurden Anweisungen zur Erstellung einer Concept Map mit dem digitalen Concept Mapping-Tool gegeben. Nach der Concept Mapping-Aufgabe wurden die Concept Maps der Lernenden je nach Bedingung ergänzt durch (a) Feedback in Form einer Expert:innenlösung zum Vergleichen, wobei die Expert:innen-Concept-Map in die eigene Lösung integriert werden konnte (kombinierte Concept Map) und Reflexionsprompts, (b) nur die kombinierte Concept Map oder (c) keine kombinierte Concept Map und keine Prompts. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied im Wissenszuwachs zwischen den Bedingungen gab und der Lernzuwachs zwischen Prä- und Posttest in allen Gruppen vergleichsweise niedrig ausfiel. Allerdings erzielten Lernende einen höheren Lernzuwachs, wenn sie die kombinierte Concept Map häufiger konsultierten als solche, die weniger Zeit in den Vergleich der Lösungen investierten. Die Autor:innen schließen daraus, dass Lernende z. T. nicht in der Lage waren, die Hilfen gewinnbringend zu nutzen und daher zusätzliche Unterstützung brauchen, um höhere Lernerfolge zu erzielen. Sie diskutieren die Nutzung der Prompts und ihre Form des Feedbacks, doch erwähnen auch den potenziellen Nutzen eines ausführlicheren methodischen Trainings im Concept Mapping, um die Qualität der Concept Maps zu steigern, die in ihrer Studie wohl verbesserungswürdig war.

Zusammenfassend ist festzustellen, dass Concept Mapping hohe kognitive Anforderungen an ihre Nutzer:innen stellt, die durch ein Training verringert werden sollen. Trainingsmaßnahmen werden in der Forschung einerseits empfohlen, andererseits ist die Auswahl bezüglich entscheidender Trainingsparameter davon abhängig, bei welcher Zielgruppe und mit welchem Ziel Concept Mapping eingesetzt wird. Die Trainings

unterscheiden sich bezüglich der Trainingsdauer (von 10 Minuten bis drei Stunden), der Trainingsintensität (eine oder mehrere Trainingssitzungen), der Trainingsmethode (indirektes Training zur Vermittlung deklarativen Wissens oder direktes Training zur Vermittlung deklarativen und prozeduralen Wissens) und dem Maß an Selbstständigkeit (Level an Unterstützung, z. B. durch Scaffolds in dem Lernmaterial; vgl. Kapitel 4, 1.3). Trotz einer generellen Anerkennung der Bedeutung von Trainingsmaßnahmen ist jedoch aufgrund der unterschiedlichen Trainingsansätze eine Entscheidung bezüglich der berichteten Lernerfolgsergebnisse nur eingeschränkt möglich. Ein Desiderat stellt somit eine auf empirischen Befunden basierte Handreichung für den Einsatz von Concept Mapping in der Praxis dar. Dies könnte durch die Variation von Trainingsparametern wie der Dauer, der Intensität und der Trainingsmethode erreicht werden, wobei die unterschiedlichen Trainingsansätze sich auf die Förderung kognitiver Prozesse und Reduzierung der lernhinderlichen kognitiven Belastung sowie den Lernerfolg auswirken sollten, da zahlreiche Studien (vgl. Chang et al., 2002; Hilbert & Renkl, 2005, 2008; Sumfleth et al., 2010) nachweisen konnten, dass die Implementierung von Trainingsmaßnahmen zu höheren Lernerfolgen bei Lernenden führen kann. Ein erstes Ziel des Forschungsprojektes ist es daher, die Frage zu beantworten, ob Lernende, die ein spezifisches Concept Mapping-Training erhalten haben, erfolgreicher im Lernen sind, wenn sie Concept Maps erstellen, als solche, die auf eine andere Weise an Concept Mapping herangeführt wurden. Darauf aufbauend sollen Möglichkeiten der Trainingsoptimierung ermittelt werden, um die Effizienz und Lernwirksamkeit von Trainingsmaßnahmen im Concept Mapping zu verbessern.

2.3.2 Einsatzmöglichkeiten von Concept Mapping

Neben der Diskussion zu Trainingsmaßnahmen wird eine weitere bezüglich des Einsatzes von Concept Mapping in unterschiedlichen Versuchsbedingungen geführt: Dabei wird die Verfügbarkeit des Lernmaterials variiert, da dem Abruf von Informationen aus dem Gedächtnis in der letzten Dekade zunehmend Aufmerksamkeit geschenkt wurde und Befunde darauf hinweisen, dass durch einen solchen gezielten Abruf (*retrieval practice*) die Gedächtnisorganisation und somit auch der Lernerfolg beeinflusst werden (Carpenter et al., 2009; Endres et al., 2017; Karpicke, 2017; Karpicke & Roediger, 2008; Pyc & Rawson, 2009; Roediger et al., 2011; Roediger & Karpicke, 2006a).

2.3.2.1 Elaborationssettings

Bei Verfügbarkeit des Lernmaterials wird erwartet, dass Lernende sich aktiv damit auseinandersetzen, um neue Informationen in ihr Vorwissen zu integrieren, ergo zu elaborieren. Versuchsbedingungen, in denen das Lernmaterial verfügbar ist, werden dieser Argumentation folgend als Elaborationssetting bezeichnet. Beim Concept Mapping steht den Lernenden in der Regel das entsprechende Lernmaterial (z. B. in Form eines Textes) zur Verfügung. Diese Herangehensweise soll einerseits Organisations- und Elaborationsprozesse fördern, da die Informationen des Textes in der Concept Map organisiert und neue Informationen in bestehende Wissensstrukturen integriert werden können (Mayer, 2002a; Nückles et al., 2010). Andererseits ermöglicht der Abgleich zwischen der eigenen Concept Map und dem Text das Erkennen von Verständnislücken und somit die Entwicklung metakognitiver Fähigkeiten (Bramwell-Lalor & Rainford, 2014; Kinchin, 2011; Mintzes et al., 1997; Ritchhart et al., 2009; Salmon & Kelly, 2014). Auf diesem Weg sollen Lernende in die Lage versetzt werden, den Text mittels Concept Mapping zu verstehen, bevor ihr Lernerfolg gemessen wird (z. B. Karpicke & Blunt, 2011b; Nesbit & Adesope, 2006; Sumfleth et al., 2010; vgl. Kapitel 5, 1.2).

2.3.2.2 Retrievalsettings

Ist das Lernmaterial nicht verfügbar, wird angenommen, dass Abruf- bzw. Retrieval-Prozesse stattfinden, wenn Lernende Informationen aus dem Gedächtnis abrufen, weshalb diese Bedingungen als Retrievalsettings bezeichnet werden (Blunt & Karpicke, 2014; Großschedl & Tröbst, 2018; Karpicke & Blunt, 2011b; Karpicke & Smith, 2012; vgl. Kapitel 5, 1.2). Bisherige Studien zu den Effekten der beiden Ansätze konnten zeigen, dass die Implementierung von Designs, die von Lernenden solche Abrufprozesse erfordern, den Lernenden nicht nur die Möglichkeit bieten, Lerninhalte zu organisieren, sondern auch ihren Lernfortschritt zu überprüfen und somit ihre Metakognition zu verbessern (z. B. Endres et al., 2017; Roediger et al., 2011; Stanton et al., 2021; Zaromb & Roediger, 2010). Ferner hat sich gezeigt, dass das Verständnis der Lernenden durch konstruktive Elaborationsprozesse während des Abrufs (Endres et al., 2017) sowie das Lernen selbst verbessert wird (z. B. Brame & Biel, 2015; Karpicke, 2009; Karpicke & Smith, 2012; Roediger & Karpicke, 2006b; Toppino & Cohen, 2009; für einen Überblick s. auch Moreira et al., 2019). Diesbezüglich zeigte sich in mehreren Studien (z. B. Carpenter, 2009; Carpenter & DeLosh, 2006; Endres et al., 2017; Roediger & Karpicke, 2006b), dass bereits ein einmaliger Abruf zwischen Lernaufgabe und dem Erinnerungstest die Behaltensleistung deutlich steigern kann.

Als Erklärungen für diese Effekte werden mehrere mögliche Optionen diskutiert (z. B. Carpenter, 2009; Carpenter & DeLosh, 2006; Lipowsky, 2015; Peters, 2012; Roediger et al., 2011). Zum einen wird angenommen, dass durch den Informationsabruf die Wahrscheinlichkeit der Elaboration steigt, ergo, dass neue Informationen in bestehende Strukturen stärker integriert und dadurch leichter erinnert werden. Ergänzend oder alternativ können durch den Abruf zusätzliche Abrufpfade gebildet werden, die „künftig den Abruf erleichtern bzw. einen erfolgreichen Abruf wahrscheinlicher machen“ (Peters, 2012, S. 102; s. auch Carpenter, 2009; Carpenter & DeLosh, 2006). Zum anderen ist es möglich, dass im Zuge des Abrufs Lernende Verständnisschwierigkeiten oder Wissenslücken bemerken, auf die sie sich in der kommenden Lernphase fokussieren können. Durch diese Feedbackfunktion würde die Wissensorganisation erleichtert und die Entstehung einer kohärenten Wissensbasis gefördert werden (Lipowsky, 2015; Roediger et al., 2011).

Die Ausnutzung der Lerneffektivität bei einem Einsatz von Concept Mapping in Elaborationssettings wurde durch eine von Karpicke und Blunt (2011b) veröffentlichte Studie infrage gestellt, welche die Befunde zur Lernwirksamkeit von kognitiven Abrufprozessen aufgriff und die Lerneffekte des Einsatzes von Concept Mapping mit denen einer *retrieval practice* Bedingung verglich, in der die Studierenden nach dem Lesen des Textes in Form von Notizen so viel niederschreiben sollten, wie sie wiedergeben konnten (*free recall*). Ihre Ergebnisse zeigen, dass Lernende in der *retrieval practice*-Gruppe in allen Tests bessere Ergebnisse erzielte als die Concept Mapping-Gruppe. Blunt und Karpicke (2014) untersuchten daraufhin die Wirksamkeit des Concept Mapping in einer *retrieval practice*-Bedingung, indem sie die Lernenden nach dem Lesen des Textes anwies, Concept Maps mit bzw. ohne Lernmaterial anzufertigen. Anhand der Ergebnisse ihrer beiden Experimente konnten sie zwei Erkenntnisse für die Praxis ableiten: Erstens, dass Concept Mapping auch innerhalb von Retrievalsettings effektiv eingesetzt werden kann, und zweitens, dass beide Lernmethoden (Concept Mapping und Notizen Erstellen) jeweils zu besseren Lernergebnissen führen können, wenn sie in einem solchen Retrievalsetting eingesetzt werden anstatt in einem klassischen Elaborationssetting. Bei einem Kurzantworttest zum deklarativen Wissen (im Sinne von Faktenwissen) fanden die Autor:innen speziell bei den Elaborationssettings einen Vorteil zugunsten von Concept Mapping, aber nicht bei einem Inferenztest (s. auch Kapitel 2.2.2). Die Erkenntnisse dieser Studien zur Lerneffektivität von Concept Mapping in Retrievalsettings griffen O'Day und Karpicke (2020) auf und implementierten ein Studiendesign, in dem die Lernmethode variiert wurde, allerdings nutzten sie eine zusätzliche Gruppe, in der die zwei Methoden kombiniert wurden: In ihren beiden Experimenten sollten die Studierenden kurze

Texte lesen und anschließend die Informationen des Textes durch Notizen (*free recall*), Concept Mapping oder beides nacheinander abrufen und schriftlich festhalten. Es zeigte sich, dass die *free-recall*-Gruppe im Vergleich zu der *Concept-Mapping*-Gruppe und der *Concept-Mapping-und-recall-Gruppe* bessere Testergebnisse erzielte.

Aus didaktischer Perspektive weisen die Studien von Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) jedoch Mängel auf, die zu einer Konfundierung der Ergebnisse geführt haben könnten: Für die schlechteren Ergebnisse der Concept Mapping-Gruppen kann als mögliche Ursache der schon bei Karpicke und Blunt (2011b) kritisierte Mangel an Übung und Vertrautheit mit Concept Mapping auf Seiten der Proband:innen angeführt werden (Mintzes et al., 2011). Die Proband:innen wurden in den Studien von Karpicke und Blunt (2011b), Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) lediglich über die Grundlagen von Concept Mapping informiert, weshalb davon auszugehen ist, dass die Proband:innen nicht ausreichend versiert im Umgang mit der Lernmethode waren, was die Lerneinbußen erklären könnte (s. Kapitel 2.3.1.1). Dafür spricht auch, dass die *Concept-Mapping-und-recall-Gruppe* bei O'Day und Karpicke (2020) schlechtere Ergebnisse erzielte als die Gruppe, die *free recall* ohne Concept Mapping nutzte, und die Concept Mapping-Aufgabe für die Kombinationsgruppe somit eine zusätzliche Belastung darstellte. Darüber hinaus könnte die Kürze der Lerntexte, die in den Studien Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) weniger als 300 Wörter lang waren, eine unrealistisch niedrige Belastung für die Lernenden dargestellt und damit ihre Motivation gesenkt haben, ihr Vorwissen zu vertiefen (vgl. Kapitel 5, 1.2). Aufgrund dieser potenziell ergebnisverzerrenden Charakteristika der Studien von Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) sind die oben genannten Ergebnisse mit Vorsicht zu betrachten.

Die Ergebnisse ändern jedoch nichts an der theoretisch fundierten Vermutung, dass Concept Mapping in Retrievalsettings die Vorteile beider Ansätze kombinieren kann: In der Theorie sollte das Erstellen der Concept Maps die Entstehung von kohärenten Gedächtnisstrukturen erleichtern, da es die organisatorische Repräsentation des Materials durch die bzw. den Lernenden verbessert (Novak & Cañas, 2008). Diese Struktur sollte wiederum Abrufprozesse unterstützen, da ein effektiver Abruf die organisatorische oder relationale Kodierung von Informationen voraussetzt, die in einer Abrufstruktur repräsentiert werden, welche sich durch eine Reihe von erstellten Abrufhilfen (*cues*) auszeichnet, die den Suchprozess während des Abrufs steuern (O'Day & Karpicke, 2020). Die erwarteten Vorteile der Concept Mapping-Gruppen zeigten sich nicht in den genannten Studien, doch ihre Ergebnisse könnten auf die angesprochenen methodischen Ansätze zurückzuführen sein.

Deshalb sind weitere Untersuchungen zum Einsatz von Concept Mapping in Retrievalsettings unter Berücksichtigung der Vertrautheit der Lernenden mit der Lernmethode nötig, um eine auf empirischen Befunden basierte Handreichung für den Einsatz von Concept Mapping in der Praxis diesbezüglich geben zu können. Dies könnte durch eine Variation der Faktoren *Setting* und *Training* erreicht werden, wobei zu erwarten ist, dass die Verfügbarkeit des Textes und die unterschiedlichen Trainingsansätze sich auf die Förderung (meta)kognitiver Prozesse und Reduzierung der lernhinderlichen kognitiven Belastung sowie die Lernleistung auswirken (s. auch Nückles et al., 2010). Den ersten Teil dieses Desiderats greifen auch O’Day und Karpicke (2020) auf und monieren, dass Studien, die *retrieval practice* und elaborative Lernstrategien kombinieren, beklagenswert rar sind. O’Day und Karpicke (2020) konzentrierten sich jedoch auf die Variation der Lernmethode und nicht auf das Training, weshalb anhand ihrer Ergebnisse keine auf die Praxis zweifelsfrei übertragbaren Empfehlungen zum Einsatz von Concept Mapping abgeleitet werden können, da die Bedeutung von Trainingsmaßnahmen u. a. für Lernleistungen in zahlreichen Studien und Metaanalysen (s. Kapitel 2.3.1) betont wird.

Das zweite Ziel des Forschungsprojektes ist es daher, zum einen die Frage zu beantworten, wie effektiv Concept Mapping ist, wenn den Lernenden bei der Erstellung ihrer Concept Maps das Lernmaterial nicht zur Verfügung steht (Retrievalsetting), im Gegensatz zum klassischen Elaborationssetting, bei dem das Lernmaterial verfügbar ist. Zum anderen soll herausgefunden werden, ob mögliche Effekte der beiden Settings von der Intensität eines vorangegangenen Trainings in der Anwendung von Concept Mapping abhängen.

3 Forschungsfragen und Hypothesen

Für einen effektiven und lernwirksamen Einsatz des Concept Mappings sind anhand der hier vorgestellten bisherigen Befunde keine klaren Empfehlungen bezüglich des Trainings (s. Kapitel 2.3.1) und des Einsatzes von Concept Mapping in Elaborations- und Retrievalsettings (s. Kapitel 2.3.2) ableitbar. Daraus ergeben sich Desiderata zur Bedeutung des Trainings und Settings, die in zwei Studien erforscht wurden. Die resultierenden Forschungsfragen und Hypothesen zu den beiden im Rahmen dieses Projektes durchgeführten Studien werden in diesem Kapitel vorgestellt. In der ersten Studie stand der Faktor *Training* im Fokus, in der zweiten Studie eine Kombination der Faktoren *Training* und *Setting*. Nach der Vorstellung der Forschungsfragen und Hypothesen wird die erste Studie in einer Kurzbeschreibung dargestellt, wonach die Originalpublikation in Englisch des dazugehörigen Manuskripts folgt (s. Kapitel 4). Im Sinne des konsistenten Aufbaus wird mit der zweiten Studie

ebenso verfahren: In Kapitel 5 wird die zweite Studie kurz beschrieben, worauf das dazugehörige Manuskript folgt.

3.1 Studie 1

Basierend auf der in Kapitel 2.3.1 vorgestellten uneinheitlichen Empfehlungen bezüglich der Implementierung von Trainingsmaßnahmen im Concept Mapping und damit assoziierter uneinheitlicher Lernleistungsergebnisse ist es das Ziel der ersten Studie, die Frage zu beantworten, ob Lernende, die ein spezifisches Concept Mapping-Training erhalten haben, generell erfolgreicher im Lernen sind, wenn sie Concept Maps erstellen, als solche, die auf andere Weise an Concept Mapping herangeführt wurden.

Für dieses spezifische Concept Mapping-Training wurden folgende Überlegungen zur Entscheidung über die Trainingsparameter angestellt (vgl. Kapitel 4, 1.3): Bezüglich der Trainingsdauer wird davon ausgegangen, dass ein mehrwöchiges Training mit wöchentlichen Sitzungen à 90 Minuten angemessen ist, um den Lernenden das nötige deklarative und prozedurale Wissen zu vermitteln, damit sie eine komplexe Lernmethode wie das Concept Mapping lernwirksam einsetzen können und u. a. ihr Textverständnis und die Wissensanwendung gefördert werden (Renkl & Nückles, 2006). Von einem solchen mehrwöchigen Ansatz und damit assoziierten positiven Effekten auf die Lernleistung berichten sowohl Einzelstudien (z. B. Ajaja, 2011; Kinchin, 2000a; Romero et al., 2017) als auch Metaanalysen (Nesbit & Adesope, 2006; Schroeder et al., 2018). Bezogen auf die Instruktionmethode, mit der die Vermittlung deklarativen und prozeduralen Wissens stattfinden soll, kann davon ausgegangen werden, dass praktische Übungsphasen mehr als rein theoretisches Input den Lernenden dabei helfen, mit dem Concept Mapping vertraut zu werden und prozedurales Wissen zu erwerben (s. Kapitel 2.1.2 und 2.1.5). Darüber hinaus hilft es den Lehrenden, da die Verlässlichkeit der Ergebnisse, die anhand der Concept Maps gewonnen werden sollen, gesteigert wird (Brandstädter et al., 2012; Jüngst & Strittmatter, 1995). Ein weiterer Vorteil wird in der schrittweisen Heranführung und der Bereitstellung von Lernhilfen gesehen, die sowohl das Ausprobieren und Einüben erleichtern als auch den Nutzen und die Vorteile der Lernstrategie den Lernenden verdeutlichen soll (Stark et al., 2008). Das Erkennen der Vorteile sollte Lernende wiederum ebenfalls motivieren, die Methode auch in späteren Situationen selbstständig und selbstreguliert anzuwenden (Brandstädter et al., 2012; Landmann & Schmitz, 2007; Leopold et al., 2006).

Basierend auf den Erkenntnissen der in Kapitel 2.3.1 vorgestellten Studien (z. B. Ajaja, 2011; Jegede et al., 1990; Jüngst & Strittmatter, 1995; Michalak & Müller, 2017; Mintzes et al., 2011; Nesbit & Adesope, 2006; Quinn et al., 2003; Schroeder et al., 2018; Sumfleth et al., 2010) wird erwartet, dass ein entsprechendes Concept Mapping-Training eine Vertrautheit mit der Methode bewirkt. Bezüglich des lernleistungsbezogenen Einflusses, den der Grad an Vertrautheit mit Concept Mapping auf kognitive Prozesse, die kognitive Belastung und die Concept Mapping-bezogenen Fertigkeiten hat, werden folgende drei Hypothesen auf Grundlage bisheriger Befunde abgeleitet (vgl. Kapitel 4, 1.4):

- H1.1 Bezüglich der für das Concept Mapping postulierten Effekte auf kognitive Prozesse (z. B. Mayer, 2002a; Renkl & Nückles, 2006) wird erwartet, dass das oben beschriebene extensive Training zu mehr Organisations- und Elaborationsprozessen während des Lernens im Sinne der Informationsintegration in vorherige Wissensstrukturen und damit verbundener Wissenskonsolidierung führt.
- H1.2 In Anlehnung an die vorgestellten Befunde (s. Kapitel 2.1.5.2 und 2.3) wird des Weiteren erwartet, dass insbesondere ein solches extensive Training dazu führt, dass weniger kognitive Ressourcen für die Prozesse, die das Concept Mapping erfordert, beansprucht werden und die für das Concept Mapping reklamierte Entlastung des Arbeitsgedächtnisses ermöglicht wird (Jonassen, 2003; Nesbit & Adesope, 2006; O'Donnell et al., 2002; Renkl & Nückles, 2006). Somit geht ein solches Training mit einer geringeren wahrgenommenen kognitiven Belastung beim Einsatz von Concept Mapping während des Lernens einher.
- H1.3 Schließlich wird, basierend auf den in Kapitel 2.3.1 thematisierten Studienberichten zu Trainingsmaßnahmen (Ajaja, 2011; Hilbert & Renkl, 2008; Jegede et al., 1990; Mintzes et al., 2011; Morse & Jutras, 2008; Sumfleth et al., 2010), erwartet, dass insbesondere ein extensives, wie oben beschriebenes Training zu einer besseren Handhabung und Bearbeitung vorgegebener bzw. Erstellung eigener Concept Maps durch die Lernenden führt.

3.2 Studie 2

In der zweiten Studie wurde die Kombination zweier Faktoren berücksichtigt, die einen potenziellen Einfluss auf den Lernerfolg beim Concept Mapping haben könnten: Zum einen der im Rahmen der ersten Studie bereits berichtete Faktor *Training* und zum anderen die unterschiedlichen Auswirkungen der Verfügbarkeit des Lernmaterials während des Concept Mappings auf die Lernleistung. Während für das Concept Mapping sowie für das Praktizieren

des Abrufs ein positiver Einfluss auf die Lernleistung postuliert wird, zeigte sich auch in diversen Studien (s. Kapitel 2.3.1), dass erst ein gewisser Grad an Vertrautheit mit Concept Mapping dazu führt, dass die Strategie lernwirksam eingesetzt werden kann. Dessen uneingedenk nutzten Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) Kurzeinführungen in das Concept Mapping und kombinierten anschließend Concept Mapping mit einem Retrievalsetting, was in den Erkenntnissen resultierte, dass Concept Mapping zwar in einem solchen Setting eingesetzt werden kann, die Lernleistung dieser Concept Mapping-Gruppen jedoch hinter dem jener Gruppen lag, die das Erinnerte in Form von Notizen festhielten. Angesichts dieser Ergebnislage wurde ein Design entwickelt, um bezüglich des Einflusses der beiden Faktoren auf den Lernerfolg mehr Klarheit zu schaffen. Das Design basiert auf vier Untersuchungsgruppen, in denen das vorgeschaltete Training (extensiv vs. Kurzeinführung) sowie die Textverfügbarkeit (verfügbar vs. nicht verfügbar) während des Concept Mappings variiert werden, und mithilfe derer versucht wird, folgende Fragen zu beantworten (vgl. Kapitel 5, 1.4):

- (1) Wie effektiv kann Concept Mapping das Lernen unterstützen, wenn den Lernenden der Lerntext während der Erstellung ihrer Concept Maps nicht zur Verfügung steht (Retrievalsetting), im Gegensatz zu dem „klassischen“ Setting, in dem sie auf das Lernmaterial zurückgreifen können (Elaborationssetting)?
- (2) Bedingt die Intensität eines vorgeschalteten methodischen Trainings in der Anwendung von Concept Mapping mögliche Effekte der beiden Settings?

Auf den bisherigen Befunden zu den Effekten der Textverfügbarkeit und unterschiedlicher Trainingsmaßnahmen basierend (s. Kapitel 2.3.1 und 2.3.2), werden die folgenden drei Hypothesen zur Beantwortung der beiden Fragen aufgestellt (vgl. Kapitel 5, 1.4):

H2.1 Basierend auf den Erkenntnissen von Blunt und Karpicke (2014) wird erwartet, dass die Verfügbarkeit des Lernmaterials entscheidend für die kognitiven Prozesse beim Concept Mapping ist. Die Textverfügbarkeit sollte in diesem Zusammenhang insofern einen Einfluss auf kognitive Prozesse haben, als die Abwesenheit eines Lerntextes gedächtnisbezogene Abrufprozesse während des Einsatzes von Concept Mapping induziert. Diese Abrufprozesse binden jedoch kognitive Ressourcen, die normalerweise für die weitere Elaboration benötigt werden. Steht den Personen der Lerntext jedoch während der Concept Map Erstellung zur Verfügung, sollten vermehrt Elaborationsprozesse stattfinden. Dies sollte sich in mehr Propositionen, in denen die

Lernenden ihr Vorwissen integriert haben, manifestieren. Aufgrund der Lernvorteile, die mit solchen kognitiven Elaborationsprozessen assoziiert sind (s. Kapitel 2.1.4), sollten Lernende, die in ihren Concept Maps mehr Elaborationspropositionen erstellen, bessere Ergebnisse in den anschließenden Wissenstests zu den Lerninhalten erzielen.

H2.2 In Anlehnung an die für Studie 1 aufgestellten Hypothese, dass ein extensives Training zu einer besseren Handhabung und Bearbeitung vorgegebener bzw. Erstellung eigener Concept Maps führt (H1.3), wird auch für Studie 2 erwartet, dass ein extensives Training in Concept Mapping bei den Lernenden eine gewisse Vertrautheit mit der Lernstrategie, und insbesondere mit ihren methodischen Anforderungen, induziert. Diese Vertrautheit wiederum ist für einen erfolgreichen Einsatz der Lernmethode hilfreich. Ein solcher Effekt sollte insbesondere bei einem Vergleich mit den Gruppen, die lediglich eine Kurzeinführung in das Concept Mapping erhalten, deutlich werden. Basierend auf dieser Annahme wird erwartet, dass diejenigen Personen, die an einem vorgeschalteten extensiven Concept Mapping-Training teilgenommen haben, das sowohl theoretisches Input als auch praktische Übungsphasen und Feedback zu den Concept Maps beinhaltet, eine höhere Concept Mapping-bezogene Selbstwirksamkeitserwartung berichten, da sie sich vertrauter mit der Lernmethode fühlen. Darüber hinaus sollten sie einerseits in der Analyse und Bearbeitung vorgegebener Concept Maps bessere Ergebnisse erzielen und andererseits weniger methodische Fehler bei der Erstellung eigener Concept Maps machen.

H2.3 In Anbetracht der Ergebnisse aus Studie 1 (s. Kapitel 4) wird, ergänzend zu den Hypothesen H2.1 und H2.2, erwartet, dass ein umfangreiches Concept Mapping-Training kombiniert mit der Verfügbarkeit des Lernmaterials während der Concept Map Erstellung die Concept Mapping- und wissensbezogenen Lernergebnisse am stärksten verbessert. Die entsprechende Wahrnehmung des eigenen Lernerfolgs während des Lernens sollte sich auch im Anschluss bei der damit verbundenen metakognitiven Vorhersage der Lernenden bezüglich ihrer erwarteten Lernleistung widerspiegeln, die höher ausfallen sollte als bei den anderen Gruppen (s. auch Großschedl & Tröbst, 2018).

4 Studie 1: Untersuchung unterschiedlicher Trainingsansätze

In der ersten Studie sollte herausgefunden werden, ob die Lernleistung beim Concept Mapping von der Art und Intensität des Trainings abhängt und das Training somit einen entscheidenden Einflussfaktor für den Wissenserwerb darstellt. Im Rahmen des von Großschedl und Tröbst (2017) beantragten DFG-Projektes (GR 47632-1) führte der Autor der vorliegenden Dissertation die Studie durch. Er war anteilig an der Konzeption und Methodologie der Studie, überwiegend an der Datenverwaltung, wesentlich an der Datenauswertung, anteilig an den Visualisierungen und wesentlich an dem Verfassen, Einreichen und Überarbeiten des Manuskripts beteiligt (s. auch Kapitel 4, *Author Contributions*). Die Inhalte sowie Abbildungen und Tabellen dieses Kapitels (mit Ausnahme von Tabelle I) stammen aus der Originalarbeit von Becker, Welter, Aschermann und Großschedl (2021)⁷, welche für die vorliegende Dissertation zur Verfügung gestellt wurden, so dass in der folgenden Kurzbeschreibung auf einen fortlaufenden Quellenverweis im Interesse der Lesbarkeit verzichtet wird.

Zur Überprüfung der in Kapitel 3.1 vorgestellten Hypothesen wurde zur Feststellung von Unterschieden in den abhängigen Variablen das Training als unabhängige Variable (UV) gewählt. Hierfür wurden drei unterschiedliche Versuchsbedingungen durch zwei Experimentalgruppen und eine Kontrollgruppe implementiert, die sich in Art und Intensität des erhaltenen Concept Mapping-Trainings unterschieden. Die Datengrundlage der statistischen Analysen beruht auf einer Stichprobe von $N = 73$ Proband:innen, bei denen es sich um Studierende aus verschiedenen Fachrichtungen handelte. Von diesen Proband:innen waren 56 % in einem naturwissenschaftlichen Studiengang eingeschrieben, während die anderen hauptsächlich ein geisteswissenschaftliches Fach studierten. Die Proband:innen waren im Durchschnitt 22,6 Jahre alt und 78 % waren weiblich. Die erste Gruppe (T++, $n = 27$) erhielt ein Concept Mapping-Training mit Scaffolding und Feedback, die zweite Gruppe (T+, $n = 21$) erhielt ein Concept Mapping-Training mit wiederholtem Üben ohne Scaffolding und Feedback, und die dritte Gruppe (T-, $n = 25$) erhielt kein spezielles Concept Mapping-Training, sondern eine kurze verbale Einweisung in das Concept Mapping. Nachfolgend wird in einer kurzen Zusammenfassung über das Studiendesign, die Erhebungsinstrumente und eingesetzte Materialien, die Stichprobe sowie die Durchführung unserer Studie informiert.

⁷ An dieser Stelle wird erneut darauf hingewiesen, dass das Manuskript der Originalarbeit von Becker, Welter, Aschermann und Großschedl (2021) in Kapitel 4 einen in sich geschlossenen Abschnitt mit einer eigenen Nummerierung der Kapitel, Abbildungen und Tabellen sowie eigenem Literaturverzeichnis darstellt. Verweise in der übrigen Arbeit auf einzelne Kapitel, Abbildungen und Tabellen des Manuskripts werden mit einem Verweis auf das Kapitel 4 eingeleitet, worauf die Nummer des Unterkapitels (z. B. „s. Kapitel 4, 3.2“) bzw. der Verweis auf die Abbildung oder Tabelle folgt (z. B. „s. Kapitel 4, Tab. 4“).

Die quasi-experimentelle Studie wurde über einen Zeitraum von sechs Wochen durchgeführt, die in drei aufeinanderfolgende Phasen (Trainings-, Lern- und Testphase) unterteilt wurden und in wöchentliche Sitzungen aufgeteilt waren: drei Trainingssitzungen, eine Lernsitzung und zwei Testsitzungen (s. Tabelle I). Zur Verhinderung von Testeffekten wurde für die Trainingsphase das nicht-biologische Thema Intelligenz, und für die Lern- und spätere Testphase das Thema Zellbiologie gewählt.

Tabelle I. Übersicht über Phasen und Gruppen der Studie 1 (verändert nach Großschedl & Tröbst, 2017)

Gr.	Trainingsphase (Lerninhalt: Intelligenz)			Lernphase (Lerninhalt: Zellbiologie)		Testphase	
	Wo. 1	Wo. 2	Wo. 3	Woche 4		Woche 5	Woche 6
	Training			Lernaktivität	Abhängige Variablen	Abhängige Variablen	
T++	Vorbefragung	Scaffolding und Feedback	Zwischenerhebung (Wissen über Lerninhalte der Trainingsphase)	Kurzeinführung , CM-Selbstwirksamkeit, CM-Treatment-Check	Concept Mapping (mit Lernmaterial) 1. während der Lernaktivität mit dem Text erstellte Concept Maps (Begriffe frei wählbar) 2. Kognitive Belastung (7 Items) 3. Metakognitive Vorhersage (1 Item)	1. Concept Maps (mit Begriffssatz, bestehend aus 22 Konzepten) 2. Strukturelles Wissen (55 Items) 3. Deklaratives Wissen (30 Items)	4. Konzeptuelles Wissen (15 Items)
T+		Wiederholtes Üben					
T-		Unspezifisches Kontrolltraining					

Anmerkungen. Gr. = Gruppe; CM = Concept Mapping; T++ = CM-Trainingsgruppe ($n = 27$), die Scaffolding und Feedback erhält; T+ = CM-Trainingsgruppe ($n = 21$), die CM wiederholt übte, aber kein Scaffolding und Feedback erhält; T- = Kontrolltrainingsgruppe ($n = 25$), die während der Trainingsphase kein CM-Training erhält, sondern andere Lernmethoden (z. B. Kleingruppendiskussionen in Woche 1) nutzt.

Um die Lernbedingungen für die Proband:innen so realitätsnah wie möglich an ihren üblichen Studienalltag anzupassen und die Kritik der mangelnden externen Validität an den Bedingungen in den Studien von Karpicke und Blunt (2011b), Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) aufzugreifen, wurde den Proband:innen in der Lernphase ein Text zur Zellbiologie vorgelegt, der ca. 2.000 Wörter sowie Abbildungen umfasste. Im Speziellen ging es in dem zu lesenden Text um Aufbau und Funktionen von eukaryotischen Zellen. Durch den Umfang des Lernmaterials sollten die Anforderungen der universitären Ausbildung im Sinne der Bewältigung von komplexen Texten angemessen repräsentiert werden (vgl. Assaraf et al., 2013; Hmelo-Silver & Azevedo, 2006; Tripto et al., 2017).

Als potenzielle Kontrollvariablen wurden im Rahmen einer Vorbefragung (Woche 1) demographische Daten sowie Vorerfahrungen mit Concept Mapping, das Vorwissen bezüglich der zellbiologischen Lerninhalte und die Lesekompetenz der Proband:innen erhoben. Die abhängige Variable in der Trainingsphase stellte das erworbene Wissen bezüglich des Lernthemas Intelligenz dar (Zwischenerhebung in Woche 3). In der Lernphase (Woche 4) wurden nach der Kurzeinführung und vor dem Lesen der Lerntexte zwei weitere Variablen erhoben: zum einen die bezüglich des Concept Mappings empfundene Selbstwirksamkeit, zum anderen die objektivere Messung der Concept Mapping-Fertigkeiten mithilfe einer vorgegebenen Concept Map (vgl. H1.3), deren integrierte Fehler von den Proband:innen erkannt und korrigiert werden sollten (CM-Treatment-Check). Des Weiteren dienten als abhängige Variablen die mithilfe der Concept Maps operationalisierten kognitiven Prozesse (Abruf, Organisation, Elaboration; vgl. H1.1) sowie die Concept Map Qualität (vgl. H1.3). Letztere sollte sowohl Aufschluss darüber geben, wie ausgeprägt die Concept Mapping-Fertigkeiten der Proband:innen nach Erhalt der unterschiedlichen Trainings waren, als auch darüber, wie viele inhaltlich falsche Propositionen erstellt wurden, die Rückschlüsse auf das Verständnis bzw. Missverständnis der Proband:innen zulassen. Nach der Erstellung der Concept Maps wurde die kognitive Belastung erhoben und die Proband:innen machten eine metakognitive Vorhersage bezüglich der eigenen Behaltensleistung (vgl. H1.2).

In der ersten Woche der Testphase (Woche 5) wurden Concept Mapping-bezogene Fertigkeiten durch eine weitere Erstellung einer eigenen Concept Map zum Lerninhalt der Lernphase erhoben (vgl. H1.3). Ferner wurden das strukturelle Wissen der Proband:innen zum Lerninhalt der Lernphase mit einem Similarity Judgments Test (55 Items) und das deklarative Wissen mit einem Multiple Choice Test (30 Items) erhoben. In der zweiten Woche der Testphase (Woche 6) wurde schließlich die letzte abhängige Variablen (konzeptuelles Wissen) mithilfe eines Tests im offenen Antwortformat (15 Items) erhoben.

**Comprehension-Oriented Learning of Cell Biology: Do Different Training Conditions
Affect Students' Learning Success Differentially?**

Becker, L. B., Welter, V. D. E., Aschermann, E., & Großschedl, J. (2021). Comprehension-oriented learning of cell biology: Do different training conditions affect students' learning success differentially? *Education Sciences*, 11, 438. doi: 10.3390/educsci11080438

Abstract

Concept Mapping (CM) is a learning strategy to organize and understand complex relationships, which are particularly characteristic of the natural science subjects. Previous research has already shown that constructing concept maps can promote students' meaningful learning in terms of deeper knowledge and its more flexible use. While researchers generally agree that students need to practice using CM successfully for learning, key parameters of effective CM training (e.g., content, structure, and duration) remain controversial. This desideratum is taken up by our study, in which three different training approaches were evaluated: a CM training with scaffolding and feedback vs. a CM training without additional elements vs. a non-CM control training. In a quasi-experimental design, we assessed the learning outcome of $N = 73$ university students who each had participated in one of the trainings before. Our results suggest that an extensive CM training with scaffolding and feedback is most appropriate to promote both CM competence and acquisition of knowledge. From an educational perspective, it would therefore be advisable to accept the time-consuming process of intensive practice of CM in order to enable students to adequately use of the strategy and thus facilitate meaningful learning in terms of achieving sustained learning success.

Keywords: concept mapping; learning strategy; training; meaningful learning; knowledge acquisition

1. Introduction

Natural science subjects, such as physics or chemistry, primarily deal with highly abstract topics including various complex relationships [1–7]. In a similar manner, the subject of biology demands the learners' conceptualization considering complex relationships [8–15]. As particular attention is paid to inter- and intra-systemic relationships to describe, explain, and predict interdependencies, for example between different levels of biological systems [16,17], a corresponding challenge for learners is to know, understand, and handle these sometimes highly abstract concepts. In contrast to mere declarative (factual) knowledge, this knowledge about the relations between different topics is called conceptual knowledge. However, empirical studies point out profound difficulties for learners regarding knowledge acquisition [6,14,15,18–20]. In particular, the demanding cognitive processes of organization and elaboration (i.e., an integration of new information into structures of prior knowledge) which are necessary when learning requires conceptual reasoning, seem to pose a particular challenge to learners [8,21–31].

Cognitive processes of organization and elaboration can be fostered by respective learning strategies [32]. Organization strategies can be used to condense and structure new information regarding its semantic composition, for example by summarizing texts or constructing concept maps, both requiring in-depth processing of learning material. Elaboration strategies, such as activating prior knowledge, taking notes, asking questions, or developing related ideas, are intended to foster semantically oriented understanding (in-depth strategies) on the one hand, as well as permanent retention via cueing (surface strategies) on the other hand. By using organization and elaboration learning strategies, new information can be meaningfully integrated into prior knowledge.

Previous empirical findings suggest that concept mapping (CM) effectively supports organization and elaboration processes as well as conceptual reasoning, and thus, can facilitate a deeper understanding and more organized knowledge acquisition [33–38]. CM can be categorized as an in-depth cognitive learning strategy that allows for analysis and reconstruction of the significance of text-based learning material by extracting its central concepts/terms and their interrelations. The result is represented by a combination of graphic and linguistic elements, the concept maps, that can be construed as structural and semantic extensions of mind maps which tend to be more associative [10,14,34,36,37,39–49]. Concept maps form diagrams, in which concepts are represented by “nodes” that are connected by arrows to indicate their interrelation. To indicate the exact nature of this interrelation, the arrows are labeled to determine the reading direction. Two linked concepts form a so-called proposition, the smallest meaningful unit of a concept map [12,47,50–53] (see Figure 1).

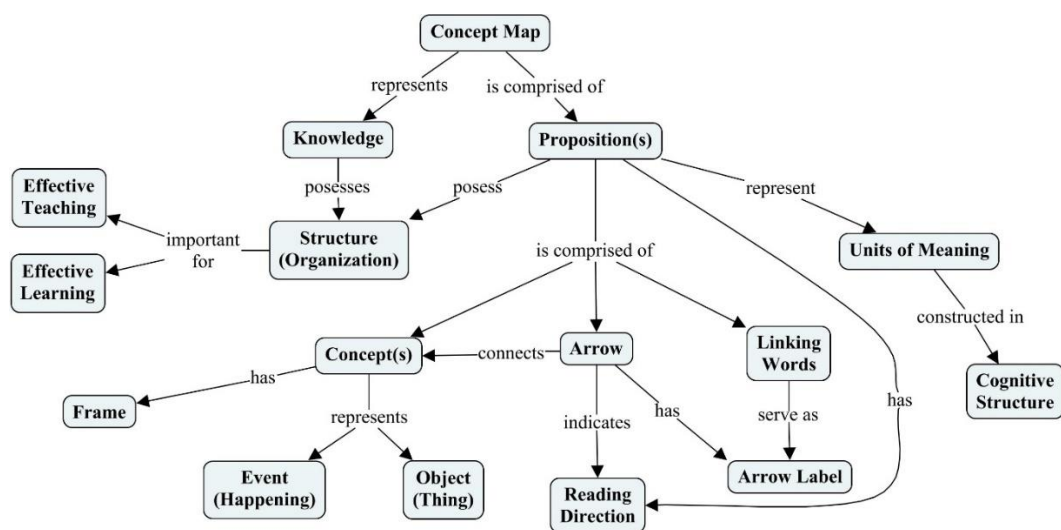


Figure 1. A concept map representing principles of the CM learning strategy.

Due to their characteristics' potential regarding comprehension-oriented learning, concept maps have received a lot of attention from educators since their invention [38,46,47,54,55]. Educational books of biology for school and university level have already started to add CM to their list of recommended learning strategies [56,57]. While teachers are trying to adopt the practice, however, some of them report that integrating CM in classrooms poses certain challenges, since implementation of CM seems quite time-consuming because students are often struggling with selecting concepts and links [58,59]. Accordingly, the main problem for learner seems to be that operating with an unfamiliar strategy for a longer time binds cognitive capacity for this strategy's monitoring although this capacity is actually required for knowledge acquisition [32]. Consequently, beginners often lack the skills to use CM effectively, and they are unable to tap its full potential [60,61]. This makes the importance of appropriate CM training an essential aspect of the strategy's effectiveness. To this, Hilbert and Renkl [60] note that "an ideal training method for students on how to use concept mapping that is effective, efficient, and directed to the typical needs of beginners in mapping, is missing at present" (p. 55). More recently, Roessger et al. [62] have pointed out that "few researchers have broached the best way to teach concept mapping" (p. 12). It is therefore hardly surprising that results from studies on the strategy's effectiveness can differ since CM has been used in different subject areas (STEM vs. non-STEM), with different goals (as a learning vs. a diagnostic tool), in different ways (studying vs. creating concept maps), and in different settings (individual vs. collaborative CM exercise) [38].

If researchers are unable to determine effective CM training conditions, even though training represents one of the strategy's key factors for its effectiveness, different results in CM research could be also attributed to inconsistent training conditions. Therefore, our study tries to determine the effectiveness of different CM training conditions by focusing on what makes CM effective to facilitate meaningful learning in terms of achieving sustained learning success.

1.1. *Concept Mapping from a Cognitive and Metacognitive Perspective*

Ausubel [63,64] linked effective processes and products of meaningful learning to certain requirements: learners should proactively seek an integration of newly learned information into their existing networks of prior knowledge and establish corresponding relationships to it, for example by using learning strategies of elaboration and organization (see Section 1). In addition, prior knowledge itself should be well organized and consolidated. Finally, the learning material should be as clearly and comprehensibly designed as possible, e.g., by including explicit indications of potential relational connections to other elements within a respective field of knowledge [50,65].

On the cognitive level, the linking of terms/concepts using CM facilitates the integration of new information into prior knowledge structures in terms of elaboration. Furthermore, CM helps the learner in terms of organization by providing a frame that can be used to structure new information. This way, all information is not only reduced to its essentials but also represented in a way that facilitates processing according to semantic network models [66–70]. Cox [71] and other researchers [72] attribute these strategies' benefits to activation of self-explanatory activities during the learning process. Consequently, the externalization of knowledge by using concept maps allows the recipient to simply read the information and access it more easily, since concept map-bound information does not have to be continually kept in mind [71]. This externalization of the assumed cognitive processes and structures of representation is also associated with the development of coherent mental knowledge structures [73] and the relief of the strain on working memory [47,74,75]. Beyond this simplification function, the concept map construction itself is expected to result in a more in-depth analysis of the learning material which, in turn, predicts better and long-lasting learning outcome [50,60,67,76,77,78]. Finally, regarding the metacognitive level, it is assumed that CM is associated with more efficient workflow monitoring skills and better identification of difficulties in understanding or knowledge gaps [32,52,60,67,79].

Concept maps, therefore, represent reliable and valid measures of how well learners organize declarative knowledge in a domain [80]. Correspondingly, certain features of a concept map can be used to draw conclusions about associated cognitive processes during learning from texts using CM. Especially the complexity of propositions specified during construction indicates (more or less) distinguishable cognitive processes: (1) recall-suggesting (R-) propositions are represented by relations that are explicitly mentioned in the learning material, (2) organization-suggesting (O-) propositions are represented by relations, that are not explicitly mentioned in the learning material, but constructed, and (3) elaboration-suggesting (E-) propositions are represented by relations between concepts of which at least one was not explicitly mentioned in the learning material but derived from the learners' prior knowledge.

1.2. Concept Mapping, Prior Knowledge, and Cognitive Load

The structural model of cognition from McClelland et al. [81] considers relevant (prior) knowledge activated during information processing as a central element of information integration. However, this prior knowledge can contradict the knowledge to be acquired, i.e., learners' prior knowledge inhibits learning, as new information cannot be meaningfully integrated into the existing network [82]. For educators, this means that the influence of

learners' prior knowledge on learning activities must be considered in terms of activating corresponding processes as precisely as possible when dealing with new tasks [83–88].

With particular regard to the importance of (structured and activated) prior knowledge for the efficacy of learning aids and learning strategies, Hasselhorn and Gold [75] point out the reverse U-functional relationship between these two variables: if the learners' prior knowledge ranges on a medium level, they can benefit from provided learning aids best, whereas higher or lower levels of prior knowledge are associated with lower profit. If teachers fail to adapt learning strategies to the learners' levels of competence, experience, and prior knowledge, their learning success decreases, even in the case of usually high-performance students. These high achievers already use learning strategies that work well for them. If they are asked to replace them by using another learning strategy offering no additional profit, they will usually show a comparatively poor learning outcome. This phenomenon has become known as the *expertise reversal effect* [89–91] and could also be an example of *aptitude treatment interaction* [90,92,93]. A detailed explanation for this decline in learning success is based on Sweller's *Cognitive Load Theory* (CLT) [94–97], which highlights the importance of working memory capacity during learning activities. Regarding the expertise reversal effect, it can be assumed that the learners' cognitive capacity is not exclusively bound by the processing of new information but rather by coping with the unknown methodical procedure when using an unfamiliar learning strategy.

1.3. Training in Concept Mapping

If learners are introduced to a new learning strategy, those without supportive prompting often fail to actively and specifically add this new strategy to their cognitive learning strategy repertoire (production deficit) or fail to gain an advantage from using it (deficit in use). The same is reported for metacognitive knowledge about learning strategies and corresponding skills for monitoring the own learning process while using a strategy [75,98]. In this respect, Cañas et al. [34] critically point out that unfamiliarity with CM may cause an unnecessary burden on a learner's cognitive load which makes successful learning impossible. Michalak and Müller [44] found that persons who lacked CM experience accordingly stated to be unfamiliar with a holistic view of a graphic and the authors concluded that methodical guidance is necessary. This is not a novel idea, as Jüngst and Strittmatter [99] have stated that the adequate use of CM requires sufficient knowledge about the strategy and its use which is why they proposed that only after 10 to 15 self-constructed maps, learners are sufficiently familiar with CM as a learning tool. Accordingly, several weeks of training with repeated intervention and

feedback measures were implemented in some studies [100,101]. In contrast, authors such as Jonassen et al. [102] state that “concept mapping is relatively easy to learn” (p. 162).

Overall, the construction of concept maps requires procedural knowledge which is defined as the “know how” (meaning some form of action carried out), whereas declarative knowledge is defined as “know that” [103] (p. 219), including conceptual knowledge regarding relationships between respective concepts [104]. Neglecting more detailed aspects of distinction, the repetition of procedural knowledge has been found to have significant learning effects: Repeating an action can even make the execution itself unnecessary, as the final result can be recalled as declarative knowledge [105]. For example, if a student had to use an instruction (e.g., step-by-step explanation from a handout) on how to construct a proposition in a concept map, by repetition, he or she would sooner or later know the answer without this instruction. In addition, special series of processes can be generated through repetition of actions and routines can be established. These routines put less strain on attentional processes and can be executed faster with time. The same is true for skills such as reading: if they have been practiced for an extended period, retrieval from long-term memory is accomplished with very little effort, which is referred to as automaticity [106], and a relief of working memory can be observed [107].

For these reasons, many researchers mention an importance of CM training before its use [60,62,66,68,69,99,108–113]. However, for such training measures to have a positive effect on learning performance, they should be carried out long enough to ensure that using the new strategy becomes sufficiently routine. In this respect, questions arise as to (1) what content should be trained (2) in which amount of time, and (3) in which intensity using (4) which specific methods? Most training approaches reported in literature concerning CM are adapted from Novak and Cañas [47]. However, most of these studies only touch on CM training since they predominantly focus on the effectiveness of the learning tool itself. Therefore, consistent recommendations concerning specific training designs could not be deduced yet. Consequently, within the empirical findings concerning the training design, central parameters vary widely and are unsystematically associated with different learning outcomes [110,114–116]. For example, these varying parameters refer to:

(a) Training duration (10 min to three hours) [18,108,117–119];

(b) Training intensity (once or several times, sometimes over several weeks) [82,120];

(c) Instruction method (direct training to acquire declarative and procedural knowledge vs. indirect training to acquire exclusively declarative knowledge) [18,21,60,69,121]; as well as

(d) Self-activity of the learners (level of support, such as scaffolding in instructions and task material) [10,36,66,67,113,122–127].

Considering the findings of Hay et al. [117], who stated that CM can be taught in a single session within 10–20 min, and that “most students will find another 20–30 min sufficient to construct a reasonable map” (p. 302), Brandstädter et al. [22] used a comparable timeframe, but implemented a training approach, that included both theory- and practice-oriented elements from the beginning. Additionally, they varied the degree of directedness (highly directed vs. non-directed) during the course of the two training sessions. Based on their results, the authors recommend using a highly directed approach, including practical elements, as the validity for an assessment of students’ conceptual reasoning using CM was positively influenced by this approach. Sumfleth et al. [69] also based their design on previous studies examining the benefits of active practice of concept map construction [36,67]. The authors developed a CM training booklet for grade 10 chemistry classes and compared a CM training group to a control group that received only a short introduction to CM. The CM training consisted of three sessions (introduction, demonstration of the CM strategy, and exercises), in which students were given 60 min to complete the booklet. The control group read the same texts but did not use CM. Results show that learners in the trained group were able to generate significantly better concept maps than the untrained learners. In contrast, Jegede et al. [110] also trained students in CM over a three-week period, including theoretical and practical phases, but focused specifically on content-related learning outcome. Their results show that their CM training led to significantly better test results in biology than a traditional teaching strategy. Regarding additional training measures (e.g., prompting), Hilbert et al. [66] compared a CM training group that received prompts to a second training group that did not. They found that participants who received prompts during CM showed a better learning outcome compared to the other group. Results of a follow-up test, however, indicated that learners, who were provided with prompts once, were not able to add them to their skill repertoire and continue using CM on their own. Another study of Pearsall et al. [120] focused on the effect of additional feedback during CM training. They trained students on CM in a single 3 h session but provided students with individual feedback on some concept maps which had to be created during their regular semester of study that followed. Analyses of the students’ concept maps showed that the students were obviously able to use the continuous feedback to construct concepts maps

indicating more and more elaborated and well-differentiated knowledge structures over the course of time.

Based on a synopsis of these empirical findings, we assume that a training period of several weeks with weekly training sessions of 90 min seems appropriate to the strategy's complexity to ensure a minimum level of familiarity, and thus, the opportunity for the successful acquisition of procedural knowledge. This is also in line with findings from Ajaja [128] and Kinchin [58], who showed that greater familiarity with CM led to significant and consistent improvement in achievement. In agreement with Brandstädter et al. [22] who specifically address the method of instruction, we further assume that practical exercises in CM will be superior to purely theoretical teaching in terms of learning success. Finally, considering the learners' activity and the degree of support, we assume that only guided training including step-by-step practicing makes the practical applicability of CM, as well as its benefits, clear to the learners, which, in turn, should motivate a later self-regulated use of the strategy [22,129].

1.4. Research Question and Hypotheses

Accordingly, one aim of our study is to answer the question of whether participants who receive a specific CM training are generally more successful in learning using CM than students who have been introduced to CM in another way.

Corresponding to this, we hypothesize that a respective CM training causes certain familiarity with the strategy, which, in turn, leads to:

- (1) More organization and elaboration processes during learning in terms of information integration into prior knowledge structures and thus knowledge consolidation.
- (2) Less perceived cognitive load during learning when using CM.
- (3) Better handling and editing prescribed and creating own concept maps, respectively.

2. Materials and Methods

Our quasi-experimental study was conducted over a six-week period and based on three experimental conditions, varying regarding the type and intensity of CM training. While the training phase dealt with the non-biological topic of intelligence, the learning and later test phases referred to the topic of *cell biology*, since we wanted to prevent test effects.

2.1. Sample

The recruitment of our sample ($N = 73$) from students at the University of Cologne (Germany) was carried out by offering the possibility of participating in an elective module to improve learning and using learning strategies successfully. The group assignment was based on self-selection for organizational reasons because the students needed to solve their everyday study program, too. Concretely, this means that interested students could choose one of three weekdays offered, in which they would each take part in our study during the following six weeks. Our resulting quasi-experimental design allowed for satisfactorily addressing our test conditions (see Section 3.1). The first group (T++, $n = 27$) received a CM training with scaffolding and feedback, the second group (T+, $n = 21$) received a CM training without scaffolding and feedback but only repeated practice, and the third group (T-, $n = 25$) received no special CM training but a short verbal instruction to its construction techniques. On average, the participants in our study were 22.6 years old, 78% were female, and 56% were enrolled in a natural science study program, whereas the others primarily studied a subject of humanities.

2.2. Experimental Design

The procedure of our study was divided into three consecutive phases (training, learning, and testing), split into six weekly sessions: three training sessions, one learning session, and two test sessions, each on one day per week (see Figure 2).

Training phase

- **Week 1:**
 - assessment of socio-demographics and potentially confounding variables
 - implementation of the three different training conditions (T++ vs. T+ vs. T-)
- **Week 2:** continuation of different of trainings (T++ vs. T+ vs. T-)
- **Week 3:**
 - completion of different training (T++ vs. T+ vs. T-)
 - treatment check: assessment of declarative knowledge about the training phase's learning objective *intelligence*



Learning phase

- **Week 4:**
 - group-independent standardized short introduction to CM for for two reasons: (1.) enabling also T- participants to use the strategy during the following learning session, and (2.) refreshing CM-related knowledge of the T++ and T+ groups
 - assessment of CM-related self-efficacy and objective assessment of the competence in handling CMs by presenting an error-including CM to be corrected
 - text-based learning on *cell biology*
 - assessment of CM competence and cognitive processes via creation of a self-constructed concept map about the learning phase's learning objective *cell biology*, assessment of the perceived cognitive load as well as a metacognitive prediction regarding the declarative knowledge remembered one week later



Test phase

- **Week 5:** assessment of CM competence via creation of self-constructed concept maps on the the learning objective *cell biology* as well as assessment of declarative and structural knowledge about *cell biology*
- **Week 6:** assessment of conceptual knowledge about *cell biology*

Figure 2. Timeline of the experimental procedure (T++ = CM training ($n = 27$), including additional scaffolding and feedback elements; T+ = CM training ($n = 21$), including no additional scaffolding and feedback elements; T- = control training ($n = 25$) including non-CM strategies).

The three experimental groups in our study varied in the intensity and content of support offered to participants during the CM training phase:

- (1) The first group (T++) was given additional scaffolding and feedback during CM practice and participants received a strictly guided training. First, the instructor gave an overview of the day's learning objectives and main topics in terms of an advance organizer. An introduction to the CM strategy followed, including declarative knowledge elements about its practical use. A list of metacognitive prompts [41] (e.g., "Did I label all arrows clearly, concisely and correctly?" or "Where can I draw new connections?"; see Supplementary Materials Sections S2.3 and S2.4) was handed out for the individual work phase to induce the use of the relevant learning strategies of elaboration and organization [18], but prompts were reduced over the course of the training phase in terms of fading [130]. During the subsequent work phase, participants constructed concept maps based on learning text passages dealing with the abstract topic of intelligence. Here, they received different scaffolds (see Supplementary Materials Sections S2.1 and 2.2): in week 1, a skeleton map allowed participants to focus entirely on the main concepts and linking terms [80]; in week 2, a given set of the 12 main

concepts taken from the learning material allowed participants to define links between these concepts on their own; and in week 3, T++-participants were able to construct concept maps completely by themselves. After each work phase, one of the participants' completed maps was transferred to a blackboard and discussed. An expert map was presented and discussed as well, so participants were able to compare it to their own. Additionally, all participants received individual verbal as well as written feedback on their constructed concept maps. Considering previous findings [131–133], we decided on a knowledge of correct results (KCR) feedback approach but limited the feedback to marking CM errors and pointing out any resulting misconceptions. In addition, a list of the most common CM errors was available (see Supplementary Materials Sections S2.5 and S2.6).

- (2) The second group (T+) constructed concept maps in each session on their own without any additional support. As analyzing and providing feedback is very time-consuming, this approach is more economic and has been found to be effective as well [134]. The training sessions' sequence was framed in almost the same manner as in group T++: all sessions started with an advance organizer, followed by an introduction to CM but without giving metacognitive prompts. During the subsequent individual work phase, participants worked on the same learning material but without any scaffolding during their own concept map construction. All participants had the opportunity to ask questions during the introduction but were not able to compare their own results to one of the participants' or an expert map. Given these characteristics, this group represents the kind of practical training most likely found in classrooms [135,136].
- (3) The third group (T-) did not receive any CM training but used common non-CM learning strategies from other studies [18,114,137] to deal with the same learning material as the T++ and T+ groups: group discussions in week 1, writing a summary in week 2, and carousel workshops in week 3. The training sessions' procedure was again framed in almost the same manner as in group T++: they started with an advance organizer, followed by an introduction to the respective learning strategy including metacognitive prompts (see Supplementary Materials Section S3). During the subsequent individual work phase, participants worked on the learning material and afterwards, they were given the opportunity to discuss their results and compare them to an expert solution.

2.3. Concept Map Scoring

There are many different scoring systems available for evaluating concept maps [138,139] due to their complexity and the implied hierarchical structure. To determine a CM score in our study, we applied two measures: On the one hand, we followed an approach from McClure et al. [140] assigning each proposition specified a value of 0 points (no relation between the concepts), 1 point (there is a relation between the concepts, but the arrow label is incorrect), 2 points (the arrow label is correct, but the arrow's direction is not), or 3 points (the whole proposition is correct and meaningful). The combined score from all propositions specified by the participants was used as a learning success measure of the test phase, since participants who specify more propositions demonstrate their ability to build meaningful connections between a larger number of concepts. On the other hand, we developed an additional ratio score: the *balanced Quality of Concept Map Index* (bQCM Index), taking the quantitative performance aspect into account (the task was limited in time) as well: due to the trivial fact that the probability of making an error increases with a larger number of propositions specified, we put the concept map score as well as the numbers of technical and content-related CM errors into relation to the number of the participants' overall propositions to provide a measure of training effectiveness. For the sake of clarity, the first-mentioned score following McClure et al.'s [140] approach, is called the *absolute Quality of Concept Map Index* (aQCM Index) below.

In addition, we also evaluated technical and content-related CM errors in the participants' concept maps. The categorization of errors allows for assessing a person's CM competence by focusing on the number of CM technical errors in relation to the overall number of propositions specified (*concept map error ratio*), and assessing a person's understanding of the learning content by focusing on the number of content errors in relation to the overall number of propositions specified (*content-related error ratio*).

Other added measures to assess the quality of concept maps constructed by participants were the numbers of recall-suggesting (R-) propositions, organization-suggesting (O-) propositions, and elaboration-suggesting (E-) propositions (see Section 1.1, Section 2.4.4, and Section 3.4).

2.4. Further Measures and Operationalizations

In the following, we describe the operationalizations of the other dependent variables as well as those, we controlled for prior to the training. Their presentation is not based on the order of the hypotheses, but on the chronology of the study's phases.

2.4.1. Measures Prior to Training Phase (Week 1)

To ensure the best comparability of our three groups (T++, T+, and T-), we first assessed the participants' demographics (gender, age, etc.) and other relevant variables we wanted to control for, such as familiarity with CM, prior knowledge in biology, or reading comprehension (see Supplementary Materials Section S1).

As the quality of students' learning is strongly determined by their prior knowledge [117], we assessed our participants' prior knowledge regarding the learning phase's learning objective of cell biology using an 18-item test, comprising established item sets from previous studies [141–143] which were partly adapted. After 3 of these 18 items were excluded from the subsequent analysis due to lack of reliability, the test showed an internal consistency of $\alpha = 0.74$ for the present sample.

In addition, the participants were asked about the extent of their biology education during the last two years of schooling: no biology vs. basic biology vs. advanced biology (the only possible options regarding the German school system).

To assess the participants' familiarity with concept maps, a seven-item scale from McClure et al. [140] was used and the following adaptations made: The answer-format ("yes" or "no") was replaced by a 5-point rating scale (1 = *never/very rarely* to 5 = *very often/always*), the prior use of concept maps in school was assessed with the first question instead of asking whether the participant had never used concept maps, and the benefit of concept maps when solving problems was specified (see Supplementary Materials Section S1.3). The internal consistency of this set was $\alpha = 0.89$.

As the topics and contents of the learning texts provided in the training (learning objective: intelligence) and the learning phase (learning objective: cell biology) were considered sophisticated, but had to be read in a given timeframe, the participants' reading speed and comprehension were additionally assessed via the *Lesegeschwindigkeits- und Verständnistest 6-12* (LGVT 6-12) [144]. In this test, participants are given four minutes to read as much as possible of a story about a woman trying to plant brussels sprouts herself. The text includes multiple sentences which are interrupted through brackets and the reader is instructed to mark the correct word out of three options in these brackets to continue the story in a meaningful way. The LGVT's internal consistency was $\alpha = 0.67$ for the present sample.

2.4.2. Measures after Training Phase/Treatment Check (Week 3)

After all participants had completed their respective training conditions on the learning objective *intelligence*, they were presented a self-designed multiple-choice test (20 items; $\alpha = 0.75$) to assess their declarative knowledge about this topic (see Supplementary Materials Section S4). The respective results (see Section 3.2) were used to check for immediate effects of the three different training conditions and therefore served as a treatment check.

2.4.3. Measures Prior to Learning Phase (Week 4)

After this training phase, the participants of all three conditions received the same short introduction to the CM strategy at the beginning of week 4 to set comparable starting conditions regarding the subsequent learning phase: on the one hand, one week had passed since the training had ended; on the other hand, the T- group had not received any CM training yet. Immediately after this short introduction, the participants of all three groups were asked to fill in a questionnaire on their CM-related self-efficacy expectations (see Supplementary Materials Section S5). This assessment of the participants' CM-related self-efficacy was used to quantify potential differences between the three different conditions since it could be expected that the T- group, which only received the short introduction, would rank their self-efficacy significantly lower. The chosen self-designed questionnaire ($\alpha = 0.88$) consisted of six items (e.g., "I feel competent in choosing the important concepts for my concept map"), which were each rated on a 5-point scale (see Supplementary Materials Section S5).

Furthermore, they completed an error detection task by working on a prescribed concept map, including 18 propositions and some errors (e.g., missing arrowhead; see Supplementary Materials Section S6). The task was to detect and correct as many errors as possible out of a total of 10 in a given timeframe of four minutes. This task was implemented to obtain an additional objective measure of the competent use of the CM strategy beside the self-efficacy ratings. Here, a total of two measures was considered: (1) the quality of error detection and (2) the quality of error correction.

2.4.4. Measures during Learning Phase (Week 4)

Participants' concept maps constructed during the learning phase on the topic of cell biology were evaluated following the bQCM Index procedure (see Section 2.3). To ensure evaluation objectivity in terms of interrater reliability, all participants' concept maps were separately coded by two researchers, following recommendations by Wirtz and Caspar [145]. The intraclass correlation coefficient (ICC) and its 95% confidence interval were interpreted using Koo and Li's [146] guidelines. The analysis yielded excellent interrater reliability

(ICC = 0.90, CI_{95%} [0.56, 0.96]) on average, indicating that participants' concept maps can be clearly judged using the coding scheme of scoring.

In addition, the participants' concept maps were analyzed with regard to the complexity of specified propositions, serving as an indicator of cognitive processes during learning. To that, we analyzed (1) how many relations were taken directly from the text (recall-suggesting [R-] propositions); (2) how many relations, that were not explicitly named in the text, were constructed between concepts of the text (organization-suggesting [O-] propositions); and (3) how many relations were constructed between a concept in the text and a concept that was not named in the text (elaboration-suggesting [E-] propositions). While a specification of R- and O-propositions can be accomplished with the text, the specification of E-propositions required an integration of prior knowledge elements into the concept map.

The participants' perceived cognitive load during CM was assessed using the *Cognitive Load Questionnaire* (CLT) by Klepsch et al. [96]. The scale comprises seven items ($\alpha = 0.69$) which refer to three distinguishable aspects of cognitive load: (1) intrinsic cognitive load (ICL, $\alpha = 0.62$) results from demands made by the learning task itself; (2) germane cognitive load (GCL, $\alpha = 0.72$) is related to the mental monitoring of learning and strategy use; and (3) extraneous cognitive load (ECL, $\alpha = 0.68$) is not directly linked to learning since it only raises if ill-designed learning material/instruction claims additional mental resources. The CLT's items were adapted by replacing the phrase "this task" with "constructing the concept map" as well as referring to the presentation of information in the text instead of the task itself in the second ECL-related question (see Supplementary Materials Section S8). All statements that had to be rated on a 7-point scale (1 = *I do not agree at all* to 7 = *I agree completely*).

Finally, in accordance with Karpicke and Blunt [147] and Blunt and Karpicke [148], participants gave metacognitive predictions in terms of a judgment of learning ("How much of the information from the text you will remember in one week from 0 to 100%?"). As CM can be a useful metacognitive learning tool [34,149], these metacognitive judgments were used to examine possible differences between participants' self-predicted performance on questions concerning the topic of cell biology and their factual declarative knowledge assessed one week later (see Section 2.4.5).

2.4.5. *Measures after Learning Phase/Learning Outcome (Weeks 5 and 6)*

Concerning the three groups' learning outcomes, we conducted several measures in a test phase that took place on two days within two consecutive weeks, covering CM competence as well as three different domains of knowledge: declarative, structural, and conceptual knowledge. While declarative knowledge consists of segregated information and facts, structural knowledge focuses more on contextual characteristics such as the relationships between these individual pieces of information and facts. Finally, conceptual knowledge categorizes the features and principles inherent in the data, facts, and relations in a highly decontextualized manner, so a learner becomes able to use his or her knowledge flexibly in various contexts [150–154].

During the first test session (week 5, 3-h testing), participants were first asked to create a concept map within 60 min using a given set of 22 concepts from the learning text on the topic of cell biology from the learning phase one week before. Concept map evaluation followed the aQCM Index procedure for measuring learning success and the bQCM Index procedure for measuring CM quality (see Section 2.3). As in the week before, all concept maps were separately coded by two researchers to assess interrater reliability [145]. The analysis yielded excellent objectivity as well (ICC = 0.93, CI_{95%} [0.88, 0.96]) [146].

After the concept map was created, a 30-min *Similarity Judgments Test* (SJT) [155] was carried out to assess the participants' structural knowledge about cell biology. Within this 55-item test, the semantic proximity of two given cell biological concepts (e.g., nucleus–bio membrane or ribosome–DNA) had to be rated on a 9-point scale (1 = *minimally related* to 9 = *strongly related*). A total of 11 central cell biological concepts was used, which were presented in a balanced manner. To prevent confounding sequence effects, two test versions were used, in which the sequence of the 55 pairs of terms was varied randomly (see Supplementary Materials Section S9). The SJT was first given to a sample of seven experts. The intraclass correlation coefficient (ICC) regarding the experts' ratings indicates excellent interrater reliability [146] (ICC = 0.95, CI_{95%} [0.93, 0.97]), so face validity of the procedure can be assumed. The mean of the experts' ratings on each item was used as an evaluation standard and the correlation of the participants' responses with those of the experts was finally used as the measure of the participants' structural knowledge (see Section 3.7).

At the end of this first test session, the participants completed a 60-min test on their declarative knowledge about cell biology. This self-designed multiple-choice questionnaire (see Supplementary Materials Section S10) comprised a total of 30 items (e.g., “Which of the following statements about the cell wall are correct?”). After two of these 30 items were

excluded from the subsequent analysis due to lack of reliability, the test showed an internal consistency of $\alpha = 0.80$.

During the second test session (week 6, 90-min testing), the participants' conceptual knowledge about cell biology was assessed. For this purpose, a self-designed test ($\alpha = 0.79$) with 15 open-answer items was used (see Supplementary Materials Section S11). An example of an item is: "The intake of poison from the destroying angel mushroom (α -Amanitin) leads to inhibition of the enzyme RNA-Polymerase II, which is responsible for the synthesis of an mRNA-copy of DNA information. Which consequences for protein synthesis as well as for the entire cell could be expected? Justify your answer." Answers were coded as wrong (0 points), partially correct (1 point), or correct (2 points).

2.5. Materials and Procedure

In the training phase, our participants worked on a learning text on the topic of intelligence, whereas the learning material in the following learning phase belonged to the topic of cell biology. The text *Theories and Models of Intelligence* comprised 3197 words (8 pages) and was divided into three parts according to the three training sessions. The text *The Structure and Function of Eukaryotic Cells* comprised 2010 words (7 pages; see Supplementary Materials Section S7) and was used in the learning phase. Regarding the ecological validity, both the subject and content of these texts were rated by experts as adequate for learning at the university level prior to the study. In this manner, consistency, comprehensibility, and stringent reasoning were ensured as well.

In order to minimize Rosenthal/Pygmalion effects, a maximum standard of behavior and situation, which was fixed in writing and secured via a second investigator, was developed, which was strictly implemented in all study-related interactions with the participants.

2.6. Statistical Analyses

The collected data were analyzed using parametric and non-parametric procedures to detect potential group differences and relevant correlations. As appropriate, we further applied post hoc procedures to provide further information on the quantity and direction of group differences concerning single dependent variables.

3. Results

Below, we report our results of the statistical analyses based on the data of our $N = 73$ participants. Herein, we orient towards the chronological order of events, i.e., the point of measurement of the respective dependent variables during the study period.

3.1. Pre-Analyses of Baseline Differences (Measures Prior to Training Phase, Week 1)

To check whether there were significant baseline differences on relevant variables between the three training groups, we carried out difference testing for both categorical (χ^2 -test; see Table 1) and metric variables (univariate ANOVA; see Table 2). The results of the χ^2 -test indicate an equal distribution regarding the participants' gender and university study program across the three groups. Regarding participation in biology courses during the last two years of schooling, however, there was a significant group difference, $\chi^2(4) = 11.96, p < 0.05$ (see Table 1).

Table 1. Results of cross-table analyses regarding possible baseline differences between the three groups.

Variable	Level	Observed n in Groups			Comparison (Groups)
		T++ ¹	T+ ²	T- ³	
Gender	female	18	18	21	$\chi^2(2) = 3.28, p = 0.19$
	male	9	3	4	
Educational level in biology ⁴	essential	5	5	7	$\chi^2(4) = 11.96, p < 0.05$
	basic	11	7	17	
	advanced	11	9	1	
University study program	B. A.	10	9	13	$\chi^2(2) = 1.19, p = 0.55$
	B. Sc.	17	12	12	

¹ T++ = CM training ($n = 27$), including additional scaffolding and feedback elements; ² T+ = CM training ($n = 21$), including no additional scaffolding and feedback elements; ³ T- = control training ($n = 25$) including non-CM strategies; ⁴ an essential educational level in biology means that the participants did not participate in biology lessons at all during their last two years of schooling, whereas “basic” and “advanced” mean that they participated in biology lessons, but to a different extent.

Table 2. Descriptive statistics and results of ANOVA regarding possible baseline differences between the three groups.

Variable	Group						Comparison (Groups)
	T ₊₊ ¹		T ₊ ²		T ₋ ³		
	<i>M</i> ⁴	<i>SD</i> ⁵	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Age	22.93	5.94	22.05	3.37	22.64	5.92	$F(2, 70) = 0.16, p = 0.85$
Final school exam grade	2.02	0.63	2.07	0.69	1.81	0.55	$F(2, 70) = 1.10, p = 0.34$
Prior knowledge in cell biology	7.74	3.86	7.52	3.22	6.64	2.93	$F(2, 70) = 0.75, p = 0.48$
Familiarity with concept maps	1.55	0.66	1.73	0.75	1.33	0.40	$F(2, 70) = 1.64, p = 0.20$
Reading speed	842.33	247.32	845.14	271.97	835.52	225.37	$F(2, 70) = 0.01, p = 0.99$
Reading comprehension	16.67	6.72	18.33	5.84	16.68	6.52	$F(2, 70) = 0.50, p = 0.61$

¹ T₊₊ = CM training ($n = 27$), including additional scaffolding and feedback elements; ² T₊ = CM training ($n = 21$), including no additional scaffolding and feedback elements; ³ T₋ = control training ($n = 25$) including non-CM strategies; ⁴ mean value; ⁵ standard deviation.

The results of the univariate ANOVA indicate an equal distribution across the three groups for all variables. The participants in the three conditions neither differ regarding their age or final school exam grade nor regarding their prior knowledge about cell biology, their familiarity with CM, or their reading competencies (see Table 2).

Overall, the results of our pre-analyses show no significant group differences regarding potentially confounding variables. We consider the statistically significant group difference concerning participation in biology classes during the last two years of schooling (see Table 1) to be irrelevant since the groups did not differ significantly in their prior knowledge about cell biology (see Table 2) and are therefore comparable to each other in terms of the dependent variables relevant to us. Accordingly, no covariates were additionally included in subsequent difference-testing analyses.

3.2. Treatment Check (Measure after Training Phase, Week 3)

After participants had completed the three training sessions working on the topic of intelligence, we assessed their declarative knowledge about this learning objective to evaluate the immediate effect of the three different training conditions (see Section 2.4.2). A univariate ANOVA showed that the three groups did not differ in their knowledge immediately after training, $F(2, 70) = 0.08, p = 0.92$, indicating that potential effects of the training conditions on knowledge acquisition do not seem to occur without any time delay.

3.3. Concept Mapping-Related Self-Efficacy and Error Detection Task (Measures Prior to Learning Phase, Week 4)

Regarding the participants' CM-related self-efficacy, we expected that the two groups who received a CM training would feel more competent in handling concept maps than the T- group without such a training (see Section 2.4.3). Contrary to this expectation, the T- group rated its CM-related self-efficacy highest. However, the Kruskal-Wallis test did not indicate any statistically significant difference between the three groups, $\chi^2(2) = 2.75, p = 0.25$ (see Table 3).

Table 3. Descriptive statistics and results of ANOVA and Kruskal-Wallis tests regarding group differences in CM-related self-efficacy and the error correction task performance.

Variable	Group						Comparison (Groups)
	T++ ¹		T+ ²		T- ³		
	<i>M</i> ⁴	<i>SD</i> ⁵	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Error detection ⁶	7.74	1.66	7.05	1.53	7.12	1.39	$F(2, 70) = 1.56, p = 0.22$
Proper error correction ⁶	7.67	1.75	7.10	1.70	6.80	1.56	$\chi^2(2) = 6.10, p < 0.05$
Improper error correction	0.48	0.64	0.86	0.96	1.08	1.00	$\chi^2(2) = 5.41, p = 0.08$
CM-related self-efficacy	5.11	1.03	5.38	1.03	5.47	0.98	$\chi^2(2) = 2.75, p = 0.25$

¹ T++ = CM training ($n = 27$), including additional scaffolding and feedback elements; ² T+ = CM training ($n = 21$), including no additional scaffolding and feedback elements; ³ T- = control training ($n = 25$) including non-CM strategies; ⁴ mean value; ⁵ standard deviation; ⁶ the maximum scores for error detection and proper error correction were each 10.

As part of the error detection task (see Section 2.4.3), participants should detect and correct as many errors as possible out of a total of 10 errors in a prescribed concept map with 18 specified propositions. In total, two measures were considered, (1) the quality of error detection and (2) the quality of error correction. Regarding the first measure (error detection) a comparison of the participants' performance showed that the T++ group performed slightly better than the groups T+ and T-. However, these differences are not statistically significant, indicating that the CM training had no significant effect on the error detection rate. Regarding the second measure (error correction), a comparison indicates a significant group difference in the mean number of properly corrected errors, $\chi^2(2) = 6.10, p < 0.05$ (see Table 3). Post hoc pairwise comparisons showed that this significance is particularly based on the difference between the groups T++ and T- ($z = 2.34, p < 0.05$). The differences between the groups T++ and T+ ($z = 1.82, p = 0.07$) and between the groups T+ and T- turned out to be insignificant ($z = 0.24, p = 0.81$). Regarding a second aspect of error correction quality, analyses showed mean differences, suggesting that the T- group made a correct proposition incorrect and/or recognized a CM error but were unable to correct it properly and therefore only replaced one error with another. However, these differences proved to be insignificant by a Kruskal-Wallis test, so the number of improperly corrected errors seems largely unaffected by the CM training (see Table 3).

3.4. Recall, Organization, and Elaboration Processes (Measures during Learning Phase, Week 4)

Within our hypotheses, we further assumed that CM training would have a positive effect on organization and elaboration processes during CM (see Section 1.4), operationalized by the number of recall-suggesting (R-) propositions, organization-suggesting (O-) propositions, and elaboration-suggesting (E-) propositions specified. Group comparisons, however, show no significant mean differences between the three groups regarding the extent to which different classes of propositions are specified (see Table 4).

Table 4. Descriptive statistics and results of ANOVA and Kruskal–Wallis tests regarding different types of specified propositions.

Variable	Group						Comparison (Groups)
	T++ ¹		T+ ²		T- ³		
	<i>M</i> ⁴	<i>SD</i> ⁵	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
R-propositions ⁶	36.74	14.77	46.95	17.43	38.08	14.36	$F(2, 70) = 2.92, p = 0.07$
O-propositions ⁷	0.41	0.89	0.81	1.63	0.44	0.65	$\chi^2(2) = 0.48, p = 0.79$
E-propositions ⁸	0.41	0.97	0.24	1.09	0.12	0.44	$\chi^2(2) = 2.50, p = 0.29$

¹ T++ = CM training ($n = 27$), including additional scaffolding and feedback elements; ² T+ = CM training ($n = 21$), including no additional scaffolding and feedback elements; ³ T- = control training ($n = 25$) including non-CM strategies; ⁴ mean value; ⁵ standard deviation; ⁶ recall-suggesting propositions; ⁷ organization-suggesting propositions; ⁸ elaboration-suggesting propositions.

3.5. Cognitive Load (Measures during Learning Phase, Week 4)

Regarding the participants' cognitive load, we assumed that greater familiarity with the CM strategy would reduce the perceived cognitive load during CM (see Section 1.4). The mean values for the three distinguishable facets of cognitive load are presented in Table 5. With regard to the dimension of extraneous cognitive load (ECL), a comparison indicates a significant group difference, $\chi^2(2) = 6.51, p < 0.05$ (see Table 5). Post hoc pairwise comparisons showed that this significance is particularly based on the difference between the groups T+ and T- ($z = -2.54, p < 0.05$). The differences between the groups T++ and T+ ($z = 1.18, p = 0.72$) and between the groups T++ and T- turned out to be insignificant ($z = -1.47, p = 0.42$). Regarding the two other dimensions of cognitive load, the results of the Kruskal–Wallis tests showed no statistically significant differences between the three groups (see Table 5). Altogether, these results indicate that participants from the T- group regarded the learning material more taxing (ECL) than the T+ group but all groups regarded the CM task itself equally complex (ICL), were equally engaged, and directed their mental resources to learning processes (GCL) to a very similar extent.

Table 5. Descriptive statistics and results of and Kruskal–Wallis tests regarding the facets of perceived cognitive load.

Variable	Group						Comparison (Groups)
	T ₊₊ ¹		T ₊ ²		T ₋ ³		
	<i>M</i> ⁴	<i>SD</i> ⁵	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Extraneous cognitive load	4.10	1.36	3.62	1.14	4.49	1.25	$\chi^2(2) = 6.51, p < 0.05$
Intrinsic cognitive load	5.87	0.88	5.41	1.00	5.86	1.22	$\chi^2(2) = 3.48, p = 0.18$
Germane cognitive load	5.67	1.16	5.43	1.09	5.60	1.42	$\chi^2(2) = 0.91, p = 0.64$

¹ T₊₊ = CM training ($n = 27$), including additional scaffolding and feedback elements; ² T₊ = CM training ($n = 21$), including no additional scaffolding and feedback elements; ³ T₋ = control training ($n = 25$) including non-CM strategies; ⁴ mean value; ⁵ standard deviation.

3.6. Metacognitive Prediction (Measures during Learning Phase, Week 4)

Analyses of the participants' judgments of learning (metacognitive predictions on how much they would remember after one week) assessed at the end of the learning phase also yielded no significant differences between the groups, $F(2, 70) = 1.82, p = 0.17$, indicating that the training did not influence the participants' metacognitive prediction. Numerically, group T₊ predicted the highest learning retention ($M = 58.1, SD = 21.12$), followed by group T₊₊ ($M = 50.74, SD = 28.14$), and T₋ ($M = 44.4, SD = 24.51$). However, beyond that, correlational analyses show significant relationships between these judgments of learning and the aQCM Index (indicating overall concept map quality) as well as the scores for declarative and structural knowledge (SJT) of the later test phase, indicating that the participants' judgments of learning were quite accurate across all groups (see Table 6).

Table 6. Correlations between the judgments of learning and factual later learning outcome.

Variable	1	2	3	4	5
1 Metacognitive prediction	-				
2 aQCM Index ¹	0.56 ***	-			
3 Structural knowledge	0.62 ***	0.66 ***	-		
4 Declarative knowledge	0.71 ***	0.61 ***	0.65 ***	-	
5 Conceptual knowledge	-0.01	0.27 *	-0.08	0.15	-

¹ absolute Quality of Concept Map Index; * = $p < 0.05$; *** = $p < 0.001$.

3.7. Concept Map Quality, Structural, Declarative, and Conceptual Knowledge (Measures after Learning Phase/Learning Outcome in Test Phase, Weeks 5 and 6)

With regard to the participants' learning outcomes, we stated that CM training would facilitate learning using CM, and thus, would cause better overall learning success (see Section 1.4). This learning outcome was assessed by evaluating the concept maps constructed by the participants using the aQCM Index, and the achieved scores in tests of structural, declarative, as well as conceptual knowledge about the learning objective of cell biology. The MANOVA carried out indicated significant overall group differences in terms of learning success $F(8, 134) = 3.36, p < 0.01, \eta^2 = 0.17$. A post hoc univariate ANOVA shows that the significant group differences belong to the dependent variables concept map quality (aQCM Index), structural (SJT score), and conceptual knowledge, whereas the dependent variable declarative knowledge only shows a clear, albeit statistically insignificant, difference. Corresponding post hoc pairwise comparisons indicate that the significant differences are largely attributable to the differences between the T- group and the two training groups T++ and T+ (see Table 7).

Table 7. Descriptive statistics and results of ANOVA and post hoc tests regarding the content-related learning outcome.

Variable	Univariate Comparisons			Post Hoc Tests ¹					
	<i>F</i> -Test	<i>p</i>	<i>H</i> ²	Group	<i>N</i> ²	<i>M</i> ³	<i>SD</i> ⁴	Comparison	<i>p</i>
aQCM Index ⁸	<i>F</i> (2, 70) = 5.67	<0.01	0.14	T++ ⁵	27	61.52	18.82	T+	1.00
								T-	0.01
				T+ ⁶	21	57.19	19.31	T++	1.00
								T-	0.08
				T- ⁷	25	45.56	14.09	T++	0.01
								T+	0.08
Structural knowledge	<i>F</i> (2, 70) = 4.38	<0.05	0.11	T++	27	0.46	0.17	T+	1.00
								T-	0.03
				T+	21	0.45	0.18	T++	1.00
								T-	0.06
				T-	25	0.33	0.15	T++	0.03
								T+	0.06
Declarative knowledge	<i>F</i> (2, 70) = 2.52	0.09	0.07	T++	27	15.19	6.12	T+	0.56
								T-	0.98
				T+	21	17.14	4.23	T++	0.56
								T-	0.09
				T-	25	13.80	4.32	T++	0.98
								T+	0.09

Table 7. (Fortsetzung)

					5.79	T+	1.00	
	0.14	T++	27	12.96		T-	0.01	
Conceptual knowledge	$F(2, 70) = 5.67$	<0.01			4.97	T++	1.00	
			T+	21	12.62		T-	0.02
						3.39	T++	0.01
			T-	25	8.80		T+	0.02

¹ Tamhane's test considering unequal variances was carried out for comparisons regarding conceptual knowledge, whereas Bonferroni's test was carried out in all other cases; ² number of participants; ³ mean value; ⁴ standard deviation; ⁵ T++ = CM training ($n = 27$), including additional scaffolding and feedback elements; ⁶ T+ = CM training ($n = 21$), including no additional scaffolding and feedback elements; ⁷ T- = control training ($n = 25$) including non-CM strategies; ⁸ absolute Quality of Concept Map Index.

In addition to this analysis of group differences, it seemed expedient to analyze the learning development of our participants between learning and test phase in terms of a training-related advantage in concept map quality. Within the corresponding repeated measure ANOVA, we considered the CM's quality (bQCM Index; see Section 2.3) as well as the proportional frequency of technical errors (concept map error ratio; see Section 2.3) as dependent variables. However, there was no statistically significant interaction effect between the three groups and the point of measurement (POM) on these two variables (see Table 8).

Table 8. Descriptive statistics and results of repeated measure ANOVA regarding concept map quality over time within the different groups.

Variable	Group	N^1	POM ²	M^3	SD^4	<i>F</i> -Test	
						(Group × Time)	
bQCM-Index ¹⁰	T++ ⁵	27	t ₁ ⁸	2.49	0.47	<i>F</i> (2, 70) = 2.27	0.11
			t ₂ ⁹	2.38	0.43		
	T+ ⁶	21	t ₁	2.61	0.45		
			t ₂	2.30	0.49		
	T- ⁷	25	t ₁	2.01	0.47		
			t ₂	1.96	0.49		
Concept map error ratio	T++	27	t ₁	0.05	0.09	<i>F</i> (2, 70) = 0.56	0.57
			t ₂	0.05	0.08		
	T+	21	t ₁	0.04	0.06		
			t ₂	0.09	0.15		
	T-	25	t ₁	0.14	0.12		
			t ₂	0.16	0.15		

¹ number of participants; ² point of measurement; ³ mean value; ⁴ standard deviation; ⁵ T++ = CM training ($n = 27$), including additional scaffolding and feedback elements; ⁶ T+ = CM training ($n = 21$), including no additional scaffolding and feedback elements; ⁷ T- = control training ($n = 25$) including non-CM strategies; ⁸ week 4 (learning phase); ⁹ week 5 (test phase); ¹⁰ balanced Quality of Concept Map Index.

Although the repeated measures analysis turned not out to be significant for the interaction effect across all groups, post hoc tests to check for individual differences were carried out considering inequality of variances. For both dependent variables, it turned out that the group T- differed significantly from both training groups (T++ and T+), whereas there were no significant differences between these two training groups (see Table 9; Figure 3 and Figure 4).

Table 9. Post hoc test results for the interaction effect group \times time.

Variable	Group	Comparison	<i>p</i>
bQCM Index ⁴	T ₊₊ ¹	T ₊	1.00
		T ₋	<0.01
	T ₊ ²	T ₊₊	1.00
		T ₋	<0.01
	T ₋ ³	T ₊₊	<0.01
		T ₊	<0.01
Concept map error ratio	T ₊₊	T ₊	1.00
		T ₋	<0.01
	T ₊	T ₊₊	1.00
		T ₋	<0.05
	T ₋	T ₊₊	<0.01
		T ₊	<0.05

¹ T₊₊ = CM training ($n = 27$), including additional scaffolding and feedback elements; ² T₊ = CM training ($n = 21$), including no additional scaffolding and feedback elements; ³ T₋ = control training ($n = 25$) including non-CM strategies; ⁴ balanced Quality of Concept Map Index.

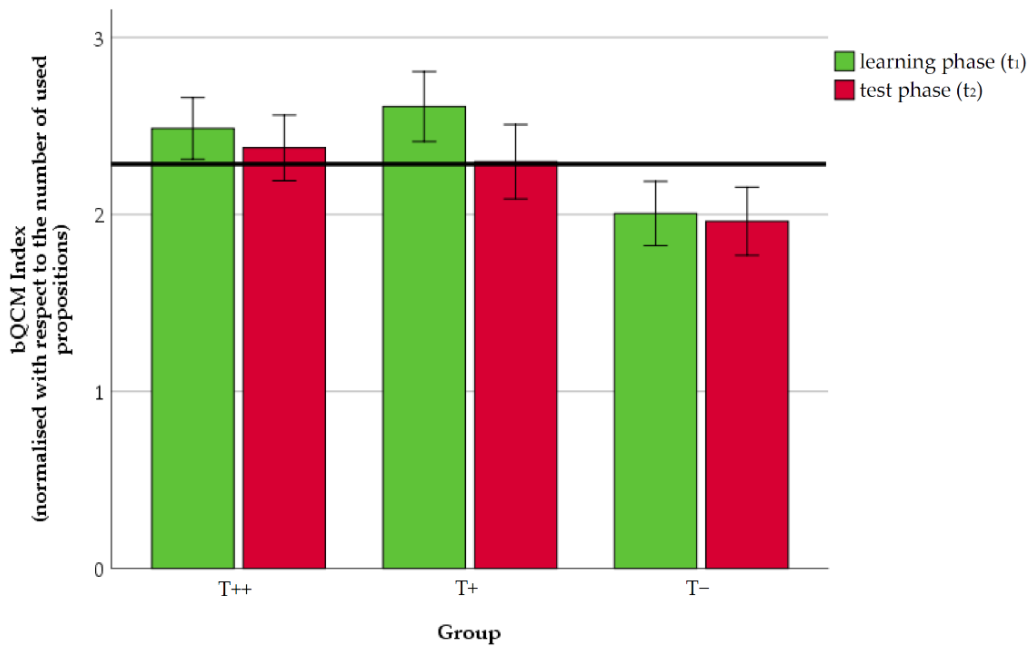


Figure 3. Bar chart representing intragroup changes in bQCM Index between the learning phase in week 4 and the test phase in week 5 (T++ = CM training ($n = 27$), including additional scaffolding and feedback elements; T+ = CM training ($n = 21$), including no additional scaffolding and feedback elements; T- = control training ($n = 25$) including non-CM strategies; error bars are based on standard error of mean; the overall mean value is represented by the bold line in black).

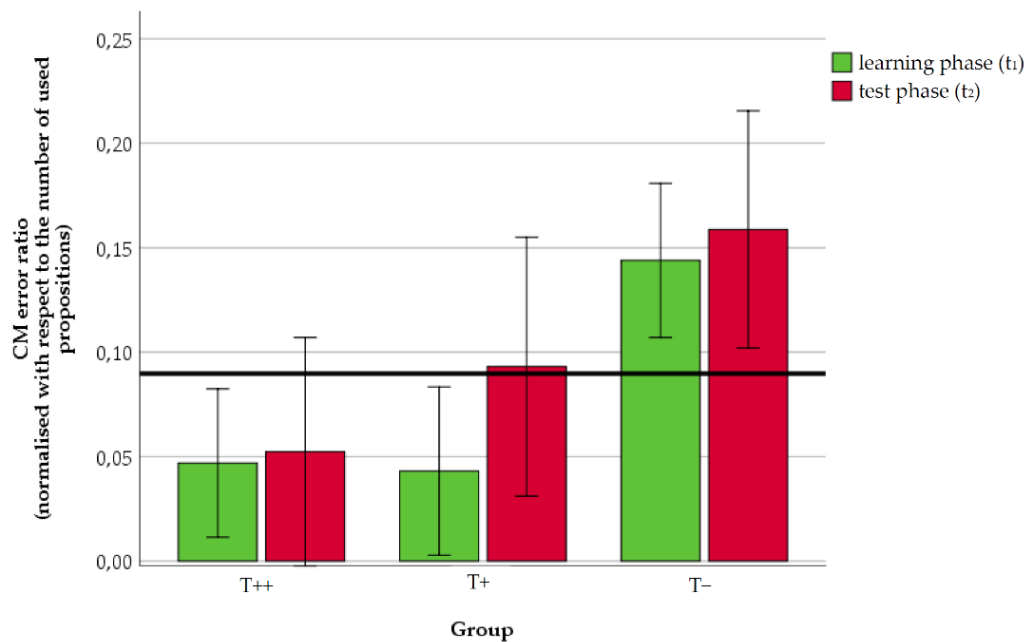


Figure 4. Bar chart representing intragroup changes in CM error ratio between the learning phase in week 4 and the test phase in week 5 (T++ = CM training ($n = 27$), including additional

Figure 4. (*Fortsetzung*)

scaffolding and feedback elements; T+ = CM training ($n = 21$), including no additional scaffolding and feedback elements; T- = control training ($n = 25$) including non-CM strategies; error bars are based on standard error of mean; the overall mean value is represented by the bold line in black).

Taking a closer look at the differences within groups between the two POMs, the bQCM Index shows that the performance of the groups T++ and T- did not change significantly over time, while the T+ participants performed significantly different at the two POMs, $t(20) = 2.96$, $p < 0.01$, $d_{\text{Cohen}} = 0.69$. Regarding the concept map error ratio, the merely numerical values of the T+ participants' performance seemed to differ between the two POMs as well, but a Wilcoxon test showed this difference to be statistically insignificant, $z = -1.71$, $p = 0.09$, $n = 21$. The performance of the groups T++ and T-, on the other hand, did neither change visibly nor statistically significantly over time (see Figure 3 and Figure 4). These results overall indicate that, for the T+ group, the CM-related performance (operationalized via bQCM Index and concept map error ratio) in the later test phase was at a significantly lower level than in the learning phase one week before.

4. Discussion

We hypothesized that an extensive CM training makes learning with CM more effective in terms of more organizational and elaborative information processing, less perceived cognitive load, and better overall learning success (see Section 1.4). Considering our statistical analyses, this assumption, which is in line with theoretical considerations and empirical findings, is partially supported by our results.

4.1. Pre-Analyses of Baseline Differences (Measures Prior to Training Phase, Week 1)

To ensure valid conclusions from our results, we first checked for a balanced distribution of relevant sample characteristics. Our results showed comparability of the three groups in terms of age, sex, final overall school exam grade, and course of study at the university as well as prior knowledge about cell biology, even though the group assignment was based on self-selection for organizational reasons. Altogether, this allows for stating participants' equal initial conditions.

4.2. Treatment Check (Measure after Training Phase, Week 3)

To evaluate an immediate effect of the three different training conditions, we assessed the participants' declarative knowledge about the learning objective of intelligence after the training sessions, which, however, did not differ significantly between the three groups (see Section 2.4.2 and Section 3.2). Neither an extensive CM training nor the corresponding conditions had an impact on the immediate (no delay) retention/reproduction performance of our participants. On the one hand, this result could indicate, that training success is only assessable after a specific time delay. On the other hand, this result could also indicate that the acquisition of knowledge about the topic of intelligence probably took place in an intuitive way common to all students, in which they are accustomed to learning. Additionally, it is noticeable that neither CM nor the training method caused any interference in terms of additional (cognitive) burdens for short-term learning and reproduction.

4.3. Concept Mapping-Related Self-Efficacy and Error Detection Task (Measures Prior to Learning Phase, Week 4)

On the first day of the learning phase, participants in all three groups received the same 30-min short introduction to the CM strategy for two reasons: (1) to enable T⁻ participants to use the strategy during the following learning session, since they had not received any training on CM yet, and (2) to refresh CM-related knowledge of the other two groups T⁺⁺ and T⁺. Immediately afterwards, all participants were asked to judge their CM-related self-efficacy expectations and to solve an error detection and correction task using a prescribed error-including concept map (see Section 2.4.3).

The CM training had no significant effect on the error detection rate, while only the T⁺⁺ group showed significant advantages over the groups T⁺ and T⁻ when correcting errors. Furthermore, the T⁻ group tended to correct errors more falsely compared to the training groups. The fact that the participants of all three groups achieved comparable results regarding the error detection rate could be due to the fact that the task we used for its assessment and especially the presented error characteristics were not difficult enough to differentiate between groups. A more complex test, also including content-related errors next to formal ones, would probably have led to clearer results.

Furthermore, we had expected that the two CM training groups (T⁺⁺ and T⁺) would report higher levels of CM-related self-efficacy expectations than participants in the T⁻ group. Surprisingly, the opposite was found. The T⁻ group reported a slightly higher (although not significantly different) self-efficacy expectation than the two training groups, and the self-efficacy expectation in the T⁺⁺ group was the lowest (see Table 3). Based on our pretests, we

assume consistent comparability of our three groups, also with regard to their level of accuracy in metacognitive prediction (see Section 3.6) as well as their cognitive load (see Section 3.5). Therefore, both findings can be attributed to sample characteristics, since our student participants were experienced in learning and in adapting new learning strategies in different settings. It is therefore conceivable all participants generally had a high level of trust in their academic abilities and were therefore equally confident that they will be able to cope with using the CM strategy successfully. On the other hand, if the merely numerical differences regarding CM-related self-efficacy are taken seriously, it is reasonable to assume that the two groups T++ and T+ were already more familiar with the complexity of the CM strategy and, therefore, were less confident than participants of the T−group, who possibly put CM on the same level with less complex mind maps they might be familiar with, and thus, underestimated the actual requirements.

4.4. Recall, Organization, and Elaboration Processes (Measures during Learning Phase, Week 4)

Contrary to our hypothesis that an extensive CM training would have a positive effect on organization and elaboration processes during CM (see Section 1.4), our analysis did not reveal any substantial group differences in this regard. However, our results show that participants in all three groups achieved comparably good results, especially regarding the number of recall-suggesting propositions they specified within their concept maps (see Table 4). Similarly to the case of CM-related self-efficacy expectations, this result could be plausibly attributed to specific characteristics of our student sample, including solely experienced learner who perform comparably well. Furthermore, the learning texts on intelligence and cell biology were comparable to learning and study material on a university level in terms of formal as well as educational characteristics. Students who are skilled in conceptual and contextual learning will therefore probably only show random differences in making and keeping terms available, which are simply to be reproduced in a concept map. The construction of O- and E-propositions, requiring skills beyond simple reproduction, was only achieved at a comparatively marginal level by all participants, unassociated with any significant group differences. If it is also taken into account that the concept of learning implies a developmental characteristic, the question arises when and under which conditions meaningful learning associated with organization and elaboration can be expected to be successful. In the context of our study, it can be plausibly assumed that it might have been too early to expect such performance from our participants. The problem can also be seen in a rather traditional way in terms of a reminder for teachers that learners must be *picked up where they are* [156]. Post hoc, it is not possible to decide whether

our participants could have been overwhelmed by a corresponding task since they had not *explicitly* been given this task (to carry out specific cognitive processes during learning). Nevertheless, a further testable hypothesis would have to consider that, when introducing and using CM as a learning strategy, it could be crucial to make the respective learning and above all competence goals explicit and to prepare them theoretically and procedurally. In this way, the learners' support could be focused on all opportunities related to CM that go beyond a simple reproduction of what has been learned and which is also discussed as scientific literacy [157–162].

4.5. Cognitive Load (Measures during Learning Phase, Week 4)

Contrary to our hypothesis that training-related familiarity with the CM learning strategy differentially reduces the perceived cognitive load during the creation of a concept map (see Section 1.2 and Section 1.4), our analyses only revealed a substantial group difference regarding extraneous cognitive load (ECL), while the reported intrinsic (ICL) and germane cognitive load (GCL) did not statistically differ between the three groups (see Table 5). A higher ECL usually indicates a learning material's influence on participants' cognitive load. However, as all groups worked on the same learning material and, beyond that, the internal consistencies of the ECL and the ICL scales, were questionable, a significant difference or the lack of it should not be overestimated (see Section 2.4.4 and Section 4.8). Beyond that, the 30-min short introduction was probably sufficient enough to undermine a differentiating influence of a special training (see Section 4.2 and Section 4.3), if it is again taken into account that our sample consisted of experienced learners.

4.6. Facilitated Acquisition of Knowledge and Skills (Measures during as Well as after Learning Phase, Weeks 4, 5, and 6)

The central hypothesis of our study focused on the supportive effect of CM training on overall learning success (see Section 1.3 and Section 1.4). We experimentally chose the concept map quality as well as structural, declarative, and conceptual knowledge as indicators of learning success. For the domain of declarative knowledge, our results show effects by trend. Regarding the other three measures, we found clear and global effects of the training on the participants' performance. Significant advantages of the T++ group over the T- group occur regarding all relevant post hoc comparisons (with the exception of the declarative knowledge scores), which demonstrates that an extensive training including scaffolding and feedback elements is clearly more effective than a short introduction (see Table 7). In addition, the differences in structural (SJT score), declarative, and conceptual knowledge between the two training groups T++ and T+ were insignificant and only offer some trends regarding concept

map quality (aQCM Index). Considering this, it could be argued that a CM training could also be implemented with a focus on efficiency and thus less effort. However, this could be countered by the fact that the statistical insignificance of the assessed knowledge domains between the T++ and T+ groups can be related to a lack of clear requests to the participants of the T++ group to focus on the development of the specific knowledge represented by the aQCM Index or concept map quality during the test procedure. An investigation that addresses this requirement could contribute a crucial aspect to the learning-related functionality of CM. Above all, the significant decrease in CM-related performance of the T+ group from the learning to the test phase one week later should be considered in this regard (see Section 3.7 and Section 4.6.1).

4.6.1. Concept Map Quality

To determine concept map quality, the two dimensions bQCM Index and CM error rate of the learning and test phases were considered. An analysis of the learning development within the three groups regarding these two variables did not reveal any significant interaction effect but this result might be due to unequal variances between the groups. Accordingly, adjusted post hoc comparisons showed a clear advantage of the two training groups (T++ and T+) over the control group (T-), which contradicts the findings of Hay et al. [118], who stated that CM can be taught in a single session within 10–20 min. A comparison within the groups between the first and second POM showed a small decrease in performance except for a significant reduction of the bQCM Index and a tendency towards an increase of the concept map error rate in group T+ (see Table 9; Figure 3 and Figure 4). Looking closer at this finding and keeping in mind that the groups T++ and T+ do not significantly differ in terms of declarative, structural, and conceptual knowledge as well as concept map quality, it can be assumed that the CM-related knowledge advantage (bQCM Index and concept map error rate) of the T+ group, which it initially had in common with the T++ group compared to the T- group, was clearly lost again after just one week (see Table 9). How and when a stable CM competence can be expected from learners, therefore, seems to be a question about the interaction of training intensity, special training elements, and causal delay.

4.6.2. Structural Knowledge

The results on structural knowledge (SJT score) indicate a significant global difference between the three groups (see Table 7). Post hoc comparisons show that the two training groups (T++ and T+) performed almost identically, but both significantly better than the participants in the T- group. Thus, it seems plausible to suppose that an extensive CM training may facilitate

the acquisition of knowledge about the relationships between central terms of a field when using CM, whereas a short 30-min introduction to CM seems to be insufficient in this respect.

4.6.3. Declarative Knowledge

For the domain of declarative knowledge, our results show only effects by trend. Although both CM training groups (T++ and T+) performed numerically better than the control group (T-), our analyses revealed no statistically significant group differences. Overall, the T+ group performed best. Accordingly, CM training could promote retention performance somewhat better than a short introduction (see Table 7). However, regarding all three groups performing *equally badly*, we may suppose that neither CM by itself nor a respective training significantly affects the acquisition of declarative knowledge.

4.6.4. Conceptual Knowledge

In terms of conceptual knowledge, our results indicate a significant global group difference (see Table 7). Post hoc comparisons reveal a comparable pattern of results as in the case of structural knowledge: While the two training groups (T++ and T+) performed almost identically, they each performed significantly better than the participants of the T- group. Accordingly, for the acquisition of conceptual knowledge, it can be assumed that familiarity with the CM strategy caused by training is a crucial and beneficial condition. Thus, a 30-min short introduction to the CM strategy seems insufficient to enable learners to effectively acquire conceptual knowledge using CM. When learning a topic/text, CM can make the learner aware of conceptual gaps. Incorrect relationships may have been recognized or learners realized that they were simply unable to integrate terms meaningfully into an existing structure. The supportive effect in the development of this metacognitive competence is an argument for the benefits of CM [163,164]. As part of a follow-up investigation, it would therefore be of interest to have a closer look at the question of which metacognitive skills are enhanced in which way by using CM, and which ones could have contributed to the advantages of our T++ group in particular.

4.6.5. Metacognitive Prediction

At the end of the learning phase on cell biology, we asked our participants to estimate how much of what they had just learned they would remember in the test phase one week later. This judgment of learning served as a measure of the participants' confidence in the success of their efforts (see Section 4.6). Interestingly, our corresponding data analyses did not indicate any significant differences. It is plausible to assume that the accurate fitting of the participants' skills to the experimentally realized requirements regarding the topic and the learning strategy

did not allow for significant differences due to different training conditions as well as with regard to the judgment of learning. Regarding our student sample, we consider that the CM learning strategy simply represented an addition, which was apparently easy to assimilate for our participants to the already existing learning experiences and corresponding skills. This homogenizing factor, in addition to the used learning material which was comparable to the material they already were familiar with, could have undermined larger differences between groups regarding their metacognitive prediction.

4.7. Summarizing Discussion and Implications

CM is effective for learning and can also promote in-depth knowledge acquisition in terms of flexibly transferable knowledge [50,60,67,76,78]. Students can use CM as a learning strategy to visualize complex relationships by organizing and connecting central concepts with each other and probably with those out of their prior knowledge. For educators, CM represents a teaching format to summarize important content, and a diagnostic tool to assess students' knowledge of a domain since concept map construction requires conceptual knowledge and its result reflects the quality of knowledge organization [80,113]. However, our results show that concept maps should only be used as a diagnostic tool if it has been trained adequately beforehand as the development of conceptual knowledge was positively influenced by CM training (see Table 7) when CM is used as a learning tool. In addition, practicing CM as a learning strategy must be preceded by a differential educational evaluation of learners and the topic as well as the learning material and objective.

Our findings, therefore, support those of studies by Ajaja [128], Kinchin [165] and Jegede et al. [110], who state that more experience in CM leads to significant and consistent improvement in achievement. In particular, the practical exercise elements during training seem to be decisive for the development of a high-level CM skills, as the participants in our T- group, who had only received a theoretical short introduction to CM, performed significantly worse in this respect. This finding is in line with those of Brandstädter et al. [22], Sumfleth et al. [69], Nesbit and Adesope [67], and Horton et al. [36]. The use of prompts and feedback in an extensive CM training has not been attempted before although has been proven to be effective in studies [66,120,166,167]. Our results show that the T+ group achieved as good performance in many cases as the T++ group, which additionally received time-consuming scaffolding and feedback, so it could be argued that a CM training could also be implemented with less effort, but in this respect, it should not be missed that there was a significant decrease in CM-related performance of the T+ group from the learning to the test phase just one week later (see Section 3.7 and Section 4.6.1). Finally, the extent of a specific training above a 30-min introduction

showed no significant effect on the acquisition of declarative knowledge, but this result must be interpreted considering the characteristics and purpose of CM, since the strategy focuses more on acquisition of structural and conceptual knowledge.

4.8. Limitations

Even though our results provide substantial clarification, they need to be evaluated in light of the study's limitations:

- (1) Regarding our student sample, we assume a high homogeneity in terms of a high degree of learning experience and academic performance. Therefore, it might be possible that the CM learning strategy was quite easy to assimilate for our participants to their already existing learning strategy repertoire. This homogenizing factor, in addition to the applied learning material, which was comparable to the material they already were familiar with, could have undermined some group differences on dependent variables regarding the three different training conditions. In this regard, it could be expected that referring to other samples than experienced learners could reveal more distinct group differences. Nevertheless, our total sample size of $N = 73$ was simply too small to draw conclusions in terms of external validity, so we recommend interpreting our results found for our small student sample size with caution.
- (2) Furthermore, the number of $N = 73$ participants in our study was not large enough to reach sufficient absolute frequency. Accordingly, a parallelization of the groups to which the participants had assigned themselves could not be achieved to our complete satisfaction, even if no initial differences could be shown. This is directly related to the power determined for our data analysis, which did not exceed about 0.30 for correlational and about 0.45 for ANOVA testing. To ensure that our expected small to medium-sized effects could be detected at a power level of 0.80 to 0.90, our sample should have included approximately $N = 150$ participants. Despite the desirability of such optimized test conditions, it seems difficult to imagine how the required intensive care of such a large number of participants could have been ensured over a six-week study period, given limited personnel and financial resources.
- (3) The instrument used to assess the participants' ability to detect errors in a prescribed concept map (see Section 2.4.3 and Section 3.3) was apparently too easy for all three groups. A more complex test, presenting formal as well as content-related errors at different levels of difficulty, could have led to clearer results.

At least, we may take the liberty of noting that, once the limitations of our study are removed in future testing, we expect our trend-wise group differences to have a higher probability of becoming statistically significant.

4.9. Prospects for Future Research

On a general level, our study was able to replicate existing findings showing that CM is an effective learning tool [34,36–38,42,59,67,122]. Beyond that, our results indicate that the overall learning success can be improved by an extensive CM training and that stable CM strategy skills in particular can be promoted by additional integration of scaffolding and feedback elements in the course of such a training. However, the question of the interaction of training intensity, special training elements, and causal delay remains unanswered in our study, since our small sample size and the homogeneity of the participants do simply not allow for giving a specific reliable recommendation in this regard. Further investigations could therefore focus on answering this question first, since it would be helpful in terms of efficiency if it could be determined which specific components of a training are particularly effective in learning of different learners and to what extent.

In particular, the learner's intrinsic motivation could be considered as a covariate in this regard, as it can be assumed that the learners' potential is directly related to it [168,169], so it is conceivable that training effectiveness increases by intentionally stimulating motivational aspects. In this connection, it would also be of interest to have a closer look at the question of which metacognitive skills are enhanced in which way by using CM, and which ones could have contributed to the advantages of our T++ group in particular.

Another question in this regard is how it could be ensured that CM competence, which was developed within a certain field (e.g., biology), can be handled flexibly by learners in terms of thematic transfers to other fields. Since such a transfer of a learning strategy primarily requires a comprehensive availability of strategy-related knowledge, it must also be taken into account how often and how intensively corresponding training sessions have to be implemented to stabilize the procedural knowledge on concept map use over time.

Finally, our results concerning the domain of declarative knowledge pose a specific question that should be answered in future studies. Since all of our three groups performed equally badly, we supposed that CM itself could be less suitable to improve the acquisition of knowledge of facts (which is plausible, since the strategy focuses rather on structural and conceptual aspects of knowledge). In order to combine the positive effects of CM on the other two domains of knowledge with a promotion of declarative knowledge, it could be advisable to practice the CM strategy in different settings of learning which are proven to specifically

promote the acquisition of different domains of knowledge. Further explanatory perspectives on the potentially relevant addition of retrieval practice as such a setting condition are given by Karpicke and Blunt [147] as well as Karpicke [170,171].

Supplementary Materials: The following material is available online at <https://www.mdpi.com/article/10.3390/educsci11080438/s1>. Section S1. Coding Manual for the Preliminary Questionnaire; Section S2. Scaffolds, Prompts, and Feedback for Group T++; Section S3. Scaffolds and Prompts for Group T-; Section S4. Coding Manual for the Treatment Check after the Training Phase; Section S5. Coding Manual for the Concept Mapping Self-Efficacy Questionnaire; Section S6. CM Familiarity Treatment Check; Section S7. Learning Booklet for the Individual Learning Phase; Section S8. Coding Manual for the Cognitive Load Questionnaire; Section S9. Coding Manual for the Structural Knowledge Test; Section S10. Coding Manual for the Declarative Knowledge Test; Section S11. Coding Manual for the Conceptual Knowledge Test.

Author Contributions: Conceptualization, J.G., L.B.B. and V.D.E.W.; Methodology, J.G., L.B.B. and V.D.E.W.; Formal Analysis, J.G., L.B.B. and V.D.E.W.; Investigation, L.B.B.; Resources, J.G., E.A. and V.D.E.W.; Data Curation, L.B.B. and V.D.E.W.; Writing—Original Draft Preparation, L.B.B. and V.D.E.W.; Writing—Review and Editing, J.G., E.A., L.B.B., and V.D.E.W.; Visualization, L.B.B. and V.D.E.W.; Supervision, J.G.; Project Administration, J.G.; Funding Acquisition, J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Research Foundation (DFG), grant number GR 4763/2-1.

Institutional Review Board Statement: Our study was approved by the German Research Foundation (DFG; project number: GR 4763/2-1). Before participation, all subjects received a detailed written subject information in accordance with the current ethical guidelines laid down by the University of Cologne (Germany) and the German Psychological Society [172], including the following information: aims and course of the investigation, absolute voluntariness of participation, possibility of dropping out of participation at any time, guaranteed protection of data privacy (collection of only anonymized data), possibility of requesting data cancelation at any time, no-risk character of study participation, and contact information in case of any questions or problems. Written informed consent was obtained from all participants prior to the study, according to the current version of the Declaration of Helsinki

[173,174]. Anonymity was ensured by using individual codes that the students created themselves. Data storage meets current European data protection regulations [175].

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Dahnke, H.; Fuhrmann, A.; Steinhagen, K. Entwicklung und Einsatz von Computersimulation und Concept Mapping als Erhebungsinstrumente bei Vorstellungen zur Wärmephysik eines Hauses [Development and use of computer simulation and concept mapping as survey instruments for ideas on the thermal physics of a house]. *Z. Didakt. Nat.* 1998, *4*, 67–80.
2. Hegarty-Hazel, E.; Prosser, M. Relationship between students' conceptual knowledge and study strategies—Part 1: Student learning in physics. *Int. J. Sci. Educ.* 1991, *13*, 303–312.
3. Perkins, D.N.; Grotzer, T.A. Dimensions of causal understanding: The role of complex causal models in students' understanding of science. *Stud. Sci. Educ.* 2005, *41*, 117–165.
4. Peuckert, J.; Fischler, H. Concept Maps als Diagnose- und Auswertungsinstrument in einer Studie zur Stabilität und Ausprägung von Schülervorstellungen [Concept Maps as a diagnostic and evaluation tool in a study on the stability and expression of student perceptions]. In *Concept Mapping in Fachdidaktischen Forschungsprojekten der Physik und Chemie [Concept Mapping in Didactic Research Projects in Physics and Chemistry]*; Fischler, H., Peuckert, J., Eds.; Studien zum Physiklernen; Logos: Berlin, Germany, 2000; Volume 1, pp. 91–116.
5. Grüß-Niehaus, T. *Zum Verständnis des Löslichkeitskonzeptes im Chemieunterricht—Der Effekt von Methoden Progressiver und Kollaborativer Reflexion [To Understand the Concept of Solubility in Chemistry Lessons—The Effect of Methods of Progressive and Collaborative Reflection]*; Logos: Berlin, Germany, 2010.

6. Markow, P.G.; Lonning, R.A. Usefulness of concept maps in college chemistry laboratories: Students' perceptions and effects on achievement. *J. Res. Sci. Teach.* 1998, *35*, 1015–1029.
7. Stracke, I. *Einsatz Computerbasierter Concept Maps zur Wissensdiagnose in der Chemie: Empirische Untersuchungen am Beispiel des Chemischen Gleichgewichts [Use of Computer-Based Concept Maps for Knowledge Diagnosis in Chemistry: Empirical Studies Using the Example of Chemical Equilibrium]*; Waxmann: Münster, Germany, 2004.
8. Assaraf, O.B.-Z.; Dodick, J.; Tripto, J. High school students' understanding of the human body system. *Res. Sci. Educ.* 2013, *43*, 33–56.
9. Bramwell-Lalor, S.; Rainford, M. The effects of using concept mapping for improving advanced level biology students' lower- and higher-order cognitive skills. *Int. J. Sci. Educ.* 2014, *36*, 839–864.
10. Feigenspan, K.; Rayder, S. Systeme und systemisches Denken in der Biologie und im Biologieunterricht [Systems and systemic thinking in biology and biology teaching]. In *Systemisches Denken im Fachunterricht [Systemic Thinking in School Subjects]*; Arndt, H., Ed.; FAU Lehren und Lernen; FAU University Press: Erlangen, Germany, 2017; pp. 139–176.
11. Hegarty-Hazel, E.; Prosser, M. Relationship between students' conceptual knowledge and study strategies—Part 2: Student learning in biology. *Int. J. Sci. Educ.* 1991, *13*, 421–429.
12. Heinze-Fry, J.A.; Novak, J.D. Concept mapping brings long-term movement toward meaningful learning. *Sci. Educ.* 1990, *74*, 461–472.
13. Kinchin, I.M.; De-Leij, F.A.A.M.; Hay, D.B. The evolution of a collaborative concept mapping activity for undergraduate microbiology students. *J. Furth. High. Educ.* 2005, *29*, 1–14.
14. Schmid, R.F.; Telaro, G. Concept mapping as an instructional strategy for high school biology. *J. Educ. Res.* 1990, *84*, 78–85.
15. Tripto, J.; Assaraf, O.B.; Snapir, Z.; Amit, M. How is the body's systemic nature manifested amongst high school biology students? *Instr. Sci.* 2017, *45*, 73–98.
16. Sadava, D.E.; Hillis, D.M.; Heller, H.C. *Purves Biologie [Purves Biology]*, 10th ed.; Springer: Berlin/Heidelberg, Germany, 2019.
17. Urry, L.A.; Cain, M.L.; Wasserman, S.A.; Minorsky, P.V.; Reece, J.B. *Campbell Biologie [Campbell Biology]*, 11th ed.; Pearson: Hallbergmoos, Germany, 2019.

18. Großschedl, J.; Harms, U. Effekte metakognitiver Prompts auf den Wissenserwerb beim Concept Mapping und Notizen Erstellen [Effects of metacognitive prompts on knowledge acquisition in concept mapping and note taking]. *Z. Didakt. Nat.* 2013, *19*, 375–395.
19. Dreyfus, A.; Jungwirth, E. The pupil and the living cell: A taxonomy of dysfunctional ideas about an abstract idea. *J. Biol. Educ.* 1989, *23*, 49–55.
20. Flores, F.; Tovar, M.E.; Gallegos, L. Representation of the cell and its processes in high school students: An integrated view. *Int. J. Sci. Educ.* 2003, *25*, 269–286.
21. Anderman, E.M.; Sinatra, G.M.; Gray, D.L. The challenges of teaching and learning about science in the twenty-first century: Exploring the abilities and constraints of adolescent learners. *Stud. Sci. Educ.* 2012, *48*, 89–117.
22. Brandstädter, K.; Harms, I.; Großschedl, J. Assessing system thinking through different concept-mapping practices. *Int. J. Sci. Educ.* 2012, *34*, 2147–2170.
23. Dekker, S. *Drift into Failure: From Hunting Broken Components to Understanding Complex Systems*; CRC Press: London, UK, 2011.
24. Grotzer, T.A.; Solis, S.L.; Tutwiler, M.S.; Cuzzolino, M.P. A study of students' reasoning about probabilistic causality: Implications for understanding complex systems and for instructional design. *Instr. Sci.* 2017, *45*, 25–52.
25. Hashem, K.; Mioduser, D. Learning by Modeling (LbM): Understanding complex systems by articulating structures, behaviors, and functions. *Int. J. Adv. Comput. Sci. Appl.* 2013, *4*, 80–86.
26. Hmelo-Silver, C.E.; Azevedo, R. Understanding complex systems: Some core challenges. *J. Learn. Sci.* 2006, *15*, 53–61.
27. Jacobson, M.J.; Wilensky, U. Complex systems in education: Scientific and educational importance and implications for the learning sciences. *J. Learn. Sci.* 2006, *15*, 11–34.
28. Plate, R. Assessing individuals' understanding of nonlinear causal structures in complex systems. *Syst. Dyn. Rev.* 2010, *26*, 19–33.
29. Proctor, R.W.; van Zandt, T. *Human Factors in Simple and Complex Systems*, 3rd ed.; CRC Press Taylor & Francis Group: Boca Raton, FL, USA, 2018.
30. Raia, F. Students' understanding of complex dynamic systems. *J. Geosci. Educ.* 2005, *53*, 297–308.
31. Yoon, S.A.; Anderson, E.; Koehler-Yom, J.; Evans, C.; Park, M.; Sheldon, J.; Schoenfeld, I.; Wendel, D.; Scheintaub, H.; Klopfer, E. Teaching about complex

- systems is no simple matter: Building effective professional development for computer-supported complex systems instruction. *Instr. Sci.* 2017, 45, 99–121.
32. Friedrich, H.F.; Mandl, H. Lernstrategien: Zur Strukturierung des Forschungsfeldes [Learning strategies: On the structuring of the research field]. In *Handbuch Lernstrategien [Handbook of Learning Strategies]*; Mandl, H., Friedrich, H.F., Eds.; Hogrefe: Göttingen, Germany, 2006; pp. 1–23.
 33. Arndt, H. (Ed.) *Systemisches Denken im Fachunterricht [Systemic Thinking in School Subjects]*; FAU University Press: Erlangen, Germany, 2017; p. 369.
 34. Cañas, A.; Coffey, J.; Carnot, M.J.; Feltovich, P.; Hoffman, R.R.; Feltovich, J.; Novak, J.D. A Summary of Literature Pertaining to the Use of Concept Mapping Techniques and Technologies for Education and Performance Support. Available online: <https://www.ihmc.us/users/acanas/Publications/ConceptMapLitReview/IHMC%20Literature%20Review%20on%20Concept%20Mapping.pdf> (accessed on 11 May 2021).
 35. Großschedl, J.; Harms, U. Metakognition—Dirigentin des Gedankenkonzerts [Metacognition—Conductor of the concert of thoughts]. In *Biologie Methodik [Biology Methodology]*, 2nd ed.; Spörhase-Eichmann, U., Ruppert, W., Eds.; Cornelsen Scriptor: Berlin, Germany, 2014; pp. 48–52.
 36. Horton, P.B.; McConney, A.A.; Gallo, M.; Woods, A.L.; Senn, G.J.; Hamelin, D. An investigation of the effectiveness of concept mapping as an instructional tool. *Sci. Educ.* 1993, 77, 95–111.
 37. Novak, J.D. The pursuit of a dream: Education can be improved. In *Teaching Science for Understanding*; Educational Psychology Series; Mintzes, J.J., Wandersee, J.H., Novak, J.D., Eds.; Elsevier Academic Press: San Diego, CA, USA, 2005; pp. 3–28.
 38. Schroeder, N.L.; Nesbit, J.C.; Anguiano, C.J.; Adesope, O.O. Studying and constructing concept maps: A meta-analysis. *Educ. Psychol. Rev.* 2018, 30, 431–455.
 39. Cañas, A.; Carff, R.; Hill, G.; Carvalho, M.; Arguedas, M.; Eskridge, T.C.; Lott, J.; Carvajal, R. Concept maps: Integrating knowledge and information visualization. In *Knowledge and Information Visualization: Searching for Synergies*; Tergan, S.-O., Keller, T., Eds.; Springer: Berlin, Germany, 2005; pp. 205–219.
 40. Feller, W.; Spörhase, U. Innere Differenzierung durch Experten-Concept-Maps [Internal differentiation through expert concept maps]. In *Biologie Methodik [Biology Methodology]*, 4th ed.; Spörhase-Eichmann, U., Ruppert, W., Eds.; Cornelsen: Berlin, Germany, 2018; pp. 77–81.

41. Großschedl, J.; Harms, U. Metakognition—Denken aus der Vogelperspektive [Metacognition—Thinking from a bird’s eye view]. In *Biologie Methodik [Biology Methodology]*, 4th ed.; Spörhase-Eichmann, U., Ruppert, W., Eds.; Cornelsen: Berlin, Germany, 2018; pp. 48–52.
42. Haugwitz, M.; Nesbit, J.C.; Sandmann, A. Cognitive ability and the instructional efficacy of collaborative concept mapping. *Learn. Individ. Diff.* 2010, *20*, 536–543.
43. Mehren, R.; Rempfler, A.; Ulrich-Riedhammer, E.M. Die Anbahnung von Systemkompetenz im Geographieunterricht [The initiation of system competence in geography lessons]. In *Systemisches Denken im Fachunterricht [Systemic Thinking in School Subjects]*; Arndt, H., Ed.; FAU Lehren und Lernen; FAU University Press: Erlangen, Germany, 2017. [Google Scholar]
44. Michalak, M.; Müller, B. Durch Sprache zum systemischen Denken [Through language to systemic thinking]. In *Systemisches Denken im Fachunterricht [Systemic Thinking in School Subjects]*; Arndt, H., Ed.; FAU Lehren und Lernen; FAU University Press: Erlangen, Germany, 2017; pp. 111–138.
45. Novak, J.D.; Cañas, A. The origins of the concept mapping tool and the continuing evolution of the tool. *Inf. Vis.* 2006, *5*, 175–184.
46. Novak, J.D.; Cañas, A.J. Theoretical origins of concept maps, how to construct them, and uses in education. *Reflecting Educ.* 2007, *3*, 29–42.
47. Novak, J.D.; Cañas, A. *The Theory Underlying Concept Maps and How to Construct Them*; Florida Institute for Human and Machine Cognition: Pensacola, FL, USA, 2008; Available online: <http://cmap.ihmc.us/docs/pdf/TheoryUnderlyingConceptMaps.pdf> (accessed on 11 April 2021).
48. Novak, J.D.; Gowin, D.B. *Learning How to Learn*; Cambridge University Press: New York, NY, USA, 1984.
49. van Zele, E.; Lenaerts, J.; Wieme, W. Improving the usefulness of concept maps as a research tool for science education. *Int. J. Sci. Educ.* 2004, *26*, 1043–1064.
50. Cañas, A.J.; Novak, J.D. Concept mapping using CmapTools to enhance meaningful learning. In *Knowledge Cartography*, 2nd ed.; Advanced Information and Knowledge Processing; Okada, A., Ed.; Springer: London, UK, 2014; pp. 23–45.
51. Cronin, P.J.; Dekkers, J.; Dunn, J.G. A procedure for using and evaluating concept maps. *Res. Sci. Educ.* 1982, *12*, 17–24.

52. Novak, J.D. Concept mapping: A useful tool for science education. *J. Res. Sci. Teach.* 1990, 27, 937–949.
53. Novak, J.D. *Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations*, 2nd ed.; Routledge: New York, NY, USA, 2010.
54. Marzetta, K.; Mason, H.; Wee, B. ‘Sometimes They Are Fun and Sometimes They Are Not’: Concept Mapping with English Language Acquisition (ELA) and Gifted/Talented (GT) Elementary Students Learning Science and Sustainability. *Educ. Sci.* 2018, 8, 13.
55. Ullah, A.M.M.S. Fundamental Issues of Concept Mapping Relevant to Discipline-Based Education: A Perspective of Manufacturing Engineering. *Educ. Sci.* 2019, 9, 228.
56. Beier, M.; Brose, B.; Gemballa, S.; Heinze, J.; Knerich, H.; Kronberg, I.; Küttner, R.; Markl, J.S.; Michiels, N.; Nolte, M.; et al. *Markl Biologie: Oberstufe [Markl Biology: Upper School]*; Klett: Stuttgart, Germany, 2018.
57. Rost, F. *Lern-und Arbeitstechniken für das Studium [Learning and Working Techniques for University Studies]*; Springer Fachmedien: Wiesbaden, Germany, 2018.
58. Kinchin, I.M. If concept mapping is so helpful to learning biology, why aren’t we all doing it? *Int. J. Sci. Educ.* 2001, 23, 1257–1269.
59. Machado, C.T.; Carvalho, A.A. Concept mapping: Benefits and challenges in higher education. *J. Contin. High. Educ.* 2020, 68, 38–53.
60. Hilbert, T.; Renkl, A. Concept mapping as a follow-up strategy to learning from texts: What characterizes good and poor mappers? *Instr. Sci.* 2008, 36, 53–73.
61. Tergan, S.-O. Digital concept maps for managing knowledge and information. In *Knowledge and Information Visualization*; Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J.M., Mattern, F., Mitchell, J.C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., et al., Eds.; Lecture Notes in Computer Science; Springer: Heidelberg, Germany, 2005; Volume 3426, pp. 185–204.
62. Roessger, K.M.; Daley, B.J.; Hafez, D.A. Effects of teaching concept mapping using practice, feedback, and relational framing. *Learn. Instr.* 2018, 54, 11–21.
63. Ausubel, D.P. *The Psychology of Meaningful Verbal Learning*; Grune & Stratton: New York, NY, USA, 1963.
64. Ausubel, D.P. *The Acquisition and Retention of Knowledge: A Cognitive View*; Springer: Dordrecht, The Netherlands, 2000.
65. Novak, J.D. *A Theory of Education*; Cornell University Press: Ithaca, NY, USA, 1977.

66. Hilbert, T.; Nückles, M.; Renkl, A.; Minarik, C.; Reich, A.; Ruhe, K. Concept Mapping zum Lernen aus Texten [Concept mapping for learning from texts]. *Z. Pädagogische Psychol.* 2008, 22, 119–125.
67. Nesbit, J.C.; Adesope, O.O. Learning with concept and knowledge maps: A meta-analysis. *Rev. Educ. Res.* 2006, 76, 413–448.
68. Renkl, A.; Nückles, M. Lernstrategien der externen Visualisierung [Learning strategies of external visualization]. In *Handbuch Lernstrategien [Handbook of Learning Strategies]*; Mandl, H., Friedrich, H.F., Eds.; Hogrefe: Göttingen, Germany, 2006; pp. 135–147.
69. Sumfleth, E.; Neuroth, J.; Leutner, D. Concept Mapping—Eine Lernstrategie muss man lernen. Concept Mapping—Learning Strategy is Something You Must Learn. *CHEMKON* 2010, 17, 66–70.
70. Weinstein, C.E.; Mayer, R.E. The teaching of learning strategies. In *Handbook of Research on Teaching*, 3rd ed.; Wittrock, M.C., Ed.; Macmillan: New York, NY, USA, 1986; pp. 315–327.
71. Cox, R. Representation construction, externalised cognition and individual differences. *Learn. Instr.* 1999, 9, 343–363.
72. VanLehn, K.; Jones, R.M.; Chi, M.T.H. A model of the self-explanation effect. *J. Learn. Sci.* 1992, 2, 1–59.
73. Hilbert, T.; Nückles, M.; Matzel, S. Concept mapping for learning from text: Evidence for a worked-out-map-effect. In *International Perspectives in the Learning Sciences: Creating a Learning World, Proceedings of the Eighth International Conference on International Conference for the Learning Sciences—ICLS 2008, Utrecht, The Netherlands, 24–28 June 2008*; Kanselaar, G., Jonker, V., Kirschner, P.A., Prins, F.J., Eds.; International Society of the Learning Sciences: Utrecht, The Netherlands, 2008; pp. 358–365.
74. Baddeley, A. Working memory: Theories, models, and controversies. *Annu. Rev. Psychol.* 2012, 63, 1–29.
75. Hasselhorn, M.; Gold, A. *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren [Educational Psychology: Successful Learning and Teaching]*, 4th ed.; Kohlhammer: Stuttgart, Germany, 2017.
76. Cadarin, L.; Bagnasco, A.; Rocco, G.; Sasso, L. An integrative review of the characteristics of meaningful learning in healthcare professionals to enlighten educational practices in health care. *Nurs. Open* 2014, 1, 3–14.

77. Hattie, J. *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*; Routledge: Abingdon, UK, 2008; p. 378.
78. Haugwitz, M. Kontextorientiertes Lernen und Concept Mapping im Fach Biologie: Eine Experimentelle Untersuchung zum Einfluss auf Interesse und Leistung unter Berücksichtigung von Moderationseffekten Individueller Voraussetzungen beim Kooperativen Lernen [Context-oriented Learning and Concept Mapping in Biology. An Experimental Investigation of the Influence on Interest and Performance, Taking into Account Moderation Effects of Individual Prerequisites in Cooperative Learning]. Ph.D. Thesis, University of Duisburg-Essen, Duisburg, Germany, 2009.
79. Gunstone, R.F.; Mitchell, I.J. Metacognition and conceptual change. In *Teaching Science for Understanding*; Educational Psychology Series; Mintzes, J.J., Wandersee, J.H., Novak, J.D., Eds.; Elsevier Academic Press: San Diego, CA, USA, 2005; pp. 133–163.
80. Ruiz-Primo, M.A. Examining concept maps as an assessment tool. In *Concept Maps: Theory, Methodology, Technology*; Cañas, A., Novak, J.D., González, F.M., Eds.; Universidad Pública de Navarra: Pamplona, Spain, 2004; pp. 555–563.
81. McClelland, J.L.; Rumelhart, D.E.; Hinton, G.E. The appeal of parallel distributed processing. In *Explorations in the Microstructure of Cognition*; Rumelhart, D.E., McClelland, J.L., the PDP Research Group, Eds.; Parallel distributed processing; Bradford: Cambridge, MA, USA, 1986; pp. 3–44.
82. Thurn, C.M.; Hänger, B.; Kokkonen, T. Concept mapping in magnetism and electrostatics: Core concepts and development over time. *Educ. Sci.* 2020, *10*, 129.
83. Amadiou, F.; Van Gog, T.; Paas, F.; Tricot, A.; Mariné, C. Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learn. Instr.* 2009, *19*, 376–386.
84. Hailikari, T.; Katajavuori, N.; Lindblom-Ylänne, S. The relevance of prior knowledge in learning and instructional design. *Am. J. Pharm. Educ.* 2008, *72*, 113.
85. Krause, U.-M.; Stark, R. Vorwissen aktivieren [Activate prior knowledge]. In *Handbuch Lernstrategien [Handbook of Learning Strategies]*; Mandl, H., Friedrich, H.F., Eds.; Hogrefe: Göttingen, Germany, 2006; pp. 38–49.
86. Rouet, J.-F. Managing cognitive load during document-based learning. *Learn. Instr.* 2009, *19*, 445–450.
87. Sedumedi, T.D.T. Prior knowledge in chemistry instruction: Some insights from students' learning of ACIDS/BASES. *Psycho-Educ. Res. Rev.* 2013, *2*, 34–53.

88. Strohner, H. *Kognitive Systeme: Eine Einführung in die Kognitionswissenschaft [Cognitive Systems: An Introduction to Cognitive Science]*; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 1995.
89. Friedrich, H.F. Vermittlung von reduktiven Textverarbeitungsstrategien durch Selbstinstruktion [Teaching reductive text processing strategies through self-instruction]. In *Lern-und Denkstrategien: Analyse und Intervention [Learning and Thinking Strategies: Analysis and Intervention]*; Mandl, H., Friedrich, H.F., Eds.; Hogrefe: Göttingen, Germany, 1992; pp. 193–212.
90. Kalyuga, S. Expertise reversal effect and its implications for learner-tailored instruction. *Educ. Psychol. Rev.* 2007, *19*, 509–539.
91. Kalyuga, S.; Ayres, P.; Chandler, P.; Sweller, J. The expertise reversal effect. *Educ. Psychol.* 2003, *38*, 23–31.
92. Cronbach, L.J. The two disciplines of scientific psychology. *Am. Psychol.* 1957, *12*, 671–684.
93. Cronbach, L.J. Beyond the two disciplines of scientific psychology. *Am. Psychol.* 1975, *30*, 116–127.
94. Sweller, J. Implications of cognitive load theory for multimedia learning. In *The Cambridge Handbook of Multimedia Learning*; Mayer, R.E., Ed.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2005; pp. 19–30.
95. Debie, N.; van de Leemput, C. What does germane load mean? An empirical contribution to the cognitive load theory. *Front. Psych.* 2014, *5*, 1099.
96. Klepsch, M.; Schmitz, F.; Seufert, T. Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Front. Psych.* 2017, *8*, 1–18.
97. Sweller, J.; van Merriënboer, J.; Paas, F. Cognitive architecture and instructional design. *Educ. Psychol. Rev.* 1998, *10*, 251–296.
98. Bannert, M. Effekte metakognitiver Lernhilfen auf den Wissenserwerb in vernetzten Lernumgebungen [Effects of metacognitive learning aids on knowledge acquisition in networked learning environments]. *Z. Pädagogische Psychol.* 2003, *17*, 13–25.
99. Jüngst, K.L.; Strittmatter, P. Wissensstrukturdarstellung: Theoretische Ansätze und praktische Relevanz [Knowledge structure representation: Theoretical approaches and practical relevance]. *Unterrichtswissenschaft* 1995, *23*, 194–207.
100. Mintzes, J.J.; Cañas, A.; Coffey, J.; Gorman, J.; Gurley, L.; Hoffman, R.; McGuire, S.Y.; Miller, N.; Moon, B.; Trifone, J.; et al. Comment on “Retrieval practice produces

- more learning than elaborative studying with concept mapping”. *Science* 2011, 334, 453.
101. Schwendimann, B. Concept Mapping. In *Encyclopedia of Science Education*; Gunstone, R., Ed.; Springer: Dordrecht, The Netherlands, 2015; pp. 198–202.
 102. Jonassen, D.H.; Beissner, K.; Yacci, M. *Structural Knowledge: Techniques for Representing, Conveying, and Acquiring Structural Knowledge*; Lawrence Erlbaum: Hillsdale, NJ, USA, 1993.
 103. Anderson, J.R. *Language, Memory, and Thought*; Lawrence Erlbaum: Hillsdale, NJ, USA, 1976.
 104. McCormick, R. Conceptual and procedural knowledge. *Int. J. Technol. Des. Educ.* 1997, 7, 141–159.
 105. Renkl, A. Wissenserwerb [Knowledge Acquisition]. In *Pädagogische Psychologie [Educational Psychology]*, 3rd ed.; Wild, E., Möller, J., Eds.; Springer: Berlin, Germany, 2020; pp. 3–24.
 106. Moors, A.; de Houwer, J. Automaticity: A theoretical and conceptual analysis. *Psychol. Bull.* 2006, 132, 297–326.
 107. Mietzel, G. *Pädagogische Psychologie des Lernens und Lehrens [Educational Psychology of Learning and Teaching]*, 9th ed.; Hogrefe: Göttingen, Germany, 2017.
 108. Haugwitz, M.; Sandmann, A. Kooperatives Concept Mapping in Biologie: Effekte auf den Wissenserwerb und die Behaltensleistung. [Cooperative concept mapping in biology: Effects on knowledge acquisition and retention]. *Z. Didakt. Nat.* 2009, 15, 89–107.
 109. Hilbert, T.S.; Renkl, A. Individual differences in concept mapping when learning from texts. In *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, Stresa, Italy, 21–23 July 2005; Bara, B.G., Barsalou, L., Bucciarelli, M., Eds.; Lawrence Erlbaum: Mahwah, NJ, USA, 2005; pp. 947–952.
 110. Jegede, O.J.; Alaiyemola, F.F.; Okebukola, P.A.O. The effect of concept mapping on students’ anxiety and achievement in biology. *J. Res. Sci. Teach.* 1990, 27, 951–960.
 111. McCagg, E.C.; Dansereau, D.F. A convergent paradigm for examining knowledge mapping as a learning strategy. *J. Educ. Res.* 1991, 84, 317–324.
 112. Romero, C.; Cazorla, M.; Buzón, O. Meaningful learning using concept maps as a learning strategy. *J. Technol. Sci. Educ.* 2017, 7, 313.
 113. Morse, D.; Jutras, F. Implementing concept-based learning in a large undergraduate classroom. *CBE Life Sci. Educ.* 2008, 7, 243–253.

114. Ajaja, O.P. Which way do we go in biology teaching? Lecturing, concept mapping, cooperative learning or learning cycle? *Electron. J. Sci. Educ.* 2013, 17, 1–37.
115. Chiu, C.-H. Evaluating system-based strategies for managing conflict in collaborative concept mapping. *J. Comput. Assist. Learn.* 2004, 20, 124–132.
116. Chularut, P.; DeBacker, T.K. The influence of concept mapping on achievement, self-regulation, and self-efficacy in students of English as a second language. *Contemp. Educ. Psychol.* 2004, 29, 248–263.
117. Hay, D.; Kinchin, I.; Lygo-Baker, S. Making learning visible: The role of concept mapping in higher education. *Stud. High. Educ.* 2008, 33, 295–311.
118. Hilbert, T.; Renkl, A. Learning how to use a computer-based concept-mapping tool: Self-explaining examples helps. *Comput. Hum. Behav.* 2009, 25, 267–274.
119. Quinn, H.J.; Mintzes, J.J.; Laws, R.A. Successive concept mapping. *J. Coll. Sci. Teach.* 2003, 33, 12–16.
120. Pearsall, N.R.; Skipper, J.E.J.; Mintzes, J.J. Knowledge restructuring in the life sciences: A longitudinal study of conceptual change in biology. *Sci. Educ.* 1997, 81, 193–215.
121. van de Pol, J.; Volman, M.; Beishuizen, J. Scaffolding in teacher–student interaction: A decade of research. *Educ. Psychol. Rev.* 2010, 22, 271–296.
122. Chang, K.E.; Sung, Y.-T.; Chen, I.-D. The effect of concept mapping to enhance text comprehension and summarization. *J. Exp. Educ.* 2002, 71, 5–23.
123. Chang, K.E.; Sung, Y.T.; Chen, S.F. Learning through computer-based concept mapping with scaffolding aid. *J. Comput. Assist. Learn.* 2001, 17, 21–33.
124. Gurlitt, J.; Renkl, A. Prior knowledge activation: How different concept mapping tasks lead to substantial differences in cognitive processes, learning outcomes, and perceived self-efficacy. *Instr. Sci.* 2010, 38, 417–433.
125. Kapuza, A. How concept maps with and without a list of concepts differ: The case of statistics. *Educ. Sci.* 2020, 10, 91.
126. Ruiz-Primo, M.A.; Schultz, S.E.; Li, M.; Shavelson, R.J. Comparison of the reliability and validity of scores from two concept-mapping techniques. *J. Res. Sci. Teach.* 2001, 38, 260–278.
127. Ruiz-Primo, M.A.; Shavelson, R.J.; Li, M.; Schultz, S.E. On the validity of cognitive interpretations of scores from alternative concept-mapping techniques. *Educ. Assess.* 2001, 7, 99–141.
128. Ajaja, O.P. Concept mapping as a study skill. *Int. J. Educ. Sci.* 2011, 3, 49–57.

129. den Elzen-Rump, V.; Leutner, D. Naturwissenschaftliche Sachtexte verstehen—Ein computerbasiertes Trainingsprogramm für Schüler der 10. Jahrgangsstufe zum selbstregulierten Lernen mit einer Mapping-Strategie [Understanding scientific texts—A computer-based training program for 10th grade students for self-regulated learning with a mapping strategy]. In *Selbstregulation Erfolgreich Fördern [Promoting Self-Regulation Successfully]*; Landmann, M., Schmitz, B., Eds.; Pädagogische Psychologie; Kohlhammer: Stuttgart, Germany, 2007; pp. 251–268.
130. Nückles, M.; Hübner, S.; Dümer, S.; Renkl, A. Expertise reversal effects in writing-to-learn. *Instr. Sci.* 2010, *38*, 237–258.
131. Bangert-Drowns, R.L.; Kulik, J.A.; Kulik, C.-L.C. Effects of Frequent Classroom Testing. *J. Educ. Res.* 1991, *85*, 89–99.
132. Heubusch, J.D.; Lloyd, J.W. Corrective feedback in oral reading. *J. Behav. Educ.* 1998, *8*, 63–79.
133. Shute, V.J. Focus on formative feedback. *Rev. Educ. Res.* 2008, *78*, 153–189.
134. Hauser, S.; Nückles, M.; Renkl, A. Supporting concept mapping for learning from text. In *The International Conference of the Learning Sciences: Indiana University 2006; Proceedings of the ICLS 2006, Bloomington, IN, USA, 27 June–1 July 2006*; Barab, S.A., Hay, K.E., Hickey, D.T., Eds.; International Society of the Learning Sciences: Bloomington, IN, USA, 2006; Volume 1, pp. 243–249.
135. Serbessa, D.D. Tension between traditional and modern teaching-learning approaches in Ethiopian primary schools. *J. Int. Coop. Educ.* 2006, *9*, 123–140.
136. Tsai, C.-C. Nested epistemologies: Science teachers' beliefs of teaching, learning and science. *Int. J. Sci. Educ.* 2002, *24*, 771–783.
137. Reader, W.; Hammond, N. Computer-based tools to support learning from hypertext: Concept mapping tools and beyond. *Comput. Educ.* 1994, *22*, 99–106.
138. Besterfield-Sacre, M.; Gerchak, J.; Lyons, M.; Shuman, L.J.; Wolfe, H. Scoring concept maps: An integrated rubric for assessing engineering education. *J. Eng. Educ.* 2004, *19*, 105–115.
139. Kinchin, I.M.; Möllits, A.; Reiska, P. Uncovering types of knowledge in concept maps. *Educ. Sci.* 2019, *9*, 131.
140. McClure, J.R.; Sonak, B.; Suen, H.K. Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *J. Res. Sci. Teach.* 1999, *36*, 475–492.
141. Kleickmann, T.; Großschedl, J.; Harms, U.; Heinze, A.; Herzog, S.; Hohenstein, F.; Köller, O.; Kröger, J.; Lindmeier, A.; Loch, C. Professionswissen von

- Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer—
Testentwicklung im Rahmen des Projekts KiL [Professional knowledge of student
teachers of mathematical and scientific subjects—Test development within the KiL
project]. *Unterrichtswissenschaft* 2014, 42, 280–288.
142. Champagne Queloz, A.; Klymkowsky, M.W.; Stern, E.; Hafen, E.; Köhler, K. Diagnostic of students' misconceptions using the Biological Concepts Instrument (BCI): A method for conducting an educational needs assessment. *PLoS ONE* 2017, 12, e0176906.
 143. Shi, J.; Wood, W.B.; Martin, J.M.; Guild, N.A.; Vicens, Q.; Knight, J.K. A diagnostic assessment for introductory molecular and cell biology. *CBE Life Sci. Educ.* 2010, 9, 453–461.
 144. Schneider, W.; Schlagmüller, M.; Ennemoser, M. *Lesegeschwindigkeits- und Verständnistest für die Klassen 6–12 [Reading Speed and Comprehension Test for Grades 6–12]*; Hogrefe: Göttingen, Germany, 2007.
 145. Wirtz, M.A.; Caspar, F. *Beurteilerübereinstimmung und Beurteilerreliabilität [Interrater Agreement and Interrater Reliability]*; Hogrefe: Göttingen, Germany, 2002.
 146. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* 2016, 15, 155–163.
 147. Karpicke, J.D.; Blunt, J.R. Retrieval practice produces more learning than elaborative studying with concept mapping. *Science* 2011, 331, 772–775.
 148. Blunt, J.R.; Karpicke, J.D. Learning with retrieval-based concept mapping. *J. Educ. Psychol.* 2014, 106, 849–858.
 149. Townsend, C.L.; Heit, E. Judgments of learning and improvement. *Mem. Cogn.* 2011, 39, 204–216.
 150. Anderson, J.R. *Cognitive Psychology and Its Implications*, 8th ed.; Worth: New York, NY, USA, 2015.
 151. Grotzer, T.A. Expanding our vision for educational technology: Procedural, conceptual, and structural knowledge. *Educ. Technol.* 2002, 42, 52–59.
 152. Kleickmann, T. Zusammenhänge Fachspezifischer Vorstellungen von Grundschullehrkräften zum Lehren und Lernen mit Fortschritten von Schülerinnen und Schülern im Konzeptuellen Naturwissenschaftlichen Verständnis [Correlations of Subject-Specific Ideas of Primary School Teachers on Teaching and Learning with Progress Made by Pupils in Conceptual Scientific Understanding]. Ph.D. Thesis, University of Münster, Münster, Germany, 2008.

153. Watson, M.K.; Pelkey, J.; Noyes, C.R.; Rodgers, M.O. Assessing conceptual knowledge using three concept map scoring methods. *J. Eng. Educ.* 2016, *105*, 118–146.
154. Weißeno, G.; Detjen, J.; Juchler, I.; Massing, P.; Richter, D. *Konzepte der Politik: Ein Kompetenzmodell [Political Concepts: A Competence Model]*; Bundeszentrale für Politische Bildung: Bonn, Germany, 2010.
155. Großschedl, J.; Harms, U. Assessing conceptual knowledge using similarity judgments. *Stud. Educ. Eval.* 2013, *39*, 71–81.
156. Nerdel, C. *Grundlagen der Naturwissenschaftsdidaktik: Kompetenzorientiert und Aufgabenbasiert für Schule und Hochschule [Basics of Science Education: Competence-Oriented and Task-Based for Schools and Universities]*; Springer: Berlin, Germany, 2017.
157. Babb, A.P.P.; Takeuchi, M.A.; Alnosó Yáñez, G.; Francis, K.; Gereluk, D.; Friesen, S. Pioneering STEM education for pre-service teachers. *Int. J. Eng. Pedagog.* 2016, *6*, 4–11.
158. Kultusministerkonferenz. *Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie (Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 05.02.2004) [Uniform Examination Requirements in the Biology Abitur Examination (Resolution of the Conference of Ministers of Education and Cultural Affairs of 01.12.1989 as Amended on 05.02.2004)]*. Germany, 2004. Available online: https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Biologie.pdf (accessed on 10 May 2021).
159. McDonald, C.V. STEM education: A review of the contribution of the disciplines of science, technology, engineering and mathematics. *Sci. Educ. Int.* 2016, *27*, 530–569.
160. Miller, S.P.; Hudson, P.J. Using evidence-based practices to build mathematics competence related to conceptual, procedural, and declarative knowledge. *Learn. Disabil. Res. Pract.* 2007, *22*, 47–57.
161. Turns, S.R.; van Meter, P.N. Applying knowledge from educational psychology and cognitive science to a first course in thermodynamics. In Proceedings of the 2012 ASEE Annual Conference & Exposition, San Antonio, TX, USA, 10–13 June 2012.
162. U.S. Department of Education. *Transforming American Education: Learning Powered by Technology*; U.S. Department of Education: Washington, DC, USA, 2010.
163. Kinchin, I.M. Visualising knowledge structures in biology: Discipline, curriculum and student understanding. *J. Biol. Educ.* 2011, *45*, 183–189.

164. Salmon, D.; Kelly, M. *Using Concept Mapping to Foster Adaptive Expertise: Enhancing Teacher Metacognitive Learning to Improve Student Academic Performance*; Peter Lang: New York, NY, USA, 2014.
165. Kinchin, I.M. Using concept maps to reveal understanding: A two-tier analysis. *Sch. Sci. Rev.* 2000, *81*, 41–46.
166. Bitchener, J. Evidence in support of written corrective feedback. *J. Second Lang. Writ.* 2008, *17*, 102–118.
167. Lyster, R.; Saito, K.; Sato, M. Oral corrective feedback in second language classrooms. *Lang. Teach.* 2013, *46*, 1–40.
168. Deci, E.L.; Richard, M. *Intrinsic Motivation and Self-Determination in Human Behavior*; Springer US: Boston, MA, USA, 1985.
169. Ryan, R.M.; Deci, E.L. Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemp. Educ. Psychol.* 2020, *61*, 101860.
170. Karpicke, J.D. Retrieval-Based Learning. *Curr. Dir. Psychol. Sci.* 2012, *21*, 157–163.
171. Karpicke, J.D. Concept Mapping. In *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*; Frey, B.B., Ed.; Sage: Thousand Oaks, CA, USA, 2018; pp. 498–501.
172. Deutsche Gesellschaft für Psychologie [German Psychological Society]. Berufsethische Richtlinien des Berufsverbandes Deutscher Psychologinnen und Psychologen e. V. und der Deutschen Gesellschaft für Psychologie [Professional Ethical Guidelines of the Professional Association of German Psychologists e. V. and the German Psychological Society]. Available online: <https://www.bdp-verband.de/binaries/content/assets/beruf/ber-foederation-2016.pdf> (accessed on 27 May 2021).
173. World Medical Association. WMA's Declaration of Helsinki Serves as Guide to Physicians. *J. Am. Med. Assoc.* 1964, *189*, 33–34.
174. World Medical Association. Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects. *J. Am. Med. Assoc.* 2013, *310*, 2191–2194.
175. General Data Protection Regulation. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Off. J. Eur. Union* 2016, *59*, 294.

5 Studie 2: Untersuchung der Faktoren *Training* und *Setting*

In der zweiten Studie sollten zum einen die Befunde aus der ersten Studie bezüglich der Wichtigkeit des Trainings für den Lernerfolg (vgl. Kapitel 4) überprüft werden. Zum anderen sollte herausgefunden werden, ob die von Blunt und Karpicke (2014) sowie O’Day und Karpicke (2020) gewählten Bedingungen, in denen eine Kurzeinführung in das CM und eine Variation der Textverfügbarkeit implementiert wurde, Auswirkungen auf kognitive Prozesse haben und einen entscheidenden Einfluss für den Wissenserwerb darstellen. Im Rahmen des von Großschedl und Tröbst (2017) beantragten DFG-Projektes (GR 47632-1) führte der Autor der vorliegenden Dissertation die Studie durch. Er war anteilig an der Konzeption und Methodologie der Studie, überwiegend an der Datenverwaltung, wesentlich an der Datenauswertung, anteilig an den Visualisierungen und wesentlich an dem Verfassen, Einreichen und Überarbeiten des Manuskripts beteiligt (s. auch Kapitel 5, *Author Contributions*). Die Inhalte sowie Abbildungen und Tabellen dieses Kapitels (mit Ausnahme von Tabelle II) stammen aus der Originalarbeit von Becker, Welter und Großschedl (2021)⁸, welche für die vorliegende Dissertation zur Verfügung gestellt wurden, so dass in der folgenden Kurzbeschreibung auf einen fortlaufenden Quellenverweis im Interesse der Lesbarkeit verzichtet wird.

Um die in Kapitel 3.2 formulierten Hypothesen zu überprüfen, wurde eine quasi-experimentelle Studie mit vier Gruppen durchgeführt. Das Design sah den gleichen Zeitraum von sechs Wochen wie in Studie 1 mit wöchentlichen Sitzungen vor, die in drei Phasen eingeteilt wurden: eine Trainingsphase mit drei Trainingssitzungen, eine Lernphase mit einer Lernsitzung und eine Testphase mit zwei Testsitzungen (s. Tabelle II). In der Trainingsphase wurde die erste von zwei unabhängigen Variablen („Intensität des CM-Trainings“; UV_1) realisiert, indem die Proband:innen ein umfangreiches Concept Mapping-Training [T+] oder ein Kontrolltraining mit anschließender kurzer Einführung in das Concept Mapping [T–] erhielten. In der Lernphase wurde die zweite unabhängige Variable („Verfügbarkeit des Lernmaterials“; UV_2) während der Erstellung der Concept Maps variiert (Lernmaterial verfügbar/Elaboration [E] vs. Lernmaterial nicht verfügbar/Retrieval [R]). Dabei wurde auf die Trainings aus der Studie 1 zurückgegriffen: Für das Ermöglichen einer hohen Vertrautheit mit der Lernstrategie wurde das effektivere Training (Gruppe T++) aus der Studie 1 genutzt,

⁸ An dieser Stelle wird erneut darauf verwiesen, dass das Manuskript der Originalarbeit von Becker, Welter und Großschedl (2021) in Kapitel 5 einen in sich geschlossenen Abschnitt mit einer eigenen Kapitel-, Abbildungs- und Tabellenummerierung sowie eigenem Literaturverzeichnis darstellt. Verweise in der übrigen Arbeit auf einzelne Kapitel, Abbildungen und Tabellen des Manuskripts werden mit einem Verweis auf das Kapitel 5 eingeleitet, worauf die Nummer des Unterkapitels (z. B. „s. Kapitel 5, 3.2“) bzw. der Verweis auf die Abbildung oder Tabelle folgt (z. B. „s. Kapitel 5, Tab. 4“).

wohingegen die Bedingung mit geringer Vertrautheit mit Concept Mapping durch die Nutzung des Trainings für die Gruppe T- aus der Studie 1 realisiert wurde. Zur Vermeidung von Testeffekten wurde auch in der Studie 2 für die Trainingsphase das nicht-biologische Thema Intelligenz, und für die Lern- und spätere Testphase das Thema Zellbiologie gewählt.

Tabelle II. Übersicht über Phasen und Gruppen der Studie 2 (verändert nach Großschedl & Tröbst, 2017)

Gr.	Trainingsphase (Lerninhalt: Intelligenz)			Lernphase (Lerninhalt: Zellbiologie)		Testphase		
	Wo. 1	Wo. 2	Wo. 3	Woche 4		Woche 5	Woche 6	
	Training			Lernaktivität	Abhängige Variablen	Abhängige Variablen		
T+E	Vorbefragung	Concept Mapping Training (mit Scaffolding und Feedback)		Kurzeinführung, CM-Selbstwirksamkeit, CM-Treatment-Check	Concept Mapping (mit Lernmaterial)	1. während der Lernaktivität mit dem Text erstellte Concept Maps (Begriffe frei wählbar) 2. Kognitive Belastung (7 Items) 3. Metakognitive Vorhersage (1 Item)	1. Concept Maps (mit Begriffsset, bestehend aus 22 Konzepten) 2. Strukturelles Wissen (55 Items) 3. Deklaratives Wissen (30 Items)	4. Konzeptuelles Wissen (15 Items)
T+R					Concept Mapping (ohne Lernmaterial)			
T-E		Unspezifisches Kontrolltraining			Concept Mapping (mit Lernmaterial)			
T-R					Concept Mapping (ohne Lernmaterial)			

Anmerkungen. Gr. = Gruppe; CM = Concept Mapping; T+E = CM-Trainingsgruppe ($n = 23$), die dem Elaborationssetting (Lernmaterial ist während des CMs verfügbar) zugewiesen ist; T+R = CM-Trainingsgruppe ($n = 25$), die dem Retrievalsetting (Lernmaterial ist während des CMs nicht verfügbar); T-E = Kontrolltrainingsgruppe ($n = 23$), die dem Elaborationssetting (Lernmaterial ist während des CMs verfügbar) zugewiesen ist; T-R = Kontrolltrainingsgruppe ($n = 22$), die dem Retrievalsetting (Lernmaterial ist während des CMs nicht verfügbar) zugewiesen ist.

Das quasi-experimentelles Design ermöglichte es, Testbedingungen mit $n = 48$ Proband:innen in den Concept Mapping-Trainingsgruppen und $n = 45$ in den Kontrolltrainingsgruppen zufriedenstellend zu erfüllen. Somit bildete die Stichprobe aus diesen $N = 93$ Proband:innen die Datengrundlage der statistischen Analysen. Die in Studie 2 genutzten Materialien waren identisch mit denen aus der Studie 1. Ebenso glich das Vorgehen in der Trainings- und der Testphase für die beiden Trainingsgruppen dem der Gruppen T++ und T- aus der ersten Studie (s. Kapitel 4); nur in der Lernphase wurden andere Bedingungen implementiert, um die Manipulation der Textverfügbarkeit vornehmen zu können. Nichtsdestotrotz wurden die abhängigen Variablen der drei Phasen wie in der Studie 1 (s. Kapitel 4) operationalisiert. In der letzten Trainingssitzung wurde das Wissen der Proband:innen zum Thema Intelligenz mit einem Multiple-Choice Test (20 Items) erfasst, doch die nicht-signifikanten Gruppenunterschiede, $t(91) = 1,34, p = 0,18$, wurden im veröffentlichten Manuskript nicht berichtet. In Woche 4 (Lernphase) erhielten alle Proband:innen eine Einführung in das Concept Mapping und lasen anschließend den aus Studie 1 stammenden Text über den Aufbau und die Funktionen von eukaryotischen Zellen. Nach dem Lesen des Textes wurden die Proband:innen in der Lernphase zufällig in Untergruppen aufgeteilt, bevor sie mit der Konstruktion ihrer Concept Maps begannen: $n = 23$ Proband:innen der Concept Mapping-Trainingsgruppen durften mit dem Text weiterarbeiten, während sie eine Concept Map erstellten (T+E; Elaborationsgruppe); $n = 25$ Proband:innen der Concept Mapping-Trainingsgruppen konstruierten eine Concept Map ohne den Text (T+R; Retrievalgruppe); $n = 23$ Proband:innen waren in der entsprechenden Elaborations-Kontrolltrainingsgruppe (T-E) und $n = 22$ Proband:innen in der Retrieval-Kontrolltrainingsgruppe (T-R). Nach dem CM füllten die Proband:innen den Fragebogen zur kognitiven Belastung aus und machten eine metakognitive Vorhersage bezüglich ihrer erwarteten Lernleistung. Ihre Lernleistung wurde in der Testphase überprüft, die nach dem gleichen Verfahren wie in Studie 1 ablief.

**Effects of Strategy Training and Elaboration vs. Retrieval Settings on Learning of Cell
Biology Using Concept Mapping**

Becker, L. B., Welter, V. D. E., & Großschedl, J. (2021). Effects of Strategy Training and Elaboration vs. Retrieval Settings on Learning of Cell Biology Using Concept Mapping *Education Sciences*, 11, 438. doi: 10.3390/educsci11090530

Abstract

Concept mapping (CM) can be an effective strategy to facilitate understanding-based learning of complex and abstract scientific content. In our study, we tried to answer two questions that have not yet been clarified: (1) How effective is CM if the learners do not have the textual learning material available when constructing their concept maps (retrieval setting) in contrast to when the material is available (elaboration setting)? (2) Do potential effects of the two settings depend on the intensity of a previous CM strategy training? To this end, we established a quasi-experimental four-group plan and collected data of $N = 93$ undergraduate students. The participants received either a detailed CM strategy training or just a short introduction and were then asked to apply CM in a retrieval or elaboration setting. The quality of constructed concept maps and content-related learning success were assessed to determine potential group differences. Overall, our results show a significantly positive but setting-independent effect of the CM strategy training on methodical CM skills. However, the different setting conditions differed neither regarding concept map quality nor content-related learning success, implying that CM in retrieval settings could be as effective as in elaboration settings.

Keywords: concept mapping; strategy training; retrieval practice; elaboration; organization; knowledge acquisition; cell biology

1. Introduction

As a result of learners' trouble with learning in STEM subjects such as biology [1–4], researchers and educators have tried to identify new methods of instruction to afford learners the best possible learning success [5–11]. While it can certainly be helpful for learners to visualize abstract systems such as the inside of a cell with an animation or even with an augmented reality device, the question remains how learners process information, how they organize it, and how they can learn in an active and meaningful way in contrast to mere passive reception [12]. Answering these questions is crucial since the learners' challenge in biology, as in all STEM subjects, is not to merely memorize individual chunks of information but to think in extensive contexts by connecting these chunks, and thus, gain a deep understanding of their overall interrelationships [7,13–16]. This, in turn, can help them to detect interdependencies between different levels of conceptual organization, for example, regarding different biological systems. In biology, such interdependencies are also typical for the area of cell biology, which is characterized, among other things, by complex interconnected concepts ranging on different system levels [17]. However, the resulting complexity of information can pose a challenge for learners, as the cognitive processes of organization and elaboration play a crucial role for understanding such relationships [18–24].

In the context of text-based learning, cognitive processes of organization involve recognizing which statements of a section and which connections between individual chunks of information/hierarchical elements are of particular importance. Organization processes can be stimulated by appropriate cognitive strategies, including underlining key messages in a text, writing summaries, or creating illustrations [25]. By applying such strategies, learners should become aware of existing connections between individual chunks of information. Such systematic organization of knowledge can also support its reconstruction when reproduction is required (e.g., within an exam). In contrast, cognitive processes of elaboration refer to joining new information and prior knowledge elements [26,27]. Corresponding elaboration strategies therefore include verbal or pictorial enrichment, linking to examples from everyday life and personal experience, or building analogies. Using such strategies can stimulate elaboration processes, which are essential for meaningful and understanding-based learning [28,29] and facilitate connected thinking in the context of science education [30].

Complementing the research on learning efficacy of cognitive strategies such as organization and elaboration, increasing attention has also been paid to the mechanism of retrieval of information from memory and specifically to a possible retroactive effect of this retrieval on memory organization and the learning performance [31–35]. Retrieval processes are assumed to occur when learners retrieve information from memory without simultaneous availability of the learning material, resulting in a setting of retrieval practice [36–39]. In this regard, studies could show that such retrieval practice not only provides learners with an opportunity to monitor their learning progress but improves learning itself, as well [39–42]. Compared to retrieval, the operationalization of elaboration processes is usually characterized by a simultaneous availability of learning material, which the learners actively deal with in order to integrate new information into their prior knowledge. Following this strain of reasoning, such a setting is called elaborative studying setting.

Consequently, a plausible approach to promote learning success could lie in using an organization or elaboration strategy within a retrieval setting. However, according to O'Day and Karpicke [43], studies combining retrieval practice and elaborative learning strategies are “woefully sparse” (p. 2).

1.1. Concept Mapping

One learning strategy that has been used for both organization and elaboration but has rarely been used specifically in a retrieval setting is concept mapping (CM). Previous findings indicate that CM facilitates learning of abstract topics and complex relationships in STEM subjects, such as biology [44–48], physics [49,50], and chemistry [51,52]. Products of applying CM, the concept maps, are network diagrams, representing the types of relationships between meaningful terms or concepts. In practical terms, such a network is created by connecting the concepts (nodes) with labelled arrows: while the label indicates the semantic relation between two concepts, the arrow direction indicates the reading direction. Two concepts connected by a labelled arrow constitute a so-called proposition, representing the smallest meaningful unit of a concept map. Areas of concept maps that are not in close proximity can be associated with each other via building cross-connections to point out an existing relationship [53–58].

The aforementioned features suggest concept maps to be an isomorphic analogy to the assumed structure of memory, which is based on Quillian’s semantic network model [59,60]. According to this model, knowledge is represented in more or less hierarchically structured semantic networks of memory [61,62]. Relationships connecting individual concepts that form the structure of these memory networks are semantic, logical, and grammatical in nature. As concept maps simplify the cognitive processing of semantic relations by explicating the logical elements between one concept and another, CM is particularly suitable to facilitate an understanding of complex issues. The specification of complex relationships during CM requires a systematic and analytical approach, leading to structured representations of contexts in resulting concept maps [58,63–65]. In addition, these processes during CM can promote organization and elaboration processes [66] and encourage learners to analyze the learning material on a deeper level, resulting in increasing learning success [44,53,67–73].

Furthermore, CM can also serve as a metacognitive tool, as noticing one’s own understanding and misunderstanding is particularly important for developing metacognitive skills, which has been named one of the main arguments in favor of teaching students CM [74–78]. By drawing the mapper’s attention to difficulties regarding the plausible integration of concepts or the specification of relationships, concept maps can help learners to recognize knowledge gaps or flawed logic, so they are more likely to become able to react to them, for example, by restudying the material.

Finally, CM can be used to assess learning success, e.g., by analyzing the number of propositions specified by learners. From a pedagogical point of view, concept maps can serve as such a diagnostic tool because they represent the learner’s individual understanding of a

respective domain [79,80]. Therefore, a concept map draws the teacher's attention to what the learner (already) knows, but also to what the learner did not understand or what may have been misunderstood. In this regard, research has shown that concept maps are useful in assessing how learners relate, organize, and structure concepts [81]. Therefore, it is not surprising that CM has been used as a diagnostic tool in many studies regarding meaningful learning [82–86].

1.2. Previous Research on the Effectiveness of Concept Mapping in Retrieval Settings

Since CM stimulates the aforementioned processes of organization and elaboration in the course of working on learning material [87], the experimental design of constructing concept maps with simultaneous availability of a learning text (elaboration setting) seems the obvious configuration to enable the learners to understand the text using CM before their learning success is measured [38,71,88].

An alternative design was used by Blunt and Karpicke [37] to explore the effectiveness of CM, a typical elaborative learning strategy, in a retrieval practice setting. In two experiments, they were able to show that CM can also be used effectively within retrieval practice formats and that the two established learning formats, CM vs. note-taking, each resulted in better learning outcome in terms of factual and inferential knowledge when used in a retrieval practice setting instead of a classical elaboration setting in which the learning material is available. Blunt and Karpicke primarily attribute this learning strategy-independent finding to generally better skills of the retrieval practice groups' participants regarding the use of retrieval cues for knowledge reconstruction [37].

For further exploration of the learning effectiveness of CM as an elaboration-supportive learning strategy in a retrieval setting, O'Day and Karpicke [43] conducted another two experiments in which students had to read short texts and practiced retrieving the information by free recall, CM, or both. Their results indicate a superiority of the free recall group over the CM group and the CM-and-recall group regarding retention of information.

However, the aforementioned results of Blunt and Karpicke [37] as well as O'Day and Karpicke [43] should be viewed with caution since they could have been biased by a lack of practice and familiarity with CM on the part of the learners, causing additional cognitive load. These studies' participants who were asked to apply CM in learning only received a short introduction about the strategy's basics, which has also been criticized by researchers such as Mintzes et al. [89]. In addition, the brevity of the learning text of approximately 250 words used in the study of O'Day and Karpicke [43] may have represented an unrealistically low burden for students and could therefore have undermined their motivation to elaborate on their prior knowledge.

Since CM is also considered an effective metacognitive tool (see Section 1.1), Karpicke and Blunt [38], Blunt and Karpicke [37], and O'Day and Karpicke [43] additionally examined the aspect of judgement of learning (JOL) by asking their participants to judge how much (0 to 100%) of the learning content they will be able to remember one week later. The results of Karpicke and Blunt [38] and Blunt and Karpicke [37] show that even though retrieval practice groups performed better in follow-up tests, students in the elaborative studying setting rendered higher JOLs beforehand. Comparable results were also found in other studies focusing on effects of the availability of learning material [90,91]. The incorrect prognoses regarding one's own learning success are often explained by the empirical finding that JOLs apparently show a positive linear dependency on the perceived ease of information processing. When the material is available to learners, information processing usually seems smooth and easy, resulting in high JOLs and somewhat unrealistic self-efficacy beliefs compared to actual test performance levels. In contrast, it is possible that active retrieval changes the basis of judgement: instead of judging how easily information can be read and/or processed, judgments could more likely be based upon the ease of memory recall of this information during retrieval practice [92–94].

1.3. Previous Research on the Effectiveness of Different Concept Mapping Training Approaches

Since CM requires abilities such as identifying important terms or concepts (e.g., in a learning text), determining hierarchical relationships among them, and specifying meaningful propositions, it seems obvious that learners need additional training and support to use the strategy successfully. Accordingly, some studies involved several weeks of training with repeated intervention and feedback measures [44,45,78,83,89,95,96]. In contrast, other researchers such as Jonassen et al. [97] state that CM is comparably easy to learn, so they regard short introductions to be sufficient when investigating the strategy [37,38,43]. Even if such short introductions have been implemented successfully in some cases [98], it is not clear to what extent such effects depend on specific characteristics of the participants (e.g., cognitive abilities, previous experience with CM) and/or methodical variations between the studies (e.g., level of requirement, intensity of instruction). Accordingly, empirical recommendations on parameters such as the content and duration of a training course/instruction vary [84,86,99,100].

In a previous study, we already have been able to show that the overall learning success can be improved by an extensive CM training and that stable CM strategy skills, in particular, can be promoted by additional integration of scaffolding and feedback elements in the course of such a training. However, we also noted that our small sample ($N = 73$) was not sufficient in terms of external validity and further studies were needed to replicate these findings [44].

Therefore, in the present study, we wanted to take the opportunity to validate our previous results.

1.4. Research Questions and Hypotheses

Considering the aforementioned different training and instruction approaches as well as the different effects of CM setting variations (elaboration vs. retrieval) on learning outcome, we designed a study to provide more clarity in this regard. Specifically, we try to answer the following questions, using a four-group plan: (1) How effective is CM if the learners do not have the textual learning material available while constructing their concept maps (retrieval setting) in contrast to the classical elaboration setting, where the learning material is available? (2) Do potential effects of the two settings depend on the intensity of a previous method training in the utilization of CM? To answer these two questions, we have translated them into three corresponding hypotheses:

- (1) Based on the findings of Blunt and Karpicke [37], we expect that the availability of the learning material is crucial for cognitive processes during CM. In absence of a learning text, a CM construction task regarding the learning content should induce memory-related recall processes which bind cognitive resources usually required for further elaboration. Vice versa, elaboration processes should increase if the learning material is available to participants during CM, which should also be reflected in a more elaborative character of the propositions participants specified in their concept maps. Since a higher level of elaboration is additionally associated with better learning outcomes in general, we expect these participants to achieve higher scores in subsequent knowledge tests regarding the covered learning content.
- (2) Furthermore, we expect that extensive training in CM is helpful for its successful use as a learning tool [44,45,95] by inducing a certain familiarity, especially with methodical aspects. This effect should become particularly evident in comparison with another condition in which the participants only receive a short introduction to the CM strategy. Accordingly, we expect participants who took part in an extensive CM training to show higher CM-related self-efficacy expectations, to achieve better results in analyzing and editing given concept maps, as well as to make fewer methodical mistakes when creating own concept maps.
- (3) Finally, we expect that the combination of both factors, an extensive CM training and the availability of the learning material during CM, will improve CM- and knowledge-related learning outcomes the most, which should also be reflected in a related metacognitive assessment of the participants regarding their own learning success [36].

2. Materials and Methods

Our quasi-experimental study was conducted as a four-group plan. On the one hand, the variable “CM training intensity” (extensive CM training [T+] vs. control training followed by a short introduction to CM [T-]) was varied; on the other hand, the variable “availability of learning material” during CM (learning material not available/retrieval [R] vs. learning material available/elaboration [E]) was varied. The implementation of this design took place over a six-week period of weekly sessions, which comprised three consecutive phases:

- (1) A training phase (weeks 1–3), in which students received an extensive CM training (T+) or a control training (T-);
- (2) A learning phase (week 4), in which students created a concept map on the topic cell biology in presence (E) or absence (R) of the respective learning material; and
- (3) A test phase (weeks 5–6), in which the students’ learning success was measured by implementation of various knowledge tests (see Section 2.4).

In order to avoid test effects of the training phase on the later learning and test phases, we have chosen different but comparably abstract learning topics for these parts: intelligence for the training phase and cell biology for the learning and test phases. This approach has proven to be useful in our previous study [44].

The learning texts that the participants worked on during the study units of the training and learning phases were designed by us prior to our previous study [44] and assessed as appropriate to the university level by 13 experts in the respective fields of biology education ($n = 9$) and psychology ($n = 4$). The text *Theories and Models of Intelligence* comprised 3197 words (8 pages) and was divided into three equal units to cover the three training sessions, whereas the text *The Structure and Function of Eukaryotic Cells* comprised 2010 words (7 pages) and was used entirely in the learning phase. Throughout the study, the same standard of behavior and situation was strictly implemented for all four groups in every study-related interaction with the participants in order to minimize investigator and context effects.

2.1. Sample

A total of $N = 93$ undergraduate university students of different fields of study (54% enrolled in a natural science study program) participated in our study. On average, our participants were 21.9 years old, and 84% of them were female.

Participation was based on self-selection in the first instance, as our study was carried out as part of an elective curricular course on learning strategies. Accordingly, the participants

had to spend their free time, so we offered each weekly session at a total of five different time slots (different days of the week and different times of the day). After all participants had individually decided in advance on the most suitable time slot for them, they were permanently assigned to it for the entire duration of the study. Afterwards, we randomly assigned each of the five time slot groups to one of the two training conditions, resulting in a total of $n = 48$ participants in the T+ and $n = 45$ participants in the T- groups. These participants were again randomly assigned to one of the two setting conditions (learning material not available/retrieval [R] vs. learning material available/elaboration [E]), so that the four final quasi-experimental groups were created (see Figure 1). The minimally unequal distribution of the participants in the respective groups is due to a dropout of a total of five subjects during the study period.

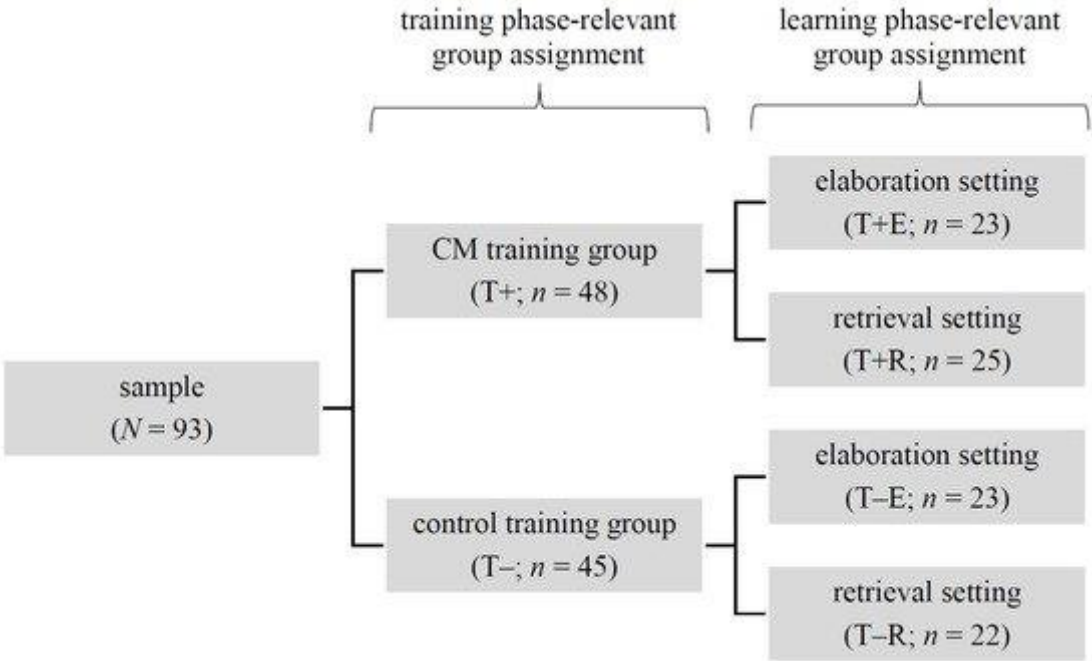


Figure 1. Composition of study groups.

2.2. Procedure

During the three training sessions (weeks 1 to 3), all participants worked on a learning text on the topic of intelligence, but the groups differed in terms of the strategy they should use in learning: whereas the T+ groups used CM, the T- groups used non-CM strategies. Accordingly, the participants in both groups received different instructions, which were partly based on that of our previous study, and thus, already proven to be useful [44]:

- (1) Participants of the T+ groups received an extensive CM training including supportive and feedback elements. Every weekly session started with a theoretical lecture on the CM strategy. In addition to the lectures, the participants received CM-specific scaffolding (see Table 1) and metacognitive prompts regarding the individual study unit following the lecture (e.g., “Did I label all arrows clearly, concisely, and correctly?”; adapted from Großschedl and Harms [101,102]). However, the number of prompts was reduced over the course of the training phase (fading) [73] to prevent unnecessary distraction. In order to give the participants the opportunity to check the correctness of the concept maps they had constructed, they received a scripted overview of the most common CM errors. Feedback by the instructor was provided continuously during the study unit: individual verbal feedback was given on request during the participants’ construction of own concept maps, and written feedback on these constructed concept maps was given after the study unit using a knowledge of correct results (KCR) approach. Here, feedback was limited to marking CM errors and pointing out possible resulting misconceptions. In addition, an expert map was discussed at the end of each training session, so participants had the opportunity to compare it to their own and ask questions.
- (2) Participants of the T- groups did not receive any CM training but rather a control training including popular non-CM learning strategies [103,104] (see Table 1). However, the training sessions’ procedure followed a similar pattern to that of the T+ groups: the instructor started with an advance organizer and introduced the learning strategy to be used that day, including metacognitive prompts. Afterwards, participants used the respective strategy to work on the learning material, whose content was identical to that of the T+ groups and related to the topic of intelligence. At the end of the training session, an expert solution was discussed, too, allowing for comparison of own learning results and asking questions.

Table 1. Group activities during the training phase.

Training Content	Week 1	Week 2	Week 3
Scaffolding in T+ group	Skeleton map	Given set of concepts	Free practice
Learning strategy in T- group	Small group discussion	Writing a summary	Carousel workshops

T+ = CM training group ($n = 48$); T- = control training group ($n = 45$).

At the beginning of the learning phase in week 4, all participants of both training conditions received a short introduction to the CM strategy as the T- groups had not yet received any CM training. For the participants of the T+ groups, this was certainly a repetition, but it seemed important to us to refresh their knowledge at the beginning of the learning phase in order to counteract a possible diminishing effect of the training. After this short introduction, all participants filled out a questionnaire on their CM-related self-efficacy expectation. We expected the T+ groups to rate this higher than the T- groups. In order to be able to check objectively whether their assessment was accurate, they received a pre-built but error-including concept map afterwards, in which as many errors as possible should be detected and corrected within a four-minute time limit. In the following study unit, all participants read a text on the structure and functions of eukaryotic cells. Afterwards, those participants of the T+ and T- groups who were assigned to the elaboration setting were allowed to keep the text while constructing a concept map on the topic of cell biology (groups T+E and T-E), whereas those who were assigned to the retrieval setting had to construct the map without the text available (groups T+R and T-R; see Figure 1). After completing CM, all participants gave a judgment of learning on how much about the topic of cell biology they would remember one week later (see Section 2.4.2).

In the test phase (weeks 5 and 6), on the one hand, the participants' CM skills were assessed again in the context of a further construction of a concept map on the topic of cell biology, and on the other hand, we evaluated their overall learning success by application of declarative, structural, and conceptual knowledge tests (see Section 2.4.3).

2.3. Concept Map Scoring

Taking into account the expectation that participants who had the learning material available during CM (elaboration setting) would potentially specify more propositions than those who worked memory-based (retrieval setting), we used two different approaches to evaluate the quality of the concept maps constructed: an absolute (aQCM) and a balanced quality of concept map (bQCM) index. This approach has proven to be useful in our previous study [44].

- (1) In order to determine the aQCM index, we followed the approach of McClure et al. [105], assigning each specified concept map proposition a value of 0 to 3 points: 0 points if there actually did not exist a relation between the concepts; 1 point if there was a relation between the concepts, but the arrow label was meaningless; 2 points if the arrow label was meaningful, but the arrow pointed in the wrong direction; 3 points if the whole proposition was correct and meaningful.
- (2) In contrast, the bQCM index considers the statistically higher probability of making mistakes when specifying more propositions by putting the aQCM index into relation to the number of the participants' overall propositions specified.

Additionally, we analyzed the type of the participants' CM mistakes by classifying them as methodical (e.g., missing arrowhead) or content-related (e.g., animals—have→cell walls). On the one hand, this categorization allows for quantifying individual methodical CM skills by focusing on the concept map error ratio (number of methodical mistakes divided by the number of overall propositions specified). On the other hand, it can be used to assess individual understanding of the learning content by focusing on the content-related error ratio (number of content-related mistakes divided by the number of overall propositions specified).

Finally, we analyzed the type of the participants' CM propositions by classifying them as recall- (R), organization- (O), or elaboration-suggesting (E). If both the relation and the associated concepts were covered completely by the learning material, this suggests an R-proposition; relations that were not explicitly named in the learning material but were constructed between concepts covered by it represent O-propositions; relations between two concepts, which included at least one that was not mentioned in the learning material, represent E-propositions, since prior knowledge needed to be integrated into the concept map (see Section 3.5).

2.4. Further Measures and Operationalizations

In the following, we describe the operationalization of the other dependent variables and some variables we controlled for baseline differences in chronological order of the consecutive study phases. Figure 2 provides an overview of all assessed variables and their relations to the study's hypotheses.

<p style="text-align: center;">Hypothesis 1</p> <p style="text-align: center;">If the learning material is available during CM, elaboration processes should increase.</p>	<p style="text-align: center;">Hypothesis 2</p> <p style="text-align: center;">Extensive training in CM is helpful for its successful use as a learning tool by inducing a certain familiarity especially with methodical aspects.</p>	<p style="text-align: center;">Hypothesis 3</p> <p style="text-align: center;">A combination of both factors, an extensive CM training and the availability of the learning material during CM, will improve overall learning success most.</p>
<ul style="list-style-type: none"> • Types of propositions (recall-, organization-, or elaboration-suggesting) • Accuracy of propositions (aQCM and bQCM indices) • Number of content-related mistakes when creating own concept maps (content-related error ratio) • Knowledge: Declarative (reproduction of learning), structural (SJT), and conceptual (transfer of learning) 	<ul style="list-style-type: none"> • CM-related self-efficacy expectation • Number of detected errors and accuracy of error correction when working on a pre-built but error-including concept map (error task) • Number of methodical mistakes when creating own concept maps (concept map error ratio) 	<ul style="list-style-type: none"> • Types of propositions (recall-, organization-, or elaboration-suggesting) • Accuracy of propositions (aQCM and bQCM indices) • Knowledge: Declarative (reproduction of learning), structural (SJT), and conceptual (transfer of learning) • Judgment of learning (JOL)

Figure 2. Overview of assessed variables (bullet points) and their relations to the study's hypotheses.

2.4.1. Measures of the Training Phase (Week 1)

At the beginning of the training phase, we first assessed socio-demographic data and other relevant variables we controlled for baseline differences: (1) familiarity with CM, (2) prior knowledge of biology, and (3) reading skills (see Section 3.1).

Familiarity with CM was assessed via a questionnaire adapted from McClure et al. [105]. Using seven items, the participants rated their previous experience with CM on a five-point Likert scale (1 = *never/very rarely* to 5 = *very often/always*). The internal consistency of this scale was $\alpha = 0.88$.

The participants' prior knowledge of biology was assessed in two different ways in order to obtain both a valid indicator for general knowledge of biology and a specific one regarding the topic of cell biology. For this purpose, we asked about the extent of their biology education during the past two years of schooling (*no biology* vs. *basic biology* vs. *advanced biology*) and used an 18-item knowledge test on the topic of cell biology. This test comprised (partly adapted) established items used in previous studies [24,44,106–108], and consisted of single-choice and grouping tasks. For the present sample, this test showed an internal consistency of $\alpha = 0.77$ (after three items were excluded due to a lack of item-total correlation).

Since the participants had to work on the (demanding) learning texts in a given time frame during the learning units, it seemed important to us to ensure that there were no systematic

variations regarding reading speed and comprehension. Therefore, we assessed both variables using a validated instrument (Lesegeschwindigkeits- und Verständnistest [LGVT] 6–12) [109]. The task was to fulfill blank spaces by marking in each case the one meaningful out of three answer options while reading the text as fast as possible. The LGVT's internal consistency was $\alpha = 0.51$ for the present sample, indicating a lack of homogeneity of the construct [110]. Consequently, we needed to exclude the scale from further analyses.

2.4.2. Measures of the Learning Phase (Week 4)

After all participants had received the short introduction to the CM strategy at the beginning of the learning phase, they filled out a questionnaire on their CM-related self-efficacy expectation (see Section 3.2), as we expected differences between the T+ and T- groups. The scale ($\alpha = 0.83$) consisted of six items (e.g., “I feel competent in choosing the important concepts for my concept map”, or “I could explain concept mapping only with words to a friend so that he could construct one on his own”), each rated on a five-point Likert scale (1 = *I do not agree at all* to 5 = *I strongly agree*). Afterwards, the participants received a pre-built but error-including concept map, in which as many errors as possible should be detected and corrected within a four-minute time limit (see Section 3.2). The pre-built concept map consisted of 18 propositions and a total of 10 common errors (e.g., missing arrowhead). Two measures were considered to obtain an objective measure of CM competence (besides the CM-related self-efficacy ratings): the number of errors detected and the accuracy of error correction.

The concept maps constructed by the participants themselves on the topic of cell biology were evaluated following the procedure described in Section 2.3 (see also Section 3.5). Interrater reliability of bQCM indexing was proven by randomly selecting one-third of the participants' concept maps for an independent re-scoring by a second rater [111]. The analysis yielded excellent interrater reliability on average ($ICC_{\text{unjust}} = 0.98$, $CI_{95\%} [0.93, 0.99]$), indicating that participants' concept maps can be clearly judged using the coding scheme provided [112].

In addition to the CM-related measures, we assessed the judgment of learning (JOL) by asking the participants to give a metacognitive prediction regarding their learning outcome one week later (see Section 1.2 and Section 3.3). As many researchers suggest that CM represents a useful metacognitive learning tool [66,113], this kind of JOL prediction has been used in several previous studies [37,38,43]. Specifically, we asked our participants to rate on a scale from 0 to 100% how much information of the learning text on cell biology they will probably remember one week later. Later on, these JOLs were compared to the corresponding objective performance measures resulting from knowledge tests of the test phase (see Section 2.4.3).

2.4.3. Measures of the Test Phase (Weeks 5 and 6)

In order to evaluate the participants' overall learning outcome, several measures were used during the two test sessions, including re-assessment of CM competence (see Section 3.5) as well as testing of declarative, structural, and conceptual knowledge (see Section 3.4). The three latter types of knowledge differ from one another: (1) declarative knowledge refers to coexisting but separate facts, (2) structural knowledge also includes relationships between such individual chunks of information, and (3) conceptual knowledge is characterized by a highly decontextualized organization of features and principles within the facts and their interrelations. Accordingly, learners with a high level of conceptual knowledge can flexibly use extensive abstract knowledge in different specific contexts [114,115].

The test phase started with the participants being given the task of constructing a concept map within 60 min by using a given set of 22 concepts taken from the text of the learning phase on the topic of cell biology. Concept map evaluation followed aQCM and bQCM indexing, as described in Section 3.3. Using the same procedure as for the corresponding measure of the learning phase, the interrater reliability was determined for one-third of the participants' concept maps; this analysis indicated it as excellent on average ($ICC_{\text{unjust}} = 0.99$, $CI_{95\%} [0.90, 1.00]$) [112].

After solving the CM task, participants' cell biology-related structural knowledge was assessed in a 30-minute time frame using a 55-item Similarity Judgments Test (SJT; adapted from Gröbelschedl and Harms [116]). This SJT consisted of pairs of overall 11 cell biological concepts (e.g., "cell membrane"–"ribosomes"; "rough endoplasmic reticulum"–"proteins"), and the participants' task was to judge the semantic proximity of these pairs on a nine-point Likert scale (1 = *minimally related* to 9 = *strongly related*). In order to control for possible sequence effects, the 55 pairs were presented in a balanced manner using two test versions A and B. The participants' individual responses for each item were compared to an average rating of $n = 7$ experts. Regarding these experts' rating, the intraclass correlation coefficient indicated excellent agreement on average ($ICC_{\text{unjust}} = 0.95$, $CI_{95\%} [0.93, 0.97]$) [112], so content validity of the procedure can be assumed. The correlation of the participants' and experts' ratings was used as indicator of structural knowledge.

The first test session (week 5) ended after a test on declarative knowledge of cell biology (60 min). This multiple-choice test comprised a total of 30 items (e.g., "Which of the following statements about the cytoplasm are correct?") and showed an internal consistency of $\alpha = 0.85$. During the last day of our study (week 6), the participants' cell biology-related conceptual knowledge was assessed. For this purpose, we have created a 15-item open answer format test

($\alpha = 0.86$) including an associated coding manual. The participants' answers were coded as either incorrect (0 points), partially correct (1 point) or completely correct (2 points). Face validity of both the test and the coding manual was checked in advance by six independent experts in the fields of biology education ($n = 3$) and test construction ($n = 3$) and found to be adequate.

3. Results

To determine the statistical correlations and differences of interest, we used parametric (Pearson correlation, *t*-test, ANOVA, and MANOVA) and non-parametric statistical analyses (Chi², Mann–Whitney, and Kruskal–Wallis test). The specific decision for a parametric or non-parametric method was based on distribution parameters of the analyzed variables. Accordingly, relevant empirical distributions (skewness, kurtosis, variance) were always considered within preliminary analyses in order to be able to decide whether the assumption of parametric statistical analyses were met. If there were significant deviations in one or more of these parameters, we applied nonparametric analyses accordingly.

In order to ensure the best possible clarity, the following results section is structured by content-related aspects, i.e., the analyses referring to related measures (e.g., all knowledge tests) are reported together.

3.1. Baseline Difference Testing

Since our sample was made up of self-selected participants and the number of subjects per test condition was comparatively small, we first checked for potential baseline differences between the groups on relevant variables using the data assessed at the beginning of the training phase.

Regarding the categorical variables educational level in biology and university study program, Chi² tests indicates an equal distribution across the four groups (see Table 2). Regarding the metric variables age, GPA, prior knowledge in cell biology, and familiarity with CM, the results of a univariate ANOVA do not indicate any baseline differences (see Table 3).

Table 2. Chi² test results regarding possible baseline differences between the four (quasi-) experimental groups.

Variable	Category	Observed <i>n</i> in groups				Chi ² Test
		T+E	T+R	T-E	T-R	
Educational level in biology	essential	2	7	6	8	$\chi^2(6) = 7.83, p = 0.25$
	basic	17	13	9	11	
	advanced	4	5	8	4	
University study program	B. A.	12	14	8	9	$\chi^2(3) = 2.75, p = 0.43$
	B. Sc.	11	11	15	13	

T+E = CM training group (*n* = 23) assigned to elaboration setting (learning material available during CM); T+R = CM training group (*n* = 25) assigned to retrieval setting (learning material not available during CM); T-E = control training group (*n* = 23) assigned to elaboration setting (learning material available during CM); T-R = control training group (*n* = 22) assigned to retrieval setting (learning material not available during CM); an essential educational level in biology means no biology education during the past two years of schooling.

Table 3. ANOVA results regarding possible baseline differences between the four (quasi-) experimental groups.

Variable	Group								ANOVA
	T+E		T+R		T-E		T-R		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Age	22.96	7.54	21.28	4.04	22.09	5.31	21.27	4.44	$F(3, 89) = 0.50, p = 0.69$
GPA	1.95	0.67	2.06	0.74	2.06	0.67	1.80	0.53	$F(3, 89) = 0.83, p = 0.48$
Prior knowledge of cell biology	8.52	3.13	7.28	3.61	7.00	4.10	7.04	3.43	$F(3, 89) = 0.92, p = 0.44$
Familiarity with CM	1.60	0.64	1.81	0.83	1.94	0.74	2.22	0.89	$F(3, 89) = 2.49, p = 0.07$
Reading speed	analyses not carried out due to the scale's lack of reliability ($\alpha = .51$; see								
Reading comprehension	Section 2.4.1)								

T+E = CM training group (*n* = 23) assigned to elaboration setting (learning material available during CM); T+R = CM training group (*n* = 25) assigned to retrieval setting (learning material not available during CM); T-E = control training group (*n* = 23) assigned to elaboration setting (learning material available during CM); T-R = control training group (*n* = 22) assigned to retrieval setting (learning

Table 3. (*Fortsetzung*)

material not available during CM); *M* = mean value; *SD* = standard deviation; a lower GPA (= final school exam grade) indicates better performance; the maximum score for prior knowledge of cell biology was 15.

Consequently, no covariates were included additionally in subsequent statistical analyses as our results do not show any significant group differences regarding potentially confounding variables.

3.2. Concept Mapping-Related Self-Efficacy and Error Detection Task (Measures Prior to Learning Phase, Week 4)

After all participants of both training conditions received a short introduction to the CM strategy at the beginning of the learning phase in week 4, CM-related self-efficacy expectations and objective measures of CM competence by working on a pre-built but error-including concept map were assessed.

Although we had expected that the T+ groups would show higher CM-related self-efficacy expectations after the training phase than the T- groups, the corresponding Mann–Whitney test surprisingly showed no statistically significant group differences (see Table 4).

Table 4. Mann–Whitney test results regarding CM-related self-efficacy and the error map task.

Variable	Group				Mann–Whitney Test
	T+		T–		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
CM self-efficacy	5.21	1.05	5.17	1.14	$U = 1071.0, p = 0.94$
Error detection	7.52	2.09	7.27	1.63	$U = 1219.5, p = 0.28$
Proper error correction	7.48	2.12	7.27	1.63	$U = 1209.0, p = 0.31$
Improper error correction	0.23	0.59	0.44	0.66	$U = 878.0, p < 0.05$

T+ = CM training group ($n = 48$); T– = control training group ($n = 45$); *M* = mean value; *SD* = standard deviation; the maximum score for error detection and proper error correction was 10.

Furthermore, Mann–Whitney test could not reveal any statistically significant differences between the T+ and T– groups regarding the number of errors detected when working on a pre-built but error-including concept map. However, a comparison of the accuracy

of error correction indicates a significant group difference in favor of the T+ groups, specifically with regard to the mean number of improperly corrected errors $U(n_1 = 48, n_2 = 45) = 878.0, p < 0.05, d_{\text{Cohen}} = 0.33$ (see Table 4).

3.3. Metacognitive Prediction/Judgment of Learning (JOL)

An ANOVA carried out to determine differences regarding the JOL showed no significant differences between the four groups, $F(3, 89) = 0.79, p = 0.50$. This means, on average, all groups were equally confident about their later learning success, even though the groups nominally differed slightly in their ratings: group T–E predicted the most success ($M = 57.39, SD = 20.72$), followed by group T–R ($M = 55.91, SD = 24.23$), group T+E ($M = 53.91, SD = 22.10$), and group T+R ($M = 47.6, SD = 27.58$). This result indicates that neither an extensive CM training nor the presence or absence of the learning material during CM influence the JOLs significantly.

In addition, correlational analyses across all groups indicate statistically significant associations between these JOLs and corresponding objective performance measures resulted from the test phase’s knowledge tests (see Table 5). However, the results show that the students’ reference point in giving their metacognitive prediction obviously related more to declarative ($r = 0.62, p < 0.001$) and structural ($r = 0.52, p < 0.001$), but less to conceptual knowledge ($r = 0.15, p = 0.16$). In this context, it is noticeable that conceptual knowledge did not correlate with neither the metacognitive prediction nor with declarative or structural knowledge and thus actually seems to represent an independent domain of knowledge.

Table 5. Correlations between JOL and objective performance measures across all groups ($N = 93$).

Variable	1	2	3	4
1 Metacognitive prediction	—			
2 Declarative knowledge	0.62 ***	—		
3 Structural knowledge	0.52 ***	0.65 ***	—	
4 Conceptual knowledge	0.15	0.09	0.14	—

Declarative knowledge refers to coexisting but separate facts; structural knowledge also includes relationships between these facts; conceptual knowledge is characterized by a highly decontextualized organization of features and principles within the facts and their interrelations; *** = $p < 0.001$.

3.4. Declarative, Structural, and Conceptual Knowledge

In order to evaluate the students' overall learning outcome, we assessed three different types of knowledge beside CM quality in the test phase (weeks 5 and 6). To determine potential group differences between all four groups, we carried out two analyses: (1) MANOVA, taking into account declarative and structural knowledge as one latent dependent variable (linear combination), as the correlation analysis had already shown that these two types of knowledge were highly correlated in our sample, and (2) a separate ANOVA for the variable of conceptual knowledge, as it has been proven to be independent from the other types of knowledge (see Section 3.3). However, these analyses could neither reveal any significant group differences in terms of declarative and structural knowledge nor in terms of conceptual knowledge (see Table 6). On the other hand, these results are consistent with our findings above that no significant group differences could be shown regarding the metacognitive prediction of learning success in terms of knowledge acquired, even though two of the three types of knowledge were partially highly correlated across all groups.

Table 6. (M)ANOVA results regarding different types of knowledge acquired.

Variable	Group								(M)ANOVA
	T+E		T+R		T-E		T-R		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Declarative Knowledge	17.48	6.25	16.96	6.56	19.52	5.55	17.32	5.35	$F(6, 178) = 0.78, p = 0.58$
Structural Knowledge	0.40	0.18	0.41	0.19	0.41	0.20	0.39	0.21	
Conceptual Knowledge	14.61	7.26	14.20	5.63	15.43	6.36	13.27	6.34	$F(3, 89) = 0.44, p = 0.72$

T+E = CM training group ($n = 23$) assigned to elaboration setting (learning material available during CM); T+R = CM training group ($n = 25$) assigned to retrieval setting (learning material not available during CM); T-E = control training group ($n = 23$) assigned to elaboration setting (learning material available during CM); T-R = control training group ($n = 22$) assigned to retrieval setting (learning material not available during CM); *M* = mean value; *SD* = standard deviation; the maximum scores for declarative knowledge and conceptual knowledge were 30; the maximum score for structural knowledge was 1.0.

In addition to these analyses taking into account all four groups, we performed specific comparisons between the two settings elaboration (E) and retrieval (R), neglecting the affiliation to the T+ and T- groups, as we expected that a higher level of elaboration is associated with better learning outcomes in general, including better knowledge. However, these comparisons also did not show any significant differences in terms of significant advantages of the elaboration groups regarding the three domains of knowledge (see Table 7).

Table 7. *t*-test results regarding different types of knowledge acquired.

Variable	Group				<i>t</i> -Test
	E		R		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Declarative Knowledge	18.50	5.94	17.13	5.96	$t(91) = 1.11, p = 0.27$
Structural Knowledge	0.40	0.19	0.40	0.20	$t(91) = 0.06, p = 0.95$
Conceptual Knowledge	15.02	6.76	13.77	5.93	$t(91) = 0.95, p = 0.34$

E = group ($n = 46$) assigned to elaboration setting (learning material available during CM); R = group ($n = 47$) assigned to retrieval setting (learning material not available during CM); *M* = mean value; *SD* = standard deviation; the maximum scores for declarative knowledge and conceptual knowledge were 30; the maximum score for structural knowledge was 1.0.

3.5. Concept Map Quality

In order to be able to evaluate the quality of the concept maps created by the participants both in the learning (week 4) and in the test phase (weeks 5 and 6), we used several indicators: (1) an absolute (aQCM) and a balanced quality of concept map (bQCM) index, (2) an analysis of the type of the participants' CM mistakes, and (3) an analysis of the type of the participants' CM propositions (see Section 2.3).

3.5.1. aQCM and bQCM Indices

Regarding the aQCM and bQCM indices, the four groups only differed in the learning phase (week 4), $F_{aQCM}(3, 88) = 4.23, p < 0.01, d_{Cohen} = 0.76$ and $\chi^2_{bQCM}(3) = 8.66, p < 0.05, d_{Cohen} = 0.52$ (see Table 8). An additional specific consideration of the learning phase's settings shows that the elaboration groups achieved significantly higher scores on average than the retrieval groups: $U_{aQCM}(n_1 = 47, n_2 = 46) = 1310.5, p < 0.05, d_{Cohen} = 0.37$ and $U_{bQCM}(n_1 = 47, n_2 = 46) = 1384.0, p < 0.05, d_{Cohen} = 0.50$. Regarding the later test phase, however, we found

no significant group differences at all, indicating equal CM quality across all groups on average (see Table 8).

Table 8. ANOVA and Kruskal–Wallis test results regarding aQCM and bQCM indices.

Variable	Group								Kruskal–Wallis Test/ANOVA
	T+E		T+R		T–E		T–R		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
aQCM Index (LP)	106.30	37.87	78.13	35.60	124.70	59.82	113.55	50.70	$F(3, 88) = 4.23, p < 0.01$
bQCM Index (LP)	2.70	0.34	2.46	0.52	2.62	0.31	2.40	0.49	$\chi^2(3) = 8.66, p < 0.05$
aQCM Index (TP)	62.00	19.27	59.40	22.08	62.72	14.22	63.27	21.90	$F(3, 89) = 0.18, p = 0.91$
bQCM Index (TP)	2.40	0.44	2.50	0.38	2.37	0.40	2.32	0.38	$\chi^2(3) = 2.94, p = 0.40$

T+E = CM training group ($n = 23$) assigned to elaboration setting (learning material available during CM); T+R = CM training group ($n = 25$) assigned to retrieval setting (learning material not available during CM); T–E = control training group ($n = 23$) assigned to elaboration setting (learning material available during CM); T–R = control training group ($n = 22$) assigned to retrieval setting (learning material not available during CM); *M* = mean value; *SD* = standard deviation; aQCM Index = absolute quality of concept map index; bQCM = balanced quality of concept map index; LP = learning phase; TP = test phase.

3.5.2. Types of Mistakes

Additionally, we analyzed the type of the participants' CM mistakes by determining the concept map error ratio (number of methodical CM mistakes divided by the number of overall propositions specified) and the content-related error ratio (number of content-related mistakes divided by the number of overall propositions specified; see Table 9).

Table 9. Means and standard deviations regarding the type of the participants' CM mistakes.

Group		Concept Map Error Ratio (%)				Content-Related Error Ratio (%)			
		Learning Phase		Test Phase		Learning Phase		Test Phase	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Training	T+	7.0	12.0	2.0	4.0	10.0	8.0	21.0	16.0
	T-	10.0	11.0	7.0	8.0	9.0	7.0	22.0	15.0
Setting	R	10.0	14.0	4.0	6.0	12.0	9.0	21.0	15.0
	E	7.0	9.0	5.0	6.0	7.0	6.0	21.0	16.0

Means and standard deviations are reported as percentages, so an error ratio of $M = 7$ indicates that, on average, 7.0% of specified propositions were inaccurate (CM- or content-related); M = mean value; SD = standard deviation; T+ = CM training group ($n = 48$); T- = control training group ($n = 45$); R = retrieval setting (learning material not available during CM; $n = 47$); E = elaboration setting (learning material available during CM; $n = 46$).

Regarding these measures, we first performed a Mann–Whitney test to evaluate whether the CM-related mistakes differed between the participants of the two training conditions T+ and T-. An analysis of all four groups, in which the setting is also taken into account, was not advisable at this point, since it cannot be plausibly assumed that the setting influences the rate of merely methodical CM mistakes (see Section 1.1 and Section 1.2). Our results show that an extensive CM training significantly reduced the CM error ratios both in the learning phase $U(n_1 = 48, n_2 = 45) = 783.5, p < 0.05, d_{\text{Cohen}} = 0.49$, and in the test phase, $U(n_1 = 48, n_2 = 45) = 656.0, p < 0.001, d_{\text{Cohen}} = 0.72$.

Afterwards, we performed a second Mann–Whitney test to evaluate whether the content-related mistakes differed between the participants of the two setting conditions elaboration (E) and retrieval (R). An analysis of all four groups, in which the training is also taken into account, was again not advisable, as this time it cannot be plausibly assumed that merely methodical CM skills influence the rate of content-related mistakes (see Section 1.3). Regarding the learning phase, our results show that the elaboration groups' content-related error ratios were significantly lower than those of the retrieval groups: $U(n_1 = 47, n_2 = 46) = 697.5, p < 0.01, d_{\text{Cohen}} = 0.64$. However, these differences go missing when looking at the test phase. Accordingly, the retrieval groups' higher content-related error ratio in the learning phase, indicating more misconceptions, does not seem to have a meaningful impact on later performance.

3.5.3. Types of Propositions

Finally, we analyzed the type of the participants' CM propositions that were specified during the learning phase in the presence or absence of the learning material. Since the total number of specified propositions differed significantly between groups, $F(3, 88) = 5.84, p < 0.01, d_{\text{Cohen}} = 0.89$, we decided to report the respective ratios of proposition categories here. A Kruskal–Wallis test taking into account all four groups showed no statistically significant differences, but non-specific trends ($p < 0.10$) regarding the E- and R-proposition ratios (see Table 10), so we decided to take a closer look at the specific differences between the settings elaboration and retrieval, neglecting the affiliation to the T+ and T– groups.

Table 10. Kruskal–Wallis test results regarding different types of CM propositions specified.

Variable	Group								Chi ² Test
	T+E		T+R		T–E		T–R		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
R-proposition ratio (%)	93.0	10.0	87.0	16.0	88.0	10.0	84.0	14.0	$\chi^2(3) = 6.98, p = 0.07$
O-proposition ratio (%)	2.0	3.0	3.0	5.0	2.0	3.0	3.0	4.0	$\chi^2(3) = 1.73, p = 0.63$
E-proposition ratio (%)	1.0	2.0	1.0	3.0	0.4	1.0	3.0	5.0	$\chi^2(3) = 6.44, p = 0.09$

T+E = CM training group ($n = 23$) assigned to elaboration setting (learning material available during CM); T+R = CM training group ($n = 25$) assigned to retrieval setting (learning material not available during CM); T–E = control training group ($n = 23$) assigned to elaboration setting (learning material available during CM); T–R = control training group ($n = 22$) assigned to retrieval setting (learning material not available during CM); *M* = mean value; *SD* = standard deviation; R-proposition = recall-suggesting proposition; O-proposition = organization-suggesting proposition; E-proposition = elaboration-suggesting proposition; means and standard deviations for proposition ratios are reported as percentages, so a R-proposition ratio of $M = 93$ indicates that, on average, 93.0% of propositions in the concept map were recall-suggesting.

In respect of these potential setting effects, the respective Mann–Whitney tests indicate still no significant difference regarding R-propositions, $U(n_1 = 47, n_2 = 46) = 1259.5, p = 0.11$, but regarding E-propositions, $U(n_1 = 47, n_2 = 46) = 866.5, p < 0.05, d_{\text{Cohen}} = 0.35$. The latter result indicates that participants of the retrieval setting obviously specified more E-propositions ($M = 1.91; SD = 3.99$), although we had plausibly assumed the opposite, namely that participants who had the learning material available during CM would specify more ($M = 0.54; SD = 1.77$). Overall, the fact that there are significant differences between the settings, but not if the affiliation to the training groups is taken into account, can be interpreted as an

indication that the presence or absence of the learning material during CM influences the quality of propositions specified considerably stronger than an extensive CM training.

4. Discussion

In our study, we tried to find out to what extent CM- and knowledge-related learning success measures differ depending on CM training intensity and the (non-) availability of learning material during the creation of concept maps in order to identify determinants for a learning-effective implementation of CM in academic contexts. Overall, our results show a significantly positive but setting-independent effect of the CM strategy training on CM-related learning outcome but not necessarily on knowledge-related learning outcome. Regarding the different setting conditions, we found both advantages and disadvantages, but in the end, the groups seem to perform equally effective, implying that CM in retrieval settings could be as effective as in elaboration settings.

4.1. Decision on Hypothesis 1

Our first hypothesis based on the assumption that non-availability of learning material during CM (R groups) could be crucial for induced cognitive processes by increasing memory-related recall [37], and thus, binding cognitive resources required for further elaboration. Accordingly, we expected an increase in elaboration processes if the learning material was available during CM (E groups), resulting in a more elaborative character of the propositions that participants specified in their concept maps and an achievement of higher scores in subsequent knowledge tests regarding the covered learning content (see Section 1.4).

Regarding the general quality of specified propositions (aQCM and bQCM indices) and the content-related error ratios within the participants' concept maps of the learning phase, the E groups achieved higher scores on average than the R groups, but both differences levelled out until the test phase one week later (see Section 3.5.1 and 3.5.2). Furthermore, regarding the types of propositions, our results surprisingly show that participants of the E groups specified significantly fewer elaboration-suggesting propositions than participants of the R groups (see Section 3.5.3). Finally, regarding declarative, structural, and conceptual knowledge, our analyses did not reveal any significant advantages of the E groups (see Section 3.4). Accordingly, hypothesis 1 is not empirically supported by the data collected in our sample.

The result that, at least in the learning phase, the E groups generally specified higher quality propositions, which also offered fewer content-related mistakes, indicates that the availability of the learning material during CM obviously enables the learners to check whether their propositions specified are consistent with the text and, if necessary, to correct mistakes.

Nevertheless, this process did not lead to stable learning effects in our study, as such effects should have been reflected in the test phase one week later. Accordingly, it seems likely that the availability of the learning material during a specific CM task can positively influence methodical quality aspects but does not have a beneficial effect on the consolidation of the learning content itself. This finding is in line with that of Blunt and Karpicke [37], who also consider their elaboration groups' participants to focus more on detailed representations of encoded knowledge than on improvement of cue diagnosticity, occurring more likely in retrieval practice groups. Additionally, these considerations imply a plausible explanation for the counterintuitive finding that participants of the E groups specified significantly fewer elaboration-suggesting propositions than participants of the R groups. The E groups' participants may have focused more on methodically converting the information of the text into high-quality propositions, neglecting the integration of this information into prior knowledge structures, which is crucial for the specification of propositions of a more elaborative character. Conversely, the higher proportion of elaboration processes in the R groups could obviously not support better knowledge consolidation, since the two groups achieved similar scores regarding subsequent tests of declarative, structural, and conceptual knowledge. Accordingly, this result is inconsistent with the previous findings of Blunt and Karpicke [37] as well as O'Day and Karpicke [43], stating the superiority of retrieval practice over elaboration regarding knowledge acquisition. However, the fact that the availability of the learning material does not have a significant impact on knowledge acquisition during CM is considerable as this suggests that CM in retrieval settings could be as effective as CM in elaboration settings [38,90].

4.2. Decision on Hypothesis 2

Our second hypothesis was based on the assumption that a short methodical introduction on CM principles (T⁻ groups) is not sufficient to enable a successful use of the strategy since applying CM as an effective learning tool requires certain familiarity, especially with methodical aspects, and thus, an extensive training in CM [89,96,117,118]. Accordingly, we expected participants who took part in such a training (T⁺ groups) to show higher CM-related self-efficacy expectations, to achieve better results in analyzing and editing given concept maps, as well as to make fewer methodical mistakes when creating their own concept maps (see Section 1.4).

Regarding CM-related self-efficacy expectations, our results surprisingly did not show any significant group differences, i.e., participants of the T⁺ groups were as confident about their CM skills as those of the T⁻ groups (see Section 3.2). Furthermore, regarding analysis and editing of given concept maps within the error task, the T⁺ and T⁻ groups performed similar

concerning the number of errors detected, but T+ participants corrected these errors considerably more adequately (see Section 3.2). This result implies that the number of detected, and somehow corrected, errors is largely unaffected by the intensity of a CM training, but an extensive CM training can obviously decrease the bias toward correcting errors improperly, which in turn indicates a better understanding of methodical aspects of CM [79,88]. Finally, regarding the number of methodical mistakes made when creating own concept maps both in the learning and in the test phase, participants of the T+ groups clearly outperform those of the T- groups by showing significantly reduced the CM error ratios (see Section 3.5.2). Accordingly, the data collected in our sample partly provide empirical evidence for hypothesis 2.

The result that participants of the T+ groups were as confident about their CM skills as those of the T- groups seems surprising, but is in line with findings of our previous study [44]. The extensive training in using the demanding strategy of CM should actually have led to higher self-efficacy expectations of the T+ groups participants, especially since preliminary analyses did not reveal any significant group differences regarding experience with CM prior to our study (see Section 3.1). Accordingly, one explanation for the missing group difference could refer to a too homogeneous structure of our sample since it solely consisted of university students having many years of practical experience with learning and knowledge acquisition as well as the constant adaptation of new learning strategies. Therefore, it seems hardly surprising that all participants, regardless of their affiliation to T+ or T- groups, hold a certain degree of basic trust in their own academic abilities [119]. This consideration also implies a plausible explanation for the error task-related findings that the T+ group only deviated positively with regard to the incorrect error correction rate, whereas both groups achieved comparable results regarding error detection and proper error correction rate. If the assumption regarding the homogeneity of the sample in terms of experience in academic performance settings is correct, the aforementioned results highlight the need for a more difficult and thus better differentiating error task, including not only methodical errors but also content-related errors (see Section 2.4.2). However, the participants of the T+ group were obviously able to transfer their knowledge of methodical CM errors from the error task's more passive reception context to an active production context, which is reflected in the result that their own concept maps contained significantly fewer methodical CM mistakes both in the learning and test phase.

4.3. Decision on Hypothesis 3

Our third hypothesis consolidated hypotheses 1 and 2, as we expected that a combination of both advantageous factors, an extensive CM training and the availability of the learning material during CM (T+E), would be most effective regarding CM- and knowledge-related learning outcomes as well as their reflection in a corresponding metacognitive assessment of the participants regarding their own learning success (see Section 1.4).

However, our analyses could not reveal any significant advantages of the T+E groups compared to the other three groups (T+R, T-E, and T-R) regarding the general concept map quality (aQCM and bQCM indices; see Section 3.5.1), the types of specified propositions (E-, O-, or R-propositions; see Section 3.5.3), or the three domains of knowledge (declarative, structural, and conceptual knowledge; see Section 3.4). Consistent with this, the metacognitive predictions in terms of judgements of leaning (JOLs) between the four groups did not differ from one another (see Section 3.3). Accordingly, hypothesis 3 does not find any empirical support from the data collected in our sample.

As this hypothesis addressed a particular effectiveness of combined training and setting conditions, it is directly dependent on the two previous hypotheses. Accordingly, concerning the homogeneous performance of the four groups with regard to CM skills and knowledge measures, the same explanations apply as we have already specified in Section 4.2. In the following, we will therefore only refer to those findings that go beyond this and thus involve further implications.

The result that the aQCM and bQCM indices of the four groups differed significantly in the learning phase in favor of the T+E and T-E groups, but not any longer in the test phase, indicates that the presence of the text could temporarily compensate a missing training on CM regarding the general quality of specified propositions, but this positive effect obviously starts fading to the same degree as the processes of forgetting occur. The assumption that significant processes of forgetting actually took place in our study is supported by the finding that the concept maps constructed by the participants of all groups in the test phase contained, on average, 2.2 times more content-related mistakes as those of the learning phase. In this context, the question arises whether a different experimental timeline (see Section 2.2) should be chosen in order to counteract such processes of forgetting. Beyond that, a potentially low level of the participants' learning motivation must be taken into account, which is particularly suggested by the results of the knowledge tests, ranking at a consistently low level of only around 40% to 60% of the achievable scores. With regard to the merely numerical scores of the knowledge test results, it seems to be slightly more advantageous for declarative knowledge acquisition if the

learning material is available to the learners during CM if they have not received an extensive CM training before (T–E group). Regarding the conceptual knowledge test, it is noticeable that, on the one hand, the control training group in an elaborative setting (T–E) performed best and, on the other hand, the control training group in a retrieval setting (T–R) performed worst. The CM training groups (T+E and T+R) are in between those two, with participants in an elaboration setting performing slightly better than those in a retrieval setting. Therefore, in terms of knowledge acquisition, it might be generally better to have the learning material available during CM. If the learning material is not available, however, it seems advantageous to be familiar with CM to support knowledge acquisition during the use of the strategy as a learning tool.

Beyond that, such a familiarity with CM also seems to be beneficial in terms of metacognitive processes, since, with regard to the merely numerical scores of their judgements of learning (JOLs), the participants of the T+ groups assessed their future learning outcome slightly more accurately than those of the T– groups. In this context, it is also interesting that the participants in elaboration settings predicted a better learning outcome on average than those in retrieval settings, which confirms the results of previous studies, showing that participants in retrieval settings obviously refer to the perceived ease of memory recall of information when giving such assessments, whereas participants in elaboration settings refer to the perceived ease of information processing [37,90,91]. Nonetheless, across all four groups, the participants assessed that one week later they would only be able to remember about 50% of the information they had learned. Since the learning text on cell biology was previously assessed by experts as understandable, structured, and appropriate for the university level, it can be assumed that the JOLs of the participants were primarily quantitatively oriented towards the simple number of words. This consideration finds support if our correlational findings are also taken into account, as they suggest that the JOLs of the participants are obviously much more related to declarative and structural than to conceptual knowledge. This interesting secondary finding can most likely be attributed to the fact that our participants are simply unaccustomed to thinking in a conceptual manner, since in their everyday university life, they tend to focus on the acquisition and reproduction of declarative knowledge and sometimes structural knowledge, while conceptual knowledge, i.e., a deeper understanding of relations and dependencies between central concepts of respective fields, seems to be less required regarding completing most university study programs successfully [120], which is also reflected in popular terms such as “bulimic learning” [121].

4.4. Limitations

The previously discussed heterogeneous and sometimes unexpected findings should additionally be evaluated against the background of the study's major limitations:

- (1) Our findings suggest that the structure of our sample could have been too homogeneous and thus might have undermined the occurrence of several training and setting effects. Our sample solely consisted of university students having many years of practical experience with learning and knowledge acquisition as well as the constant adaptation of new learning strategies. In addition, it could be assumed that our student participants may not elaborate new information habitually, but are generally more familiar with rote learning, which in turn is supported by the segregated state of conceptual knowledge in our data. Therefore, we predominantly interpret the unexpectedly missing significant group differences as an effect of sampling. In this respect, it is to be expected that referring to other samples than students could reveal more distinct group differences if the participants belong to a population in which learning activities are a less central everyday topic. Furthermore, if this assumption regarding the homogeneity of our sample in terms of experience in academic performance settings is correct, our results also highlight the need for better differentiating assessments regarding few measures. In this regard, especially the error task should probably include content-related errors in addition to the methodical ones (see Section 2.4.2), as, in our previous study, we already found a similar result, indicating the task's low difficulty [44].
- (2) Additionally, it should be taken into account that the final number of $N = 93$ participants was too small to reach sufficient power, ranking below 0.50 for the majority of the statistical analyses of differences performed. Therefore, a replication of our study would be desirable, taking a larger sample in order to either detect small but possibly still relevant effects or to be able to negate them without fail. Post-hoc power analyses have shown that our sample should have consisted of at least $N = 150$ participants in order to be able to detect small to medium-sized effects at a power level of 0.80 to 0.90. Considering the intensive and time-consuming care of the participants over a period of six weeks, it was not possible for us to carry out our study on a larger sample due to limited resources.
- (3) Finally, the LGVT's [109] reliability was completely insufficient (<0.60) for inexplicable reasons, so we had to exclude these scales assessing reading speed and comprehension from further statistical analyses to avoid impairment of our conclusions' validity. Of course, such an exclusion is always associated with a loss of information, as

we were not able to determine any baseline differences regarding reading speed and comprehension. Nevertheless, we think that this did not affect internal validity excessively, since it can be assumed that students in general offer comparable reading skills at a high level in order to be able to successfully complete their studies at all.

4.5. Prospects for Future Research

Even if our results turned out to be more heterogeneous than expected, they allow for tentative conclusions and imply suggestions for the design of prospective studies in this area that go beyond solely removing the previously discussed limitations of our study.

Regarding the setting conditions, the two groups, elaboration vs. retrieval, have both advantages and disadvantages, and in the end, they seem to perform equally effectively. This finding implies the need for taking a closer look at considerably more differentiated constellations of experimental conditions in order to uncover possible differential effects of individual parameters in the context of learning. For example, it is conceivable that corresponding effects only come into operation after a specific number of repetitions of learning phases and/or another time delay between learning and test phase, which would require respective variations and repeated short-interval follow-up measurements.

Regarding the training conditions, we were able to show that an extensive CM training has a positive effect on methodical CM skills, in particular, but not necessarily on knowledge-related learning success. This result is partly inconsistent with the results of our previous study, in which we found both significantly better CM skills and knowledge-related advantages of participants who received an extensive training [44]. This missing effect of an extensive training on knowledge acquisition in our present study allows for drawing two possible conclusions:

- (1) Possibly, the 30-minute short introduction of the T- groups regarding to the most important CM principles was just as effective as the extensive CM training of the T+ group. Such an effect could be attributed in two different ways. Either it again must be interpreted against the background of our sample of experienced and generally successful learners (which seems unlikely, since the overall performance of the groups ranged on a comparatively low level), or it is due to a generally low level of learning motivation, since incentives (e.g., study credits) for participation in our study were not offered to the students (see Section 2.1). If the latter should be true, it would be necessary to consider the participants' learning motivation as a covariate in future studies.
- (2) Additionally, it is once more possible that resounding effects of an extensive CM training could be more successfully activated by a different experimental timeline than the one we chose (see Section 2.2). For example, it is conceivable that the training's success

differs depending on the overall number of repetitions and/or the duration of individual training sessions and/or the time interval between these sessions, which could be clarified in future studies by systematic variation.

Finally, it would be interesting to investigate the potential effect of implementing an effective training with alternating phases of text availability vs. non-availability on overall learning success. In the studies of Karpicke and Blunt [38], Blunt and Karpicke [37], and O'Day and Karpicke [43], such an alternation was implemented, but the participants received only a short introduction to CM (~T- condition in our study) and furthermore, these studies' textual learning material consisted of approximately only 300 words, so ecological validity in terms of appropriateness regarding university level can be questioned. Future studies could, therefore, focus on experimental conditions that are closer to university learning and performance contexts, as we did in our study.

The design and implementation of such studies would help identify the conditions for an efficient use of CM and thus provide learners with an effective strategy that can support a deeper understanding of a field.

Author Contributions: Conceptualization, J.G., L.B.B. and V.D.E.W.; Methodology, J.G., L.B.B. and V.D.E.W.; Formal Analysis, J.G., L.B.B. and V.D.E.W.; Investigation, L.B.B.; Resources, J.G. and V.D.E.W.; Data Curation, L.B.B. and V.D.E.W.; Writing—Original Draft Preparation, L.B.B. and V.D.E.W.; Writing—Review and Editing, J.G., L.B.B. and V.D.E.W.; Visualization, L.B.B. and V.D.E.W.; Supervision, J.G.; Project Administration, J.G.; Funding Acquisition, J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Research Foundation (DFG), grant number GR 4763/2-1.

Institutional Review Board Statement: Our study was approved by the German Research Foundation (DFG; project number: GR 4763/2-1). Before participation, all subjects received a detailed written subject information in accordance with the current ethical guidelines laid down by the University of Cologne (Germany) and the German Psychological Society [122], including the following information: aims and course of the investigation, absolute voluntariness of participation, possibility of dropping out of participation at any time, guaranteed protection of data privacy (collection of only anonymized data), possibility of requesting data cancelation at any time, no-risk character of study participation, and contact information in case of any questions or problems. Written informed consent was obtained from all participants prior to the study, according to the current version of the Declaration of Helsinki

[123,124]. Anonymity was ensured by using individual codes that the students created themselves. Data storage meets current European data protection regulations [125].

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We thank Ellen Aschermann (Department of Psychology, University of Cologne) for the constructive discussions we had on our project as well as for her helpful support in organizational matters and in acquiring participants.

Conflicts of Interest: The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Conradt, C.; Sotiriou, S.A.; Bogner, F.X. How Creativity in STEAM Modules Intervenes with Self-Efficacy and Motivation. *Educ. Sci.* 2020, *10*, 70.
2. Mulnix, A.B.; Vandegrift, E.V.H. A Tipping Point in STEM Education Reform. *J. Coll. Sci. Teach.* 2014, *43*, 14–16.
3. Tripto, J.; Assaraf, O.B.Z.; Snapir, Z.; Amit, M. How is the body's systemic nature manifested amongst high school biology students? *Instr. Sci.* 2017, *45*, 73–98.
4. Udovic, D.; Morris, D.; Dickman, A.; Postlethwait, J.; Wetherwax, P. Workshop Biology: Demonstrating the Effectiveness of Active Learning in an Introductory Biology Course. *BioScience* 2002, *52*, 272–281.
5. Arwood, L. Teaching Cell Biology to Nonscience Majors Through Forensics, Or How to Design a Killer Course. *Cell Biol. Educ.* 2004, *3*, 131–138.
6. Jesionkowska, J.; Wild, F.; Deval, Y. Active Learning Augmented Reality for STEAM Education—A Case Study. *Educ. Sci.* 2020, *10*, 198.
7. Labov, J.B.; Reid, A.H.; Yamamoto, K.R. Integrated Biology and Undergraduate Science Education: A New Biology Education for the Twenty-First Century? *CBE Life Sci. Educ.* 2010, *9*, 10–16.
8. Lathwesen, C.; Belova, N. Escape Rooms in STEM Teaching and Learning—Prospective Field or Declining Trend? A Literature Review. *Educ. Sci.* 2021, *11*, 308.

9. Silveira, L.F.; Xavier, C.S.; Perez, M.A.; Bozza, D.A.; Furtado-Alle, L.; Hass, I.; Tureck, L.V. Unravelling the Central Dogma of Biology in an active way: A case study. *J. Biol. Educ.* 2021, *55*, 1–14.
10. Swaid, S.I. Bringing Computational Thinking to STEM Education. *Procedia Manuf.* 2015, *3*, 3657–3662.
11. Young, P.W. Corrigendum: Student-Produced Video of Role-Plays on Topics in Cell Biology and Biochemistry: A Novel Undergraduate Group Work Exercise. *Front. Educ.* 2020, *5*, 115.
12. McGreevy, K.M.; Church, F.C. Active Learning: Subtypes, Intra-Exam Comparison, and Student Survey in an Undergraduate Biology Course. *Educ. Sci.* 2020, *10*, 185.
13. Aguilera, D.; Ortiz-Revilla, J. STEM vs. STEAM Education and Student Creativity: A Systematic Literature Review. *Educ. Sci.* 2021, *11*, 331.
14. Preszler, R. Cooperative Concept Mapping: Improving Performance in Undergraduate Biology. *J. Coll. Sci. Teach.* 2004, *33*, 30–35.
15. Tullis, J.G.; Goldstone, R.L. Instruction in Computer Modeling Can Support Broad Application of Complex Systems Knowledge. *Front. Educ.* 2017, *2*, 4.
16. Verhoeff, R.P.; Knippels, M.-C.P.J.; Gilissen, M.G.R.; Boersma, K.T. The Theoretical Nature of Systems Thinking. Perspectives on Systems Thinking in Biology Education. *Front. Educ.* 2018, *3*, 40.
17. Campbell, N.A.; Urry, L.A.; Cain, M.L.; Wasserman, S.A.; Minorsky, P.V.; Orr, R.B. *Biology: A Global Approach*, 12th ed.; Pearson Education Limited: Harlow, UK, 2021; ISBN 978-1292341637.
18. Barak, J.; Sheva, B.; Gorodetsky, M.; Gurion, B. As ‘process’ as it can get: Students’ understanding of biological processes. *Int. J. Sci. Educ.* 1999, *21*, 1281–1292.
19. Brandstädter, K.; Harms, U.; Großschedl, J. Assessing System Thinking through Different Concept-Mapping Practices. *Int. J. Sci. Educ.* 2012, *34*, 2147–2170.
20. Garvin-Doxas, K.; Klymkowsky, M.; Elrod, S. Building, Using, And Maximizing the Impact of Concept Inventories in the Biological Sciences: Report on a National Science Foundation-sponsored Conference on the Construction of Concept Inventories in the Biological Sciences. *CBE Life Sci. Educ.* 2007, *6*, 277–282.
21. Grotzer, T.A.; Solis, S.L.; Tutwiler, M.S.; Cuzzolino, M.P. A study of students’ reasoning about probabilistic causality: Implications for understanding complex systems and for instructional design. *Instr. Sci.* 2017, *45*, 25–52.

22. Haslam, F.; Treagust, D.F. Diagnosing secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier multiple choice instrument. *J. Biol. Educ.* 1987, *21*, 203–211.
23. Hmelo-Silver, C.E.; Azevedo, R. Understanding Complex Systems: Some Core Challenges. *J. Learn. Sci.* 2006, *15*, 53–61.
24. Shi, J.; Wood, W.B.; Martin, J.M.; Guild, N.A.; Vicens, Q.; Knight, J.K. A Diagnostic Assessment for Introductory Molecular and Cell Biology. *CBE Life Sci. Educ.* 2010, *9*, 453–461.
25. Mandl, H.; Friedrich, H.F. *Handbuch Lernstrategien [Handbook of Learning Strategies]*; Hogrefe: Göttingen, Germany, 2006; ISBN 978-3801718138.
26. Anderson, J.R. *The Architecture of Cognition*; Lawrence Erlbaum Associates, Inc.: Hillsdale, NJ, USA, 1983; ISBN 978-0674044258.
27. Weinstein, C.E.; Mayer, R.E. The teaching of learning strategies. In *Handbook of Research on Teaching*, 3rd ed.; Wittrock, M.C., Ed.; Macmillan: New York, NY, USA, 1986; pp. 315–327.
28. Kalyuga, S. Knowledge elaboration: A cognitive load perspective. *Learn. Instr.* 2009, *19*, 402–410.
29. Wild, K.-P.; Schiefele, U. Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens [Learning strategies in academic studies: Results on factorial structure and reliability of a new questionnaire]. *J. Individ. Differ.* 1994, *15*, 185–200.
30. Wadouh, J. Vernetzung und Kumulatives Lernen im Biologieunterricht der Gymnasialklasse 9 [Cross-Linked Thinking and Cumulative Learning in Grade 9 Secondary School Biology Classes]. Ph.D. Thesis, University Duisburg-Essen, Duisburg, Germany, 2007.
31. Carpenter, S.K.; Pashler, H.; Cepeda, N.J. Using tests to enhance 8th grade students' retention of U.S. history facts. *Appl. Cogn. Psychol.* 2009, *23*, 760–771.
32. Karpicke, J.D. Retrieval-based learning: A decade of progress. In *Cognitive Psychology of Memory*, 2nd ed.; Wixted, J.T., Ed.; Academic Press: Oxford, UK, 2017; pp. 487–514.
33. Karpicke, J.D.; Roediger, H.L. The Critical Importance of Retrieval for Learning. *Science* 2008, *319*, 966–968.

34. Pyc, M.A.; Rawson, K.A. Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *J. Mem. Lang.* 2009, *60*, 437–447.
35. Roediger, H.L.; Karpicke, J.D. The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspect. Psychol. Sci.* 2006, *1*, 181–210.
36. Großschedl, J.; Tröbst, S. Biologie lernen durch Concept Mapping: Bedeutung eines Lernstrategietrainings für kognitive Belastung, kognitive Prozesse und Lernleistung—Kurzdarstellung des DFG-Projekts [Learning biology using concept mapping: Importance of a learning strategy training for cognitive load, cognitive processes and learning performance—Brief description of the DFG project]. *Zeitschrift für Didaktik der Biologie (ZDB)-Biologie Lehren und Lernen* 2018, *22*, 20–30.
37. Blunt, J.R.; Karpicke, J.D. Learning with retrieval-based concept mapping. *J. Educ. Psychol.* 2014, *106*, 849–858.
38. Karpicke, J.D.; Blunt, J.R. Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. *Science* 2011, *331*, 772–775.
39. Karpicke, J.D.; Smith, M.A. Separate mnemonic effects of retrieval practice and elaborative encoding. *J. Mem. Lang.* 2012, *67*, 17–29.
40. Brame, C.J.; Biel, R. Test-Enhanced Learning: The Potential for Testing to Promote Greater Learning in Undergraduate Science Courses. *CBE Life Sci. Educ.* 2015, *14*, es4.
41. Karpicke, J.D. Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *J. Exp. Psychol. Gen.* 2009, *138*, 469–486.
42. Moreira, B.F.T.; Pinto, T.S.S.; Starling, D.S.V.; Jaeger, A. Retrieval Practice in Classroom Settings: A Review of Applied Research. *Front. Educ.* 2019, *4*, 5.
43. O’Day, G.M.; Karpicke, J.D. Comparing and combining retrieval practice and concept mapping. *J. Educ. Psychol.* 2021, *113*, 986–997.
44. Becker, L.B.; Welter, V.D.E.; Aschermann, E.; Großschedl, J. Comprehension-Oriented Learning of Cell Biology: Do Different Training Conditions Affect Students’ Learning Success Differentially? *Educ. Sci.* 2021, *11*, 438.
45. Morse, D.; Jutras, F. Implementing Concept-based Learning in a Large Undergraduate Classroom. *CBE Life Sci. Educ.* 2008, *7*, 243–253.
46. Okebukola, P.A. Attaining meaningful learning of concepts in genetics and ecology: An examination of the potency of the concept-mapping technique. *J. Res. Sci. Teach.* 1990, *27*, 493–504.

47. Okebukola, P.A.; Jegede, O.J. Cognitive preference and learning mode as determinants of meaningful learning through concept mapping. *Sci. Educ.* 1988, *72*, 489–500.
48. Pruet, J.L.; Weigel, E.G. Concept Map Assessment Reveals Short-Term Community-Engaged Fieldwork Enhances Sustainability Knowledge. *CBE Life Sci. Educ.* 2020, *19*, ar38.
49. Hegarty-Hazel, E.; Prosser, M. Relationship between Students' Conceptual Knowledge and Study Strategies—Part 2: Student Learning in Biology. *Int. J. Sci. Educ.* 1991, *13*, 303–312.
50. Martínez, G.; Pérez, Á.L.; Suero, M.I.; Pardo, P.J. The Effectiveness of Concept Maps in Teaching Physics Concepts Applied to Engineering Education: Experimental Comparison of the Amount of Learning Achieved with and without Concept Maps. *J. Sci. Educ. Technol.* 2013, *22*, 204–214.
51. Gahr, A.A. Cooperative Chemistry: Concept Mapping in the Organic Chemistry Lab. *J. Coll. Sci. Teach.* 2003, *32*, 311.
52. Jack, G.U. Concept Mapping and Guided Inquiry as Effective Techniques for Teaching Difficult Concepts in Chemistry: Effect on Students' Academic Achievement. *J. Educ. Pract.* 2013, *4*, 9–15.
53. Cañas, A.J.; Novak, J.D. Concept Mapping Using CmapTools to Enhance Meaningful Learning. In *Knowledge Cartography: Software Tools and Mapping Techniques*, 2nd ed.; Okada, A., Buckingham Shum, S.J., Sherborne, T., Eds.; Springer: London, UK, 2014; pp. 23–45.
54. Cronin, P.J.; Dekkers, J.; Dunn, J.G. A procedure for using and evaluating concept maps. *Res. Sci. Educ.* 1982, *12*, 17–24.
55. Derbentseva, N.; Safayeni, F.; Cañas, A.J. Concept maps: Experiments on dynamic thinking. *J. Res. Sci. Teach.* 2007, *44*, 448–465.
56. Novak, J.D. Concept mapping: A useful tool for science education. *J. Res. Sci. Teach.* 1990, *27*, 937–949.
57. Novak, J.D. *Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations*, 2nd ed.; Routledge: New York, NY, USA, 2010; ISBN 978-0415991858.

58. Novak, J.D.; Cañas, A. *The Theory Underlying Concept Maps and How to Construct Them*; Florida Institute for Human and Machine Cognition: Pensacola, FL, USA, 2008; Available online: <http://cmap.ihmc.us/docs/pdf/TheoryUnderlyingConceptMaps.pdf> (accessed on 21 August 2021).
59. Quillian, M.R. Semantic networks. In *Semantic Information Processing*; Minsky, M.L., Ed.; MIT Press: Cambridge, MA, USA, 1968; pp. 227–270.
60. Collins, A.M.; Quillian, M.R. Retrieval time from semantic memory. *J. Verbal Learn. Verbal Behav.* 1969, 8, 240–247.
61. Buchner, A.; Brandt, M. Gedächtniskonzeptionen und Wissensrepräsentationen [Conceptions of memory and representations of knowledge]. In *Allgemeine Psychologie [General Psychology]*; Müsseler, J., Rieger, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 401–434.
62. Renkl, A.; Nückles, M. Lernstrategien der externen Visualisierung [Learning strategies of external visualization]. In *Handbuch Lernstrategien [Handbook of Learning Strategies]*; Mandl, H., Friedrich, H.F., Eds.; Hogrefe: Göttingen, Germany, 2006; pp. 135–147.
63. den Elzen-Rump, V.; Leutner, D. Naturwissenschaftliche Sachtexte verstehen—Ein computerbasiertes Trainingsprogramm für Schüler der 10. Jahrgangsstufe zum selbstregulierten Lernen mit einer Mapping-Strategie [Understanding scientific texts—A computer-based training program for 10th grade students for self-regulated learning with a mapping strategy]. In *Selbstregulation Erfolgreich Fördern [Promoting Self-Regulation Successfully]*; Landmann, M., Schmitz, B., Eds.; Kohlhammer: Stuttgart, Germany, 2007; pp. 251–268.
64. Jüngst, K.L. Studien zur didaktischen Nutzung von Concept Maps [Studies on the didactic use of concept maps]. *Unterrichtswissenschaft* 1995, 3, 229–250.
65. Jüngst, K.L.; Strittmatter, P. Wissensstrukturdarstellung: Theoretische Ansätze und praktische Relevanz [Representation of knowledge structures: Theoretical approaches and practical relevance]. *Unterrichtswissenschaft* 1995, 23, 194–207.

66. Cañas, A.; Coffey, J.; Carnot, M.J.; Feltovich, P.; Hoffman, R.R.; Feltovich, J.; Novak, J.D. A Summary of Literature Pertaining to the Use of Concept Mapping Techniques and Technologies for Education and Performance Support. Available online: <https://www.ihmc.us/users/acanas/Publications/ConceptMapLitReview/IHMC%20Literature%20Review%20on%20Concept%20Mapping.pdf> (accessed on 21 August 2021).
67. Cadorin, L.; Bagnasco, A.; Rocco, G.; Sasso, L. An integrative review of the characteristics of meaningful learning in healthcare professionals to enlighten educational practices in health care. *Nurs. Open* 2014, *1*, 3–14.
68. Hattie, J. *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*, 1st ed.; Routledge: New York, NY, USA, 2009; ISBN 978-0415476171.
69. Haugwitz, M. Kontextorientiertes Lernen und Concept Mapping im Fach Biologie [Context-Oriented Learning and Concept Mapping in Biology]. Ph.D. Thesis, University Duisburg-Essen, Duisburg, Germany, 2009.
70. Hilbert, T.S.; Renkl, A. Concept mapping as a follow-up strategy to learning from texts: What characterizes good and poor mappers? *Instr. Sci.* 2008, *36*, 53–73.
71. Nesbit, J.C.; Adesope, O.O. Learning with Concept and Knowledge Maps: A Meta-Analysis. *Rev. Educ. Res.* 2006, *76*, 413–448.
72. Mayer, R.E. Rote Versus Meaningful Learning. *Theory Pract.* 2002, *41*, 226–232.
73. Nückles, M.; Hübner, S.; Dümer, S.; Renkl, A. Expertise reversal effects in writing-to-learn. *Instr. Sci.* 2010, *38*, 237–258.
74. Bramwell-Lalor, S.; Rainford, M. The Effects of Using Concept Mapping for Improving Advanced Level Biology Students' Lower- and Higher-Order Cognitive Skills. *Int. J. Sci. Educ.* 2014, *36*, 839–864.
75. Kinchin, I.M. Visualising knowledge structures in biology: Discipline, curriculum and student understanding. *J. Biol. Educ.* 2011, *45*, 183–189.
76. Mintzes, J.J.; Wandersee, J.H.; Novak, J.D. Meaningful learning in science: The human constructivist perspective. In *Handbook of Academic Learning: Construction of Knowledge*; Phye, G.D., Ed.; Academic Press: San Diego, CA, USA, 1997; pp. 405–447.
77. Ritchhart, R.; Turner, T.; Hadar, L. Uncovering students' thinking about thinking using concept maps. *Metacogn. Learn.* 2009, *4*, 145–159.

78. Salmon, D.; Kelly, M. *Using Concept Mapping to Foster Adaptive Expertise: Enhancing Teacher Metacognitive Learning to Improve Student Academic Performance*; Peter Lang: New York, NY, USA, 2014; ISBN 978-1433122705.
79. Kapuza, A. How Concept Maps with and without a List of Concepts Differ: The Case of Statistics. *Educ. Sci.* 2020, *10*, 91.
80. Ritter, R.; Wehner, A.; Lohaus, G.; Krämer, P. Pre-service Teachers' Beliefs about Inclusive Education before and after Multi-Compared to Mono-professional Co-teaching: An Exploratory Study. *Front. Educ.* 2019, *4*, 101.
81. Marzetta, K.; Mason, H.; Wee, B. 'Sometimes They Are Fun and Sometimes They Are Not': Concept Mapping with English Language Acquisition (ELA) and Gifted/Talented (GT) Elementary Students Learning Science and Sustainability. *Educ. Sci.* 2018, *8*, 13.
82. Andrews, K.E.; Tressler, K.D.; Mintzes, J.J. Assessing environmental understanding: An application of the concept mapping strategy. *Environ. Educ. Res.* 2008, *14*, 519–536.
83. Martin, B.L.; Mintzes, J.J.; Clavijo, I.E. Restructuring knowledge in Biology: Cognitive processes and metacognitive reflections. *Int. J. Sci. Educ.* 2000, *22*, 303–323.
84. Mintzes, J.J.; Wandersee, J.H.; Novak, J.D. Assessing understanding in biology. *J. Biol. Educ.* 2001, *35*, 118–124.
85. Pearsall, N.R.; Skipper, J.E.J.; Mintzes, J.J. Knowledge restructuring in the life sciences: A longitudinal study of conceptual change in biology. *Sci. Educ.* 1997, *81*, 193–215.
86. Quinn, H.J.; Mintzes, J.J.; Laws, R.A. Successive Concept Mapping: Assessing Understanding in College Science Classes. *J. Coll. Sci. Teach.* 2003, *33*, 12–16.
87. Hilbert, T.; Nückles, M.; Matzel, S. Concept mapping for learning from text: Evidence for a worked-out-map-effect. In Proceedings of the 8th International Conference for the Learning Sciences, Utrecht, The Netherlands, 23–28 June 2008; International Society of the Learning Sciences: Utrecht, The Netherlands, 2008; Volume 1, pp. 358–365.
88. Sumfleth, E.; Neuroth, J.; Leutner, D. Concept Mapping—eine Lernstrategie muss man lernen [Concept mapping—A learning strategy has to be learned]. *CHEMKON* 2010, *17*, 66–70.
89. Mintzes, J.J.; Canas, A.; Coffey, J.; Gorman, J.; Gurley, L.; Hoffman, R.; McGuire, S.Y.; Miller, N.; Moon, B.; Trifone, J.; et al. Comment on “Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping”. *Science* 2011, *334*, 453.

90. Karpicke, J.D. Retrieval-Based Learning: Active Retrieval Promotes Meaningful Learning. *Curr. Dir. Psychol. Sci.* 2012, 21, 157–163.
91. Roediger, H.L.; Karpicke, J.D. Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention. *Psychol. Sci.* 2006, 17, 249–255.
92. Karpicke, J.D.; Grimaldi, P.J. Retrieval-Based Learning: A Perspective for Enhancing Meaningful Learning. *Educ. Psychol. Rev.* 2012, 24, 401–418.
93. Koriat, A. Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *J. Exp. Psychol. Gen.* 1997, 126, 349–370.
94. Koriat, A. Metacognition and consciousness. In *Cambridge Handbook of Consciousness*; Zelazo, P.D., Moscovitch, M., Thompson, E., Eds.; Cambridge University Press: New York, NY, USA, 2007; pp. 289–325.
95. Ajaja, O.P. Concept mapping as a study skill. *Int. J. Educ. Sci.* 2011, 3, 49–57.
96. Schwendimann, B. Concept Mapping. In *Encyclopedia of Science Education*; Gunstone, R., Ed.; Springer: Dordrecht, The Netherlands, 2015; pp. 198–202.
97. Jonassen, D.H.; Beissner, K.; Yacci, M. *Structural Knowledge: Techniques for Representing, Conveying, And Acquiring Structural Knowledge*; Lawrence Erlbaum Associates, Inc.: Hillsdale, NJ, USA, 1993; ISBN 978-0805810097.
98. Chularut, P.; DeBacker, T.K. The influence of concept mapping on achievement, self-regulation, and self-efficacy in students of English as a second language. *Contemp. Educ. Psychol.* 2004, 29, 248–263.
99. Arnaudin, M.W.; Mintzes, J.J. Students' alternative conceptions of the human circulatory system: A cross-age study. *Sci. Educ.* 1985, 69, 721–733.
100. Hilbert, T.S.; Renkl, A. Learning how to use a computer-based concept-mapping tool: Self-explaining examples helps. *Comput. Human Behav.* 2009, 25, 267–274.
101. Großschedl, J.; Harms, U. Effekte metakognitiver Prompts auf den Wissenserwerb beim Concept Mapping und Notizen Erstellen [Effects of metacognitive prompts on knowledge acquisition in concept mapping and note taking]. *Z. Didakt. Nat.* 2013, 19, 375–395.
102. Großschedl, J.; Harms, U. Metakognition—Denken aus der Vogelperspektive [Metacognition—Thinking from a bird's eye view]. In *Biologie Methodik [Methodology of Biology]*, 4th ed.; Spörhase-Eichmann, U., Ruppert, W., Eds.; Cornelsen: Berlin, Germany, 2018; pp. 48–52.
103. Ajaja, O.P. Which way do we go in biology teaching? Lecturing, concept mapping, cooperative learning or learning cycle? *Electron. J. Sci. Educ.* 2013, 17, 1–37.

104. Reader, W.; Hammond, N. Computer-based tools to support learning from hypertext: Concept mapping tools and beyond. *Comput. Educ.* 1994, 22, 99–106.
105. McClure, J.R.; Sonak, B.; Suen, H.K. Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *J. Res. Sci. Teach.* 1999, 36, 475–492.
106. Champagne Queloz, A.; Klymkowsky, M.W.; Stern, E.; Hafen, E.; Köhler, K. Diagnostic of students' misconceptions using the Biological Concepts Instrument (BCI): A method for conducting an educational needs assessment. *PLoS ONE* 2017, 12, e0176906.
107. Großschedl, J.; Mahler, D.; Harms, U. Construction and Evaluation of an Instrument to Measure Content Knowledge in Biology: The CK-IBI. *Educ. Sci.* 2018, 8, 145.
108. Kleickmann, T.; Großschedl, J.; Harms, U.; Heinze, A.; Herzog, S.; Hohenstein, F.; Köller, O.; Kröger, J.; Lindmeier, A.; Loch, C. Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer—Testentwicklung im Rahmen des Projekts KiL [Professional knowledge of student teachers of mathematical and scientific subjects—Test development within the KiL project]. *Unterrichtswissenschaft* 2014, 42, 280–288.
109. Schneider, W.; Schlagmüller, M.; Ennemoser, M. *Lesegeschwindigkeits- und Verständnistest für die Klassen 6–12 [Reading Speed and Comprehension Test for Grades 6–12]*; Hogrefe: Göttingen, Germany, 2007; ISBN 978-3403066385.
110. Döring, N.; Bortz, J.; Poeschl-Guenther, S. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften [Research Methods and Evaluation in the Social and Human Sciences]*, 5th ed.; Springer: Berlin/Heidelberg, Germany, 2016.
111. Wirtz, M.; Caspar, F. *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen [Interrater Consistency and Interrater Reliability: Methods for Determining and Improving the Reliability of Assessments Using Observational Systems and Rating Scales]*, 1st ed.; Hogrefe: Göttingen, Germany, 2002; ISBN 978-3801716462.
112. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* 2016, 15, 155–163.
113. Townsend, C.L.; Heit, E. Judgments of learning and improvement. *Mem. Cognit.* 2011, 39, 204–216.
114. Anderson, J.R. *Cognitive Psychology and Its Implications*, 9th ed.; Worth Publishers: New York, NY, USA, 2020; ISBN 978-1319067113.

115. Watson, M.K.; Pelkey, J.; Noyes, C.R.; Rodgers, M.O. Assessing Conceptual Knowledge Using Three Concept Map Scoring Methods. *J. Eng. Educ.* 2016, *105*, 118–146.
116. Großschedl, J.; Harms, U. Assessing conceptual knowledge using similarity judgments. *Stud. Educ. Eval.* 2013, *39*, 71–81.
117. Roessger, K.M.; Daley, B.J.; Hafez, D.A. Effects of teaching concept mapping using practice, feedback, and relational framing. *Learn. Instr.* 2018, *54*, 11–21.
118. Schroeder, N.L.; Nesbit, J.; Anguiano, C.J.; Adesope, O.O. Studying and Constructing Concept Maps: A Meta-Analysis. *Educ. Psychol. Rev.* 2018, *30*, 431–455.
119. Sears, D.O. College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *J. Pers. Soc. Psychol.* 1986, *51*, 515–530.
120. Erpenbeck, J.; Sauter, W. *Stoppt die Kompetenzkatastrophe! Wege in eine neue Bildungswelt [Stop the Competence Disaster! Paths to A New World of Education]*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2019; ISBN 978-3662596777.
121. Wild, E.; Esdar, W. Beiträge der Pädagogischen Psychologie zur Wissenschaftspolitik [Contributions of educational psychology to science policy]. In *Handbuch Wissenschaftspolitik [Handbook Science Policy]*, 2nd ed.; Simon, D., Knie, A., Hornbostel, S., Zimmermann, K., Eds.; Springer: Wiesbaden, Germany, 2016; pp. 191–205.
122. Berufsethische Richtlinien des Berufsverbandes Deutscher Psychologinnen und Psychologen e. V. und der Deutschen Gesellschaft für Psychologie [Professional Ethical Guidelines of the Professional Association of German Psychologists e. V. and the German Psychological Society]. Available online: <https://www.bdp-verband.de/binaries/content/assets/beruf/ber-foederation-2016.pdf> (accessed on 21 August 2021).
123. World Medical Association. WMA's Declaration of Helsinki Serves as Guide to Physicians. *J. Am. Med. Assoc.* 1964, *189*, 33–34.
124. World Medical Association. Declaration of Helsinki. Ethical Principles for Medical Research Involving Human Subjects. *J. Am. Med. Assoc.* 2013, *310*, 2191–2194.
125. European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Off. J. Eur. Union* 2016, *59*, 294.

6 Diskussion

Basierend auf den Erkenntnissen zu den Problemen Lernender im Umgang mit naturwissenschaftlichen Theorien und Konzepten sowie im Verständnis von komplexen Themen (z. B. Jacobson & Wilensky, 2006; Yoon, 2008; s. Kapitel 1), wurden im Rahmen des hier vorgestellten Projektes zwei Studien durchgeführt, in denen die komplexen Inhalte der Zellbiologie exemplarisch als Lerninhalte dienten. Dafür wurden die Probleme der Lernenden unter Verwendung kognitionspsychologischer Theorien zum Wissenserwerb und zur Wissensorganisation analysiert und mögliche Ursachen vorgestellt (s. Kapitel 2.1). Im Speziellen wurden Defizite der Lernenden im vernetzten Denken identifiziert. Das vernetzte Denken erfordert kognitive Prozesse der Organisation und Elaboration und ermöglicht Lernenden, Zusammenhänge zwischen Konzepten zu erkennen. Diese Prozesse sind ebenfalls für das bedeutungsvolle Lernen wichtig, das die Verknüpfung neuer Informationen mit dem Vorwissen (Elaboration) voraussetzt (Ausubel, 2000). Durch die Integration in ihre Vorwissensstrukturen soll das Verständnis der Konzepte von Lernenden gefördert und das neue Wissen auf andere Kontexte übertragen werden können (Novak, 2003, Novak & Cañas, 2008). Als ein Grund für die genannten Defizite wurde ein inadäquater Einsatz von Diagnoseinstrumenten und Lernstrategien, die ein vernetztes Denken und bedeutungsvolles Lernen fordern und fördern können, vorgestellt. In diesem Zusammenhang wurde festgestellt, dass an Schulen und Universitäten häufig Diagnoseinstrumente zur Erfassung der Lernleistung eingesetzt werden, die rein auswendig gelerntes und somit unverknüpftes Wissen erheben (z. B. DiSessa, 2001; Novak, 2003). Der Einsatz solcher Instrumente hat dazu geführt, dass Lernende ihren Lernstrategieinsatz dahingehend angepasst haben: Statt Tiefenstrategien zu verwenden, welche kognitive Prozesse wie z. B. Organisation und Elaboration anregen, nutzen Lernende oberflächliche Strategien, die auf eine Speicherung isolierter Informationen abzielen. Da sie auf diese Weise ausreichende Erfolge erzielen, ziehen sie das Auswendiglernen dem aufwändigeren bedeutungsvollen Lernen vor (Mayer, 2002b; Montfort et al., 2009; Novak, 2003).

Die Anpassung seitens der Lernenden und die Tendenz zum Auswendiglernen wird wiederum als Grund dafür gesehen, dass Lernende Defizite im vernetzten Denken entwickeln, da ihnen das Auswendiglernen keine entsprechenden kognitiven Prozesse der Organisation und Elaboration abverlangt (s. Kapitel 2.1.3). Diese Prozesse können durch den Einsatz entsprechender kognitiver Lernstrategien induziert und gefördert werden (z. B. Mandl & Friedrich, 2006). Allerdings wird bezüglich des Einsatzes dieser Lernstrategien berichtet, dass sie nur unzureichend eingesetzt werden. Ferner wurde bezogen auf den metakognitiven

Lernstrategiegebrauch festgestellt, dass metakognitive Prozesse, wie das Überwachen und Evaluieren ihres Lernfortschritts den Lernenden schwerfallen (z. B. Bonner & Holliday, 2006; Isaacson & Fujita, 2006). Der Einsatz metakognitiver Lernstrategien vermag den Einsatz kognitiver Lernstrategien zu unterstützen und Lernleistungen zu steigern, doch Probleme diesbezüglich können z. B. dazu führen, dass Lernende ihren Lernfortschritt falsch einschätzen und durch ein weniger effektives Zeitmanagement ihr Lernpotenzial nicht voll ausschöpfen (z. B. Nelson & Dunlosky, 1991; Jersakova et al., 2017; s. Kapitel 2.1.4 und 2.2.1.3). Der unzureichende Einsatz der kognitiven und metakognitiven Lernstrategien konnte auf eine mangelnde Vertrautheit mit den Strategien zurückgeführt werden, welche zu Lerneinbußen führt. Der Vertrautheitsmangel kann wiederum Folge eines unzureichenden Trainings in der Ausübung der Strategie sein, da durch den ungeübten Strategieeinsatz eine zusätzliche kognitive Belastung entsteht. Ein ausreichendes Strategietraining, das diese Belastung reduziert und Lernenden hilft, die nötige Vertrautheit mit der Strategie zu erlangen, um sie souverän und lernwirksam einzusetzen, wird daher als entscheidender Faktor für die Effektivität ihres Einsatzes gesehen (z. B. Cañas et al., 2003; Schwendimann, 2015; s. Kapitel 2.3).

Als potenzielle Lösung für den Einsatz von Lern- und Diagnoseinstrumenten wurde anschließend das in der Lernpsychologie und Didaktik diskutierte Concept Mapping vorgestellt. Concept Mapping gilt als Tiefenstrategie, da es kognitive Prozesse der Organisation und Elaboration induziert, und als metakognitives Werkzeug, weshalb es sowohl an Schulen als auch an Universitäten eingesetzt wird (Brüning & Saum, 2017; Nesbit & Adesope, 2006; Novak, 2010; Novak & Cañas, 2008; Schroeder et al., 2018). In diesem Zusammenhang kann Concept Mapping genutzt werden, um bedeutungsvolles Lernen und den Wissenserwerb sowohl zu fördern als auch zu erfassen (Kinchin, 2011; Renkl & Nückles, 2006; Schroeder et al., 2018; s. Kapitel 2.2). Komplementär zu den Erkenntnissen bezüglich der Bedeutung eines Trainings für einen effektiven Einsatz von Lernstrategien (s. Kapitel 2.1.5) wird auch für den Einsatz des Concept Mappings von vielen Forscher:innen (z. B. Cañas et al., 2003; Mintzes et al., 2011) ein Training empfohlen. Allerdings herrscht Uneinigkeit darüber, wie Concept Mapping trainiert werden sollte, da entscheidende Parameter wie Dauer und Intensität der Trainingsansätze in den Studien stark variieren (Nesbit & Adesope, 2006; Schroeder et al., 2018; s. Kapitel 2.3.1). Des Weiteren wird in der Forschung diskutiert, wie Concept Mapping am lernwirksamsten eingesetzt werden sollte: Während viele Forscher:innen Lernende mit dem verfügbaren Lernmaterial in einem Elaborationssetting Concept Maps erstellen lassen, um z. B. Textverstehensprozesse zu induzieren (Nesbit & Adesope, 2006; Schroeder et al., 2018), setzen andere Concept Mapping in einem Retrievalsetting ein, in dem das Lernmaterial beim Concept

Mapping nicht verfügbar ist, um lernwirksame Abrufprozesse zu fördern (z. B. Blunt & Karpicke, 2014; O'Day & Karpicke, 2020; s. Kapitel 2.3.2). Aufgrund uneinheitlicher Befunde zu den Lernleistungen und fehlender Vergleichsstudien konnte bisher nicht geklärt werden, welches Setting lernwirksamer ist. Folglich existieren sowohl zu Trainings- als auch Einsatzmöglichkeiten des Concept Mappings unterschiedliche Empfehlungen und die Berichte lassen auf Leistungsunterschiede schließen, die auf die unterschiedlichen Lernbedingungen zurückzuführen sein könnten.

Das Ziel der vorliegenden Dissertation war es, bezüglich der Bedeutung des Trainings und der Einsatzmöglichkeiten von Concept Mapping für kognitive Prozesse, kognitive Belastung und die Lernleistung Folgendes herauszufinden: Erstens, ob Lernende, die ein spezifisches Concept Mapping-Training erhalten haben, generell erfolgreicher im Lernen mit Concept Mapping sind als solche, die auf andere Weise an Concept Mapping herangeführt wurden. Darauf aufbauend sollte zweitens herausgefunden werden, wie lerneffektiv der Einsatz des Concept Mappings ist, wenn den Personen der Lerntext nicht zur Verfügung steht, während sie ihre Concept Maps erstellen (Retrievalsetting), im Gegensatz zum klassischen Elaborationssetting, bei dem das Lernmaterial verfügbar ist. Drittens sollte herausgefunden werden, ob mögliche Effekte der beiden Settings von der Intensität eines vorangegangenen Methodentrainings im Concept Mapping abhängen (vgl. Kapitel 3). Die Anpassung des Einsatzes von Concept Mapping in der Praxis an die Ergebnisse hätte nicht nur Konsequenzen für den Einsatz von Concept Mapping im Rahmen der Vermittlung zellbiologischer Inhalte, sondern auch für andere Wissenschaftsbereiche, deren Lerninhalte ein vernetztes Denken erfordern und in denen sich Concept Mapping als interdisziplinäre Lernmethode eignet.

Anhand der Ergebnisse der Vorbefragungen zu Beginn der Trainingsphase konnten in beiden Studien vergleichbare Grundvoraussetzungen der Proband:innen angenommen werden bezüglich ihres Alters und Abiturdurchschnitts, ihres Vorwissens über Zellbiologie, ihren Vorerfahrungen mit Concept Mapping und ihrer Lesekompetenz (s. Kapitel 4, 3.1 und 5, 3.1). Des Weiteren zeigte sich in der ersten Studie, dass die unterschiedlichen Trainingsbedingungen in der dreiwöchigen Trainingsphase keine Auswirkungen auf den Wissenserwerb zum Thema *Intelligenz* hatten, da die Gruppen keine signifikanten Unterschiede im dazugehörigen Test aufwiesen. Von einer zusätzlichen Belastung, die sich in Lernnachteilen der Concept Mapping-Gruppen manifestiert hätte, kann demnach nicht ausgegangen werden (s. Kapitel 4, 3.2 und 4, 4.2).

Im Folgenden werden die zentralen Ergebnisse der Lern- und Testphasen sowie Diskussionspunkte aus beiden Studien hypothesengeleitet aufgegriffen und zusammenfassend diskutiert. Um die Bedeutung der Faktoren *Training* und *Setting* für die kognitiven Prozesse, kognitive Belastung und Lernleistung vorzustellen, wird zuerst eine Aufteilung der beiden vorgenommen: Zu Beginn wird die Bedeutung des Trainings für die abhängigen Variablen diskutiert, da das Training in beiden Studien eine unabhängige Variable darstellte. Anschließend wird die Bedeutung unterschiedlicher Einsatzmöglichkeiten des Concept Mappings unter dem Aspekt der Textverfügbarkeit diskutiert, die als unabhängige Variable in der zweiten Studie diente. Anhand der gewonnenen Erkenntnisse werden im jeweiligen Abschnitt Empfehlungen für die Praxis abgeleitet. Abschließend werden die Limitationen der beiden Studien sowie Bereiche genannt, in denen weitere Erkenntnisse in Zukunft gewonnen werden könnten, um den Einsatz von Concept Mapping in der Praxis und die Lernwirksamkeit der Methode zu optimieren.

6.1 Concept Mapping-Trainingsbezogene Aspekte

Die Hypothesen H1.1 bis H1.3 sowie H2.1 und H2.2 gehen davon aus, dass ein extensives Concept Mapping-Training zu einem effektiveren Lernen mit Concept Mapping führt, und zwar in Bezug auf eine bessere organisatorische und elaborative Informationsverarbeitung, eine geringere wahrgenommene kognitive Belastung und eine höhere Gesamtleistung beim Lernen (s. auch Kapitel 6, 1.4). Diese Annahmen werden durch die Ergebnisse der beiden Studien teilweise gestützt (s. Kapitel 4 und 5) und im Folgenden hypothesengeleitet, aber studienübergreifend diskutiert.

6.1.1 Kognitive Prozesse

Die kognitiven Prozesse der Proband:innen wurden mithilfe der Propositionen in den Concept Maps der Lernphase in die drei Kategorien *Abruf-*, *Organisation-* und *Elaboration-suggestierend* geteilt (vgl. Kapitel 4, 2.4.4 und 5, 2.3). Die Ergebnisse bezüglich der ersten Hypothese aus Studie 1 (H1.1) zeigen, dass es numerische Unterschiede zugunsten der Concept Mapping-Trainingsgruppen in der mittleren Anzahl erstellter Elaborationspropositionen in den Concept Maps während der Lernphase gab, sich die drei Gruppen aber weder bezüglich Organisations- noch Elaborationspropositionen signifikant unterscheiden. Auch in der zweiten Studie zeigt ein Vergleich der beiden Trainingsgruppen, denen das Lernmaterial während des Concept Mappings zur Verfügung stand (T+E und T-E), dass die Concept Mapping-Trainingsgruppe (T+E) im Durchschnitt einen circa doppelt so hohen Elaborations-Anteil in ihren Propositionen hatte wie die Kontrolltrainingsgruppe; diese Unterschiede waren jedoch statistisch nicht signifikant (s. Kapitel 5, 3.5.3). Eine größere Vertrautheit mit der Lernmethode,

die durch eine Teilnahme am extensiven Training angenommen werden kann, hat demnach scheinbar keinen Einfluss auf die in den Concept Maps dargestellten, kognitiven Prozesse. Ein Grund dafür könnte sein, dass die Studienteilnehmer:innen wahrscheinlich alle aufgrund ihres Studiums viel Lernerfahrung besaßen und sich daher im Ausmaß der gezeigten kognitiven Prozesse nicht signifikant voneinander unterschieden. Weitere Gründe für mangelnde Unterschiede könnten zum einen die fehlende explizite Aufforderung zur Elaboration in den Aufgaben gewesen sein, zum anderen kann davon ausgegangen werden, dass der Aufwand, welcher durch die Übertragung der Textinformationen in Propositionen entstand, entsprechende kognitive Ressourcen in Anspruch nahm (vgl. Kapitel 4, 4.4). Dieser Aufwand wird sowohl in folgenden beiden Abschnitten als auch in Kapitel 7.2 erneut aufgegriffen, da er auch hier einen möglichen Einflussfaktor darstellt.

6.1.2 Kognitive Belastung

Bezüglich der intrinsischen und lernbezogenen Belastung (ICL und GCL) ergaben sich entgegen der Hypothese (H1.2) keine statistisch signifikanten Unterschiede zwischen den Gruppen in der ersten Studie. Dies lässt darauf schließen, dass die Gruppen die Concept Mapping-Aufgabe als ähnlich komplex (ICL) ansahen, gleichsam aktiv während der Bearbeitung der Aufgabe waren und ihre kognitiven Ressourcen in einem sehr ähnlichen Ausmaß auf die Lernprozesse richteten (GCL). Allerdings empfand die T- Gruppe in der ersten Studie das Lernmaterial als statistisch signifikant anstrengender (ECL) als die T+ Gruppe. Eine extrinsische kognitive Belastung wird jedoch i. d. R. durch ungeeignetes Lernmaterial oder andere externe Faktoren wie das Lernumfeld verursacht (Klepsch et al., 2017). Da alle Gruppen das gleiche Lernmaterial im gleichen Lernumfeld erhielten, sind die gemessenen Unterschiede möglicherweise auf die Reliabilität des Fragebogens zurückzuführen, die für die Skalen des ECL und ICL in der ersten Studie fragwürdig war (s. Kapitel 4, 3.5 und 4, 4.5). Ähnliche oder noch niedrigere Reliabilitätswerte wiesen die Skalen zur kognitiven Belastung in der zweiten Studie auf, weshalb sie weder berichtet noch interpretiert wurden (s. Kapitel 5, 3) und die komplementäre Hypothese (H2.2) nicht falsifiziert werden kann. Eine Entlastung des Arbeitsgedächtnisses, wie sie für das Concept Mapping diskutiert wird (Jonassen, 2003; Nesbit & Adesope, 2006; O'Donnell et al., 2002; Renkl & Nückles, 2006), ist somit nicht auszuschließen; allerdings sollte berücksichtigt werden, dass die Stichproben in beiden Studien aus Studierenden bestanden, die mehrjährige Lernerfahrung besaßen (s. Kapitel 7.3). Vor diesem Hintergrund deuten zumindest die Ergebnisse der ersten Studie bezüglich der kognitiven Belastung darauf hin, dass die Kurzeinführung wahrscheinlich ausreichend war, um

einen differenziellen Einfluss eines extensiven Trainings zu verhindern (s. Kapitel 4, 3.5 und 4, 4.5).

6.1.3 Metakognitive Vorhersage

Die Ergebnisse der ersten Studie zu den metakognitiven Vorhersagen bezüglich ihres erwarteten Lernerfolgs (JOLs) zeigen, dass Studienteilnehmer:innen, die ein extensives Training erhalten hatten, höhere Lernleistungen erwarteten als die Kontrolltrainingsgruppe; in der zweiten Studie gaben die Concept Mapping-Trainingsgruppen jedoch wider Erwarten niedrigere JOLs als die Kontrolltrainingsgruppen an. Die Unterschiede waren allerdings in beiden Studien statistisch nicht signifikant (s. Kapitel 4, 3.6 und 5, 3.5). Wenn jedoch die JOLs mit den Ergebnissen des deklarativen Wissenstests verglichen werden, fällt auf, dass die Concept Mapping-Trainingsgruppen in der ersten Studie ihre Lernleistung akkurater einschätzen konnten als die Kontrolltrainingsgruppe. Auch ein Vergleich der beiden komplementären Gruppen aus der zweiten Studie, die den Text während des Concept Mappings verfügbar hatten, zeigt, dass die CM-Trainingsgruppe (T+E) ihre Lernleistung exakter einschätzte als die Kontrolltrainingsgruppe (T-E). Daraus lässt sich ableiten, dass das Concept Mapping-Training den Lernenden auf metakognitiver Ebene hilft, ihr Lernen zu reflektieren und ihre Lernleistung besser einzuschätzen, wenn das Lernmaterial beim Concept Mapping zur Verfügung steht. Dies wiederum unterstreicht die Bedeutung eines umfassenden Trainings für metakognitive Evaluationsprozesse beim Lernen mit Concept Mapping in Elaborationssettings und spricht für einen metakognitiven Ansatz, wie ihn Hilbert und Renkl (2008, 2009; s. Kapitel 2.3.1) vorschlagen.

Die metakognitiven Vorhersagen lieferten zudem eine weitere wichtige Erkenntnis: Im Durchschnitt schätzten die Proband:innen in beiden Studien unabhängig von ihrer Gruppenzugehörigkeit, dass sie sich nur an etwa zwischen 40 und 60 % des Inhalts erinnern würden. Das bedeutet, dass die hier untersuchten Studierenden, die sowohl eine mehrjährige Lernerfahrung als auch schulbiologisches Vorwissen besaßen, in fast zwei Stunden Lernzeit nur etwa die Hälfte der zu lernenden Inhalte glaubten behalten zu können. Gründe für diese Urteile der Lernenden liefern einerseits die in JOLs berücksichtigte Verständlichkeit und andererseits die Schwierigkeit der Lerninhalte: Zaromb et al. (2010) konnten zeigen, dass Lernende höhere JOLs vergeben, wenn sie der Meinung sind, dass die Informationen relativ leicht zu verstehen waren, während Koriat (1997) herausfand, dass Lerninhalte, die als schwierig zu lernen eingeschätzt werden, niedrigere JOLs erhalten. Dies deutet darauf hin, dass die Teilnehmer:innen beider Studien den Text entweder als ziemlich schwer verständlich empfanden oder den Text für verständlich hielten, die Informationsmenge aber so groß war,

dass sie sich voraussichtlich nur die Hälfte davon merken konnten. Von diesen beiden Möglichkeiten ist letztere die wahrscheinlichere, da der Lerntext zur Zellbiologie zuvor von Expert:innen als für diese Zielgruppe verständlich, strukturiert und adäquat befunden wurde (vgl. Kapitel 4, 2.5 und 5, 2). Somit kann davon ausgegangen werden, dass sich die JOLs der Proband:innen primär auf die Wortanzahl und die dadurch bedingte Informationsmenge bezogen (vgl. Kapitel 5, 4.3).

Diese Überlegung wird zusätzlich gestützt, wenn die Korrelationswerte der Lernleistungsergebnisse mit den JOLs aus den jeweiligen Studien betrachtet werden: Hier zeigte sich, dass die JOLs der Proband:innen offensichtlich viel stärker auf deklaratives und strukturelles als auf konzeptuelles Wissen bezogen waren. Diese Unabhängigkeit von konzeptuellem Wissen zeigte sich auch im direkten Vergleich zu den drei anderen Lernleistungsmaßen in den Ergebnissen (s. Kapitel 4, Tab. 6; 5, Tab. 5). Auch wenn der Lernstrategiegebrauch der Proband:innen nicht erhoben wurde, zeigen die Ergebnisse der beiden Studien, dass die hier untersuchten Personen ihre JOLs eher auf den Erwerb und die Reproduktion von deklarativem und manchmal auch strukturellem Wissen beziehen. Dies wiederum deutet darauf hin, dass sie in ihrem universitären Alltag kaum dazu angehalten sind, vernetzt zu denken und ihr Wissen auf neue Kontexte zu übertragen. Das deklarierte Bildungsziel, bei Lernenden ein tieferes Verständnis von Zusammenhängen und Abhängigkeiten zwischen zentralen Konzepten der jeweiligen Fachgebiete für Transferleistungen zu fördern (Dobson et al., 2019; Mayer, 2002a), scheint demnach für den erfolgreichen Abschluss der meisten Studiengänge weniger erforderlich zu sein (vgl. Kapitel 5, 4.3).

Bei einer mangelnden Einforderung solchen Wissenserwerbs durch entsprechende Diagnoseinstrumente (Novak, 2003; Watson et al., 2016) ist es jedoch nicht verwunderlich, dass die metakognitiven Vorhersagen der Lernenden bezüglich ihrer erwarteten Lernleistung nur ein Minimum dessen abdecken, was sie beherrschen sollten, wie z. B. die Kenntnis von Fakten, ihrer impliziten hierarchischen Struktur und den Beziehungen zwischen Begriffen (KMK, 2004, 2020). Dieses Anforderungsspektrum wird in der Praxis nicht zwangsläufig durch jene Diagnoseinstrumente, wie z. B. Multiple-Choice-Tests, abgedeckt, die am meisten an Universitäten eingesetzt werden (DiSessa, 2001; Hasnain & Bhamani, 2014; Hasnor et al., 2013; Novak, 2003). Das Spektrum soll jedoch in direktem Zusammenhang mit Concept Mapping stehen und könnte durch eine Integration dessen in den Schul- und Universitätsalltag als Diagnoseinstrument gefördert werden (z. B. Hilbert, Nückles, Renkl, et al., 2008; Nesbit & Adesope, 2006; Renkl & Nückles, 2006; Sumfleth et al., 2010; Weinstein & Mayer, 1986).

Das Potential von Concept Mapping wird in diesem Zusammenhang ebenfalls durch die hier berichteten Ergebnisse gestützt, die zeigen, dass die JOLs signifikant mit drei Operationalisierungen der Lernleistung korrelierten, und zwar der Concept Map-Qualität (aQCM-Index), dem deklarativen Wissen und dem strukturellen Wissen (s. Kapitel 4, Tab. 6 und 5, Tab. 5). Dies unterstützt das Modell einer isomorphen Analogie zwischen dem propositionalen Netzwerkmodell des Gedächtnisses (Collins & Quillian, 1969; Quillian, 1968) und Concept Maps als Externalisierung kognitiver Repräsentationen (Cox, 1999). Zusätzlich wird der Zusammenhang des konzeptuellen Wissens und den Concept Mapping-Fertigkeiten durch Ergebnisse der ersten Studie gestützt, die zeigen, dass die Entwicklung von konzeptuellem Wissen durch das Concept Mapping-Training positiv beeinflusst wurde (s. Kapitel 4, Tab. 7) und das Training in der zweiten Studie zu keinen Lernnachteilen führte (s. Kapitel 5, 4). Concept Mapping kann folglich effektiv für das Lernen sein und auch den Erwerb von vertieftem, flexibel übertragbarem Wissen im Sinne des deklarierten Bildungsziels fördern, was für eine stärkere Einbettung des Concept Mappings in Schulen und Universitäten als Lerninstrument spricht, wie sie von diversen Forscher:innen (z. B. Cañas et al., 2003; Kinchin, 2001; Mintzes et al., 2011; Novak & Cañas, 2006, 2008) befürwortet wird.

6.1.4 Lernleistung

Aufgrund uneinheitlicher Befunde zum Einfluss von Concept Mapping-Trainings auf die Concept Map-Qualität und den Wissenserwerb (s. Kapitel 2.2; 2.3) wurde die Lernleistung der Proband:innen einerseits mithilfe von Concept Mapping-bezogenen Variablen und andererseits durch Wissenstests zum deklarativen, strukturellen und konzeptuellen Wissen erhoben (s. Kapitel 4, 2.4.5 und 5, 2.4.3).

- **Concept Mapping-bezogene Leistung**

Da aufgrund bisheriger Befunde (z. B. Sumfleth et al., 2010) von einem Einfluss der Vertrautheit mit Concept Mapping auf die Concept Mapping-Fertigkeiten ausgegangen werden konnte, wurde erwartet, dass Studienteilnehmer:innen, die ein extensives Training erhalten hatten, eine höhere Selbstwirksamkeit berichten würden als solche, die lediglich eine Kurzeinführung bekamen (s. Kapitel 4, 2.4.3). In beiden Studien berichteten jedoch die Teilnehmer:innen der unterschiedlichen Trainingsgruppen keine signifikanten Unterschiede in ihrer Selbstwirksamkeit bezüglich Concept Mapping (s. Kapitel 4, Tab. 3 und 5, Tab. 4).

Nichtsdestotrotz kann aufgrund der Auswertungsergebnisse zu den anderen Concept Mapping-bezogenen Lernleistungsmaßen davon ausgegangen werden, dass insbesondere die Teilnehmer:innen der Kontrolltrainingsgruppen ihre Concept Mapping-Fertigkeiten inkorrekt einschätzten und dass das Training einen Effekt auf diese Fertigkeiten hatte. Gründe für diese

Annahmen liefern erstens die Ergebnisse der vorgegebenen Concept Maps, die zu korrigierende Fehler enthielten und zweitens die von den Lernenden selbst erstellten Concept Maps. Im Rahmen der ersten Studie (s. Kapitel 4) wurde gezeigt, dass ein extensives Concept Mapping-Training mit Scaffolding- und Feedbackelementen Lernenden am besten dabei hilft, eine Vertrautheit mit der Lernmethode zu erlangen und sie gewinnbringend über die Lernphase hinaus einzusetzen. Die vorteilhafte Vertrautheit zeigte sich auch in Studie 2, zum einen in statistisch signifikant besseren Ergebnissen im Vergleich zur Kontrolltrainingsgruppe bei der Aufgabe, Fehler in einer vorgegebenen Concept Map zu korrigieren; zum anderen machte die genannte Concept Mapping-Trainingsgruppe in beiden Studien signifikant weniger methodische Fehler in ihren selbsterstellten Concept Maps während der Lern- und der Testphase.

Die Vorteile der beiden Concept Mapping-Trainingsgruppen (Studie 1) zeigten sich auch in qualitativ besseren Concept Maps, wobei die Ergebnisse Hinweise darauf geben, dass die Gruppe mit Scaffolding und Feedback stärker ausgeprägte Concept Mapping-Fertigkeiten auch im Vergleich zur Gruppe erwarb, die das Concept Mapping ohne Lernhilfen wiederholt übte (s. Kapitel 4, 3.7). In der zweiten Studie konnten die Befunde zur Wirksamkeit der Trainingsmaßnahmen bezüglich Concept Mapping-Fertigkeiten bestätigt werden: Die Ergebnisse zeigen, dass Teilnehmer:innen, die das extensive Concept Mapping-Training erhalten hatten, sowohl in der Lern- als auch in der Testphase statistisch signifikant niedrigere Concept Mapping-Fehlerquoten in ihren Concept Maps hatten als diejenigen, die am Kontrolltraining teilgenommen und lediglich eine Kurzeinführung erhalten hatten.

Diese Befunde sprechen einerseits dafür, dass die Lernmethode für einen souveränen Einsatz entsprechend ausführlich trainiert werden muss und stützen bisherige Befunde zum positiven Einfluss extensiver Concept Mapping-Trainingsmaßnahmen auf die Concept Map-Qualität (z. B. Jegede et al., 1990; Mintzes et al., 2011; Nesbit & Adesope, 2006; Schroeder et al., 2018; Sumfleth et al., 2010). Andererseits sprechen die Ergebnisse gegen die Verwendung des ökonomischeren extensiven Trainings ohne Lernhilfen (s. Kapitel 2.3.1) und bestätigen die Ergebnisse anderer Studien zur Wirksamkeit von Scaffolds (z. B. *skeleton maps*; Chang et al., 2002; s. auch van de Pol et al., 2010) sowie metakognitiver Prompts (z. B. Pressley et al., 1992; Stark et al., 2008) und Feedback (z. B. Roessger et al., 2018; s. auch Hattie, 2008).

- Erwerb deklarativen, strukturellen und konzeptuellen Wissens

Die Unterschiede in der Vertrautheit mit Concept Mapping und einer daraus resultierenden Concept Map Qualität spiegeln in Studie 1 teilweise auch die Leistungen in den anderen Tests wider und ermöglichen eine Verifizierung der entsprechenden Hypothese (H1.3): Der im Vergleich zur Kontrollgruppe lernwirksamere Einsatz des Concept Mappings durch Lernende, die vorher ein Training mit Scaffolding und Feedback oder ein Training mit wiederholtem Üben erhalten hatten, zeigte sich in statistisch signifikant besseren Ergebnissen in den Tests zum strukturellen und konzeptuellen Wissen. Auch bezüglich des Faktenwissens erzielten die Concept Mapping-Trainingsgruppen bessere Ergebnisse in der Testphase, allerdings stellten sich diese Unterschiede als nicht signifikant heraus, ließen jedoch, im Vergleich zur Kontrollgruppe, eine Tendenz zugunsten der T+ Gruppe erkennen (s. Kapitel 4, Tab. 7).

Die Lernleistung in den Domänen des strukturellen und konzeptuellen Wissens wurde entgegen der Hypothese (H2.3) jedoch in Studie 2 nicht durch die Art des Trainings beeinflusst. Ferner konnten die statistischen Trends in den Ergebnissen aus Studie 1 zum Faktenwissen zugunsten der T+ Gruppe in Studie 2 nicht nachgewiesen werden. Die Gruppen, die ein Concept Mapping-Training erhalten hatten, das dem der T++ Gruppe aus der Studie 1 gleich, erzielten keine signifikant besseren Ergebnisse in den Tests zu den drei Wissensdomänen als die Kontrolltrainingsgruppen, die lediglich eine Kurzeinführung in das Concept Mapping zu Beginn der Lernphase erhalten hatten. Die Kurzeinführung reichte demnach für die Lernenden scheinbar aus, um einen vergleichbaren Wissensstand in den jeweiligen Wissensdomänen zu erlangen. Diese Ergebnisse stehen somit im Gegensatz zu denen von Studie 1, in der die Teilnehmer:innen, die ein extensives Training erhalten hatten, signifikant bessere Ergebnisse in den Tests zum strukturelles und konzeptionelles Wissen erzielten als diejenigen, die eine Kurzeinführung in das Concept Mapping erhalten hatten (s. Kapitel 5, 3.4).

Die besseren Leistungen der Concept Mapping-Trainingsgruppen in den Tests zum strukturellen und konzeptuellen Wissen sowie der Concept Mapping-bezogenen Fertigkeiten in Studie 1 spiegeln Ergebnisse aus Berichten anderer Forscher:innen (z. B. Jegede et al., 1990; Kinchin, 2000a, 2000c; Nesbit & Adesope, 2006; Schroeder et al., 2018) wider, die ebenfalls einen positiven Einfluss von extensiven Trainingsmaßnahmen auf Lernleistungen fanden. Des Weiteren legen diese spezifischen Ergebnisse nahe, dass die hier verwendete Kurzeinführung in das Concept Mapping, die einen vergleichbaren Zeitrahmen hatte wie Einführungen in anderen Studien (z. B. Hay et al., 2008; Hilbert & Renkl, 2009), nicht ausreicht, um das Maß an Vertrautheit mit der Methode zu ermöglichen, das von Lernenden benötigt wird, um Concept

Mapping lernwirksam einzusetzen. Die Ergebnisse solcher Studien und potenzielle Nachteile durch den Einsatz von Concept Mapping wären daher kritisch zu sehen, da die Lernenden vermutlich nicht in der Lage waren, die Lernaufgabe adäquat zu lösen. Diese Annahmen sind jedoch aufgrund der Ergebnisse der zweiten Studie nicht haltbar: Hier deutet der Mangel an Unterschieden bezüglich des deklarativen, strukturellen und konzeptuellen Wissens auf die ausreichende Effektivität von Kurzeinführungen. Die Kritik von Forscher:innen wie Mintzes et al. (2011) an Kurzeinführungen zum Concept Mapping wäre mittels der vorliegenden Daten nicht stützbar und würde für eine ausreichende Vertrautheit der Studienteilnehmer:innen bei Karpicke und Blunt (2011), Blunt und Karpicke (2014) sowie O'Day und Karpicke sprechen. Dies würde für die Nutzung derartiger Kurzeinführungen zur ausreichenden Lernförderung durch Concept Mapping sprechen. Ein solcher Einsatz wäre jedoch ausschließlich dann vertretbar, wenn die Concept Mapping-bezogenen Fertigkeiten der Lernenden keine Rolle spielen und von einer Betrachtung der Qualität der Concept Maps abgesehen wird.

Von einem solchen Vorgehen wäre aus folgenden Gründen abzuraten: Zum einen hat sich gezeigt, dass ein Mangel an Vertrautheit in Unsicherheit der Lernenden resultieren kann (Allen & Tanner, 2003), die wiederum zu Angst führen und Einfluss auf die Lernleistung haben kann (Jegede et al., 1990). Zum anderen ist eine Grundvoraussetzung für die Interpretation von Ergebnissen aus Concept Maps die Vertrautheit der Lernenden mit der Strategie und ihren Fertigkeiten, methodisch korrekte Propositionen zu erstellen. Wenn eine Person eine fehlerhafte Proposition erstellt, kann nicht eindeutig festgestellt werden, ob der Fehler auf einem inhaltlichen Missverständnis oder der mangelnden Vertrautheit mit Concept Mapping basiert (vgl. Kapitel 2.2.1.2). Dies ist nicht nur für den Einsatz von Concept Mapping in der Zellbiologie, sondern auch in anderen naturwissenschaftlichen Bereichen wichtig, da inhaltliche Fehlvorstellungen in Naturwissenschaften generell weit verbreitet sind und sie als solche korrekt identifiziert werden müssen (Driver, 1989; Ruiz-Gallardo & Reavey, 2019). Durch einen Trainingsmangel verursachte Unklarheiten verringern daher die Reliabilität und Validität von Concept Maps als Diagnoseinstrument, da die kognitiven Strukturen der Lernenden nicht adäquat repräsentiert werden konnten (s. auch Brandstädter et al., 2012; Ruiz-Primo, Schultz, et al., 2001). Bei der Auswertung der Concept Maps im Rahmen dieses Projektes hat sich zudem gezeigt, dass eine Auswertung von qualitativ hochwertigeren Maps weniger Zeitaufwand bedeutet, da klar formulierte und korrekt erstellte Propositionen in kürzerer Zeit analysiert und interpretiert werden können als solche, die diese Kriterien nicht erfüllen. Die nachträgliche Klärung von Unklarheiten in Concept Maps kann durch weitere schriftliche Tests (wie in den hier vorgestellten Studien) und/oder Interviews mit den Personen

stattfinden. Die Durchführung und Auswertung der beiden jeweiligen Ansätze bedeutet jedoch einen erheblichen Zeitaufwand (z. B. Bortz & Döring, 2016; Egger-Rainer, 2019) und je nachdem, wie viele Concept Maps der Klärung bedürfen, könnte eine vorgeschaltete extensive Einführung in das Concept Mapping mit insgesamt weniger Zeitaufwand verbunden sein als eine nachträgliche Analyse und Klärung von Unklarheiten.

Zur Steigerung der Validität von Concept Map-Ergebnissen empfehlen Brandstädter et al. (2012) eine Reduzierung der Freiheitsgrade durch ein hochgradig gerichtetes Mapping, z. B. durch die Begrenzung der verfügbaren Pfeile oder die Vorgabe einer vorbereiteten Concept Map. Allerdings geht eine solche starke Strukturvorgabe immer mit einem Informationsverlust bezüglich der kognitiven Strukturen der Lernenden einher (Ruiz-Primo, Schultz, et al., 2001; s. Kapitel 2.2.1). Eine weniger starke Strukturierung wurde in den Testphasen der hier vorgestellten Studien mit einem vorgegebenen Begriffssatz implementiert, das die Teilnehmer:innen für die Erstellung ihrer Concept Maps nutzen sollten. Einige Propositionen in diesen Concept Maps deuten darauf hin, dass die Lernenden in der Lage waren, die Begriffe in ihre Concept Maps methodisch korrekt zu integrieren. Bezüglich der Relationen zwischen den Begriffen stellten sie jedoch teilweise Vermutungen an, so dass sie methodisch korrekte Propositionen erstellten, deren Inhalt jedoch falsch war. Diese Propositionen belegen zwar, dass die Lernenden sich nicht an die Inhalte der Lernphase erinnern konnten, jedoch wären ihre Wissenslücken durch zu starke Strukturierungsvorgaben möglicherweise verborgen geblieben. Im Rahmen von Lehrveranstaltungen können solche Fehler genutzt werden, um Missverständnisse zu klären und ihnen möglichst früh entgegenzuwirken. Dies spricht wiederum einerseits dafür, dass Lernende darin geschult werden sollten, ihre eigenen Concept Maps für das Lernen zu konstruieren (s. auch Hilbert, Nückles, & Matzel, 2008); andererseits zeigen die Ergebnisse der hier vorgestellten Studien, dass diese Schulung so lang und intensiv wie nötig sein muss, um den nützvollen Einsatz des Concept Mappings durch die Lernenden zu gewährleisten (s. Kapitel 4, 3 und 5, 3).

6.2 Einsatz von Concept Mapping in Elaborations- und Retrievalsettings

Auf Grundlage empirischer Befunde zur Lernwirksamkeit von Abrufprozessen (z. B. Blunt & Karpicke, 2014) und den Ergebnissen der ersten in dieser Arbeit vorgestellten Studie wurde erwartet, dass neben der Art des Concept Mapping-Trainings auch die Textverfügbarkeit während des Concept Mappings einen Einfluss auf kognitive Prozesse, die kognitive Belastung und die Lernleistung hat (H2.1; H2.3).

6.2.1 Kognitive Prozesse

Den Erwartungen entsprechend ergaben sich signifikante Unterschiede zwischen den Gruppen bezüglich der in den Concept Maps erstellten Propositionen. Entgegen der Erwartung (H2.1) wurden jedoch die Elaborationsgruppen ihrem Namen nicht gerecht und wiesen einen signifikant geringeren Anteil an Elaborationspropositionen in ihren Concept Maps auf als die Proband:innen der Retrievalgruppen, die nach dem Lesen des Textes ihre Concept Map ohne diesen erstellten. Die Abwesenheit des Lernmaterials scheint demnach nicht dazu zu führen, dass Lernende ihre kognitiven Ressourcen vermehrt für den Abruf zu Lasten der Elaboration einsetzen müssen. Stattdessen stellen Lernende in einem Retrievalsetting zusätzliche Verbindungen zwischen Konzepten aus dem Text her und beziehen ihr Vorwissen stärker in die Concept Map ein, was für eine aktive Wissenskonsolidierung spricht. Elaborationsprozesse dagegen werden erwartungswidrig nicht durch das verfügbare Lernmaterial gefördert (s. Kapitel 5, 3.5.3).

Das höhere Maß an Elaborationspropositionen der Retrievalgruppen ist vermutlich darauf zurückzuführen, dass ihre Aufmerksamkeit nicht durch die Arbeit mit dem Text gebunden wurde. Dadurch griffen sie mehr auf ihr Vorwissen zurück und integrierten es in ihre Concept Maps. Im Gegensatz dazu konzentrieren sich die Elaborationsgruppen mehr darauf, die Informationen des Textes in Propositionen zu transferieren. Im Sinne der dualen Kodierungstheorie von Paivio (1986) ist dies einer der Vorteile des Concept Mappings, da die Übersetzung unterschiedliche Gedächtnissysteme aktiviert (s. auch Nesbit & Adesope, 2006). Bei einer begrenzten Lernzeit führt dieses Verhalten der Elaborationsgruppen jedoch dazu, dass sie dazu tendieren, die Informationen des Textes in ihren Concept Maps abzubilden, anstatt sie im Sinne der Elaboration mit ihren Vorwissensstrukturen zu verbinden. Die Übersetzung von Textinhalten aus dem Gedächtnis in einem Retrievalsetting führt somit zu keinen Nachteilen – im Gegenteil: Die enge Arbeit mit dem Text im Elaborationssetting könnte sich als Nachteil erweisen, da die Elaboration generell mit besseren Lernleistungen in Verbindung gebracht wird, da dieses Wissen flexibel angewandt werden kann (J. R. Anderson, 1983; Mayer, 2002a; Weinstein & Mayer, 1986).

Allerdings deuten die Ergebnisse zum Erwerb deklarativen, strukturellen und konzeptuellen Wissens nicht auf einen Nachteil der Elaborationsgruppen durch weniger Elaborationspropositionen hin (vgl. Kapitel 5, 3 und 6.2.4). Ein Grund hierfür könnte das studienübergreifend auftretende niedrige Maß an derartigen Propositionen sein: Die Lernenden erstellten sowohl in der ersten als auch in der zweiten Studie nur sehr wenige Organisations- und Elaborationspropositionen. Daher könnte das Maß der Elaboration – trotz statistisch

signifikanter Unterschiede in den Concept Maps in Studie 2 – nicht ausreichend gewesen sein, um sich auch in statistisch signifikanten Unterschieden in den Lernleistungsmaßen zu manifestieren. Verantwortlich für das niedrige Maß an Organisations- und Elaborationspropositionen kann zum einen die angesprochene limitierte Zeit für die Erstellung einer Concept Map gewesen sein – insbesondere für die Elaborationsgruppen, da sie sowohl mit dem (recht umfangreichen) Text arbeiten als auch eine Concept Map erstellen mussten. Zum anderen kann die Formulierung der Aufgabenstellung (s. Kapitel 6.1), die ebenfalls nicht ausreichend gewesen sein könnte, um die Lernenden zu mehr Elaboration zu motivieren, als Grund angeführt werden.

Für die Praxis lässt sich daher ableiten, dass Lehrkräfte das Vorwissen der Lernenden sowohl in der Gestaltung der Lernmaterialien als auch in den Aufgabenstellungen berücksichtigen müssen (z. B. Hay et al., 2008; Jacobs, 2002; Krause et al., 2004; Nerdel, 2017). Die Manipulation der Textverfügbarkeit zeigte, dass ein Retrievalsetting zu mehr Elaboration führt, aber die wichtige interdisziplinäre Frage bleibt bestehen, wann und unter welchen Bedingungen bedeutungsvolles Lernen im Kontext von Organisation und Elaboration erreicht werden kann und welche Bedingungen für die Konstruktion von Concept Maps erfüllt sein müssen, um den Lernenden zu helfen, über die Reproduktion des Gelernten hinauszugehen (vgl. Kapitel 4, 4.4).

6.2.2 Kognitive Belastung

Gemäß den in den Hypothesen formulierten Erwartungen (H2.2) wurde untersucht, ob Lernende, die ein vorgeschaltetes Training im Concept Mapping erhalten und das Lernmaterial in der Lernphase verfügbar hatten, eine niedrigere kognitive Belastung berichten. Die Reliabilität der Skalen vom Instrument zur Erfassung der kognitiven Belastung (s. Kapitel 4, 2.4.4) hatte sich in der ersten Studie als akzeptabel erwiesen, in der zweiten Studie jedoch als inakzeptabel. Daher war keine Interpretation der entsprechenden Ergebnisse möglich. Die Formulierung der Items scheint folglich ungeeignet für die Erhebung der kognitiven Belastung in Versuchsbedingungen, die sich in der Textverfügbarkeit unterscheiden. Deshalb wäre eine Anpassung der Formulierungen oder die Nutzung anderer Instrumente zu empfehlen, die sich in anderen Studien zum Concept Mapping bewährt haben (s. Kapitel 6.3).

6.2.3 Metakognitive Vorhersage

Bezüglich des Einflusses der Textverfügbarkeit auf die erwartete Lernleistung der Studienteilnehmer:innen zeigte sich in der hier vorgestellten zweiten Studie ein Phänomen, das auch von anderen Forscher:innen (Blunt & Karpicke, 2014; Karpicke, 2012; Roediger & Karpicke, 2006b) berichtet wird: Die Gruppen im Elaborationssetting sagten einen höheren

Lernerfolg voraus als diejenigen im Retrievalsetting. Bei Blunt und Karpicke (2014) zeigte sich, dass die Studierenden in den Retrievalsettings ihren Lernerfolg fälschlicherweise niedriger einschätzten, da sie bessere Leistungen erzielten als die Gruppen, die das Lernmaterial verfügbar hatten. Ein derartiger Leistungsvorteil zugunsten der Retrievalgruppen trat in der hier vorgestellten zweiten Studie nicht auf, da sich die Gruppen bezüglich der Wissenstestergebnisse nicht signifikant voneinander unterschieden (s. Kapitel 5, Tab. 6). Trotzdem unterstützen die JOLs der hier untersuchten Proband:innen die Annahme, dass Lernende in Retrievalsettings ihre Entscheidungen bezüglich der JOLs auf anderen Informationen basieren als die Elaborationsgruppen (Karpicke & Grimaldi, 2012; King et al., 1980; Koriat, 1997; s. Kapitel 2.2.1.3 und 5, 1.2): Während die Retrievalgruppen ihre Urteile darauf beziehen, wie leicht bzw. schwer die Informationen aus dem Gedächtnis abgerufen werden konnten, bewerten Lernende in Elaborationssettings die wahrgenommene Leichtigkeit der Informationsverarbeitung, die bei verfügbarem Lernmaterial einfacher fiel (s. auch Karpicke, 2009; Koriat et al., 2006). Die auf dem Abruf basierten Einschätzungen scheinen jedoch akkurater zu sein, da ihre JOLs die Ergebnisse des deklarativen Wissenstests besser widerspiegeln als die der Elaborationsgruppen. Sowohl Lernende als auch Pädagogen sollten demnach ein Retrievalsetting wählen, um eine realistischere Einschätzung des (eigenen) Lernfortschritts zu erlangen. Dafür spräche auch, dass dieses Setting eher den Bedingungen der Prüfung entspricht, für die der Lernfortschritt im Regelfall erzielt werden soll.

6.2.4 Lernleistung

Bezogen auf die Lernleistung zeigen die Ergebnisse der zweiten Studie, dass neben der Art des vorgeschalteten Concept Mapping-Trainings (s. Kapitel 6.1) auch die Verfügbarkeit des Textes während des Concept Mappings einen Einfluss auf die Qualität von Concept Maps hat.

- **Concept Mapping-bezogene Leistung**

In diesem Zusammenhang zeigten sich gruppenspezifische Unterschiede in Bezug auf den Anteil inhaltlicher Concept Mapping-Fehler: Die Elaborationsgruppen machten signifikant weniger inhaltliche Fehler in ihren Concept Maps und wiesen eine höhere Concept Map-Qualität (bQCM-Index) auf als die Retrieval-Gruppen (s. Kapitel 5, 3.5). Dies bedeutet, dass die Elaborationsgruppen in der Lage waren, die Informationen aus dem Text zu nutzen und inhaltliche Fehler leichter zu vermeiden als die Gruppen ohne Text. Die Vorteile, mit dem Text arbeiten zu können, weniger inhaltliche Fehler in den Concept Maps zu machen und eine höhere Concept Map-Qualität (bQCM-Index) in der Lernphase zu haben, führten jedoch nicht zu vergleichbaren Ergebnissen in der Testphase, in der keine diesbezüglichen signifikanten Qualitätsunterschiede mehr zwischen den Gruppen bestanden. Dies deutet darauf hin, dass die

Textverfügbarkeit einen signifikanten Einfluss auf die Qualität der Propositionen in der Lernphase hatte, nicht aber in der nachgeschalteten Concept Mapping-Aufgabe. Dass der Anteil der inhaltlichen Fehler in allen vier Gruppen im Durchschnitt um das 2,2-fache anstieg, als sie aufgefordert wurden, Concept Maps mit einem vorgegebenen Begriffsset zu erstellen, deutet auf ein Vergessen der Inhalte oder das Wiederaufgreifen früherer Missverständnisse hin. Dies resultierte in der Testphase in entsprechend mehr Fehlern (s. Kapitel 5, 3 und 5, Tab. 8 und 5, 4) und lässt wiederum darauf schließen, dass dieses Maß der Concept Map-Qualität eine limitierte Aussagekraft bezüglich längerfristiger Lernleistungen hat. In Vergleichen solcher Settings sollte dieses Maß demnach nicht als einziges Lernindiz dienen.

- Erwerb deklarativen, strukturellen und konzeptuellen Wissens

Die Ergebnisse zum Wissenserwerb der Lernenden in beiden Textverfügbarkeits-Bedingungen weisen eine umfassende Kongruenz auf, die sich in allen Lernmaßen der Testphase wiederfindet: Die fehlenden Unterschiede in den metakognitiven Vorhersagen der Proband:innen spiegeln die Ergebnisse bezüglich der Concept Map-Qualität wider und zeigen sich auch in den Ergebnissen zum deklarativen, strukturellen und konzeptuellen Wissen (s. Kapitel 5, 3.4 und 5, 3.5.1). Dies deutet darauf hin, dass der höhere Anteil an Elaborationsprozessen in den Concept Maps der Retrievalgruppen offensichtlich nicht zu einer besseren Wissenskonsolidierung beitragen konnte, da die beiden Gruppen in den Tests ähnliche Werte erzielten (s. Kapitel 5, Tab. 6 und 5, 4.1 und 6.2.1).

Eine Überlegenheit des Abrufs in Retrievalsettings beim Wissenserwerb gegenüber der Elaboration in entsprechenden Bedingungen, wie sie z. B. von Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) berichtet wird, konnte hier nicht festgestellt werden. Die Möglichkeit des einmaligen Abrufs zwischen einer Lernaufgabe und dem folgenden Erinnerungstest hatte in mehreren Studien (z. B. Carpenter, 2009; Carpenter & DeLosh, 2006; Endres et al., 2017; Roediger & Karpicke, 2006b) dazu geführt, dass die Behaltensleistung deutlich gesteigert wurde. Die Ergebnisse der zweiten Studie dieses Projektes deuten jedoch darauf hin, dass ein einmaliger Abruf durch eine Concept Mapping-Aufgabe nicht zu besseren Ergebnissen im Vergleich zu Elaborationsgruppen führte, die mit dem Lernmaterial ihre Concept Maps erstellten. Dieser Mangel an Wissensunterschieden zwischen den Gruppen deutet darauf hin, dass Concept Mapping in Retrievalsettings genauso effektiv sein könnte wie Concept Mapping in Elaborationssettings (Karpicke, 2012; Karpicke & Blunt, 2011b). Erst der Blick auf die numerischen Unterschiede lässt die Vermutungen zu, dass es für den Wissenserwerb der Lernenden generell besser sein könnte, das Lernmaterial während des Concept Mappings zur Verfügung zu haben. Steht das Lernmaterial den Lernenden jedoch nicht

zur Verfügung, scheint es eher von Vorteil für ihren Wissenserwerb zu sein, wenn sie ein extensives Training mit Caffolding und Feedback im Concept Mapping erhalten und vertrauter mit dem Einsatz der Methode sind (s. Kapitel 5, Tab. 6 und 5, 4.1).

6.3 Limitationen der beiden Studien

Da die Limitationen studienspezifisch in den jeweiligen Kapiteln vorgestellt wurden (s. Kapitel 4, 4.8 und 5, 4.4), soll die folgende Gesamtbetrachtung der Befunde helfen, Gemeinsamkeiten und Unterschiede festzustellen sowie mögliche Problemfelder zukünftiger Forschungsprojekte aufzeigen. Zu den studienübergreifenden Limitationen gehören u. a. die Reliabilitätswerte einiger verwendeter Instrumente als auch Stichprobencharakteristika, wobei letztere neben dem quasiexperimentellen Design für eine eingeschränkte externe Validität der hier vorgestellten Ergebnisse verantwortlich sind.

Zu den Stichprobencharakteristika gehören die Voraussetzungen der Proband:innen und die Stichprobengröße. In beiden Studien nahmen ausschließlich Studierende teil, die mehrjährige Lernerfahrung besaßen. Das Concept Mapping war für diese Lernenden daher leichter in ihre bestehenden Lernstrategierepertoires integrierbar und dieser homogenisierende Faktor könnte Unterschiede u. a. bezüglich der kognitiven Belastung verschleiert haben. Die Untersuchung anderer Stichproben mit weniger erfahrenen Lernenden könnte andere Ergebnisse hervorbringen. Darüber hinaus waren in beiden Studien die Stichprobengrößen zu gering, weshalb Schlüsse aus den Ergebnissen mit Vorsicht betrachtet werden sollten. In beiden Studien wären mindestens 150 Proband:innen nötig gewesen, um die Effekte mit einer hinreichenden Power feststellen zu können (s. Kapitel 4, 4.8 und 5, 4.4). Solche Zahlen an Studienteilnehmer:innen konnten im Rahmen des Projektes nicht rekrutiert werden, könnten jedoch in Folgeuntersuchungen erreicht werden.

Bezogen auf die in beiden Studien verwendeten Instrumente zeigte sich, dass das Instrument zur Messung der Concept Mapping-Fertigkeiten anscheinend in beiden Studien nicht ausreichend anspruchsvoll für die Studienteilnehmer:innen war. Ein komplexerer Test mit methodischen und inhaltlichen Fehlern hätte zu klareren Ergebnissen führen können und sollte in zukünftigen Studien in Betracht gezogen werden. Ferner erwies sich die Reliabilität der Instrumente zur Erfassung des Leseverständnisses (LGVT; Schneider et al., 2007; s. Kapitel 4, 2.4.1) und der kognitiven Belastung (verändert nach Klepsch et al., 2017; s. Kapitel 4, 2.4.4) in der ersten Studie als akzeptabel, in der zweiten Studie jedoch als teilweise inakzeptabel (s. Kapitel 5, 2.4.1 und 5, 4.4 und 6.2.2). Bezogen auf diese Variablen war im hier vorgestellten Projekt keine Interpretation der Ergebnisse möglich. Im Zusammenhang mit Concept Mapping sollte daher in zukünftigen Studien, die mit ähnlichen Versuchsbedingungen und Stichproben

arbeiten, ein anderer Fragebogen zur Erfassung der kognitiven Belastung in Erwägung gezogen werden. Eine Alternative stellt z. B. der NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) dar, der ein mehrdimensionales Bewertungsverfahren ist und im Zusammenhang mit Concept Mapping an Computern von Hilbert und Renkl (2009) eingesetzt wurde. Der Gesamtwert der Arbeitsbelastung (max. Punktzahl: 100) basiert auf einem gewichteten Durchschnitt der Bewertungen auf sechs Subskalen (geistige Anforderung, körperliche Anforderung, zeitliche Anforderung, Leistung, Anstrengung und Frustration). Auch wenn sich die berichtete kognitive Belastung der Proband:innen in der Studie von Hilbert und Renkl (2009) im unteren Bereich befand, konnten zwischen den Gruppen signifikante Gruppenunterschiede festgestellt werden. Diesbezüglich gestehen die Autor:innen jedoch ein, dass die Arbeit mit Computern selbst eine kognitive Belastung darstellen kann, und diese zusätzliche Belastung wird von ihnen selbst als möglicher Grund für die niedrigen Lernleistungen der Proband:innen in Erwägung gezogen.

7 Zusammenfassung

Das Ziel der Dissertation und des vorgestellten Forschungsprojektes war es, einen substanziellen Beitrag zur Identifikation lernwirksamer Faktoren des Trainings und Einsatzes von Concept Mapping zu leisten. Im Hinblick auf den generellen Einsatz von Concept Mapping konnte anhand der beiden vorgestellten Studien in Übereinstimmung mit bisherigen Befunden gezeigt werden, dass Concept Mapping eine lerneffektive Methode darstellt (z. B. Cañas et al., 2003; Nesbit & Adesope, 2006; Schroeder et al., 2018). Die Ergebnisse der ersten Studie zeigen, dass Lernende durch ein extensives Training ein Maß an Concept Mapping-Fertigkeiten und Vertrautheit mit Concept Mapping erreichen können, um die Methode lernwirksam anzuwenden. Die lernwirksame Anwendung äußerte sich u. a. in signifikant besseren Ergebnissen der Concept Mapping-Trainingsgruppen in den Tests zum strukturellen und konzeptuellen Wissen im Vergleich zur Kontrolltrainingsgruppe. Zudem konnte gezeigt werden, dass insbesondere ein extensives Training mit Scaffolding und Feedback den Lernenden hilft, methodisch korrekte Propositionen zu erstellen. Dies ist nicht nur für sie selbst, sondern auch für die Lehrkraft hilfreich, um eine korrekte Vorstellung bezüglich der kognitiven Strukturen der Lernenden und deren Verständnis der Lerninhalte zu erlangen. Die Bedeutung des Trainings für Concept Mapping-Fertigkeiten wurde auch in der zweiten Studie bestätigt, allerdings kann anhand der Ergebnisse davon ausgegangen werden, dass auch eine Kurzeinführung ausreicht, um ein Maß an Vertrautheit zu erlangen, welches in keinen Nachteilen beim Erwerb deklarativen, strukturellen und konzeptuellen Wissens resultiert. Da

der Trainingsansatz der Kurzeinführung jedoch in beiden Studien zu mehr methodischen Concept Mapping-Fehlern geführt hat, welche die Reliabilität und Validität der Concept Maps als Diagnoseinstrument einschränken können (Ruiz-Primo, Schultz, et al., 2001), ist die Kurzeinführung in das Concept Mapping in einem solchen Rahmen nicht zu empfehlen. Lehrkräfte sollten daher Concept Mapping extensiv trainieren und Lernende dabei durch Scaffolding und Feedback unterstützen, so dass die Concept Map-Ersteller:innen in der Lage sind, methodisch korrekte Propositionen zu erstellen. Auf die in den Concept Maps dargestellten kognitiven Prozesse des Abrufs, der Organisation und Elaboration hatte das Concept Mapping-Training entgegen den Erwartungen keine signifikanten Auswirkungen. Propositionen, die Abruf suggerieren, waren am häufigsten, wohingegen Propositionen zur Organisation und Elaboration nur in geringem Maße erstellt wurden, was jedoch auf eine fehlende explizite Aufforderung dazu in den Aufgaben zurückgeführt werden kann.

Die Ergebnisse zur kognitiven Belastung aus beiden Studien konnten aufgrund der eingeschränkten Reliabilität der Skalen lediglich in der ersten Studie interpretiert werden. Hier zeigte sich ein signifikanter Unterschied aufgrund der verschiedenen Bedingungen bei der extrinsischen kognitiven Belastung, die von der Concept Mapping-Trainingsgruppe mit wiederholtem Üben signifikant niedriger eingeschätzt wurde als von der Kontrolltrainingsgruppe. Da jedoch alle Studienteilnehmer:innen mit den gleichen Materialien arbeiteten, auf die sich die Items bezüglich der extrinsischen kognitiven Belastung bezogen, kann dieser Unterschied auch auf die Reliabilität der Skala zurückführbar sein.

Bezüglich des Einsatzes von Concept Mapping in unterschiedlichen Textverfügbarkeitsbedingungen zeigte sich in der zweiten Studie, dass Elaborationsprozesse wider Erwarten mehr durch ein Retrievalsetting als durch ein Elaborationssetting gefördert werden. Das höhere Maß an Elaboration führte zu keinen Vorteilen der Retrievalgruppen in den Tests zum Wissenserwerb, da alle Gruppen ähnliche Erfolge erzielten. Dies steht im Kontrast zu anderen Studien (z. B. Blunt & Karpicke, 2014; O'Day & Karpicke, 2020), in denen sich der Abruf in Retrievalsettings beim Wissenserwerb als wirksamer im Vergleich zur Elaboration in entsprechenden Bedingungen erwies. Ferner deuten die Ergebnisse der zweiten hier vorgestellten Studie darauf hin, dass metakognitive Prozesse von Lernenden in einem Retrievalsetting stärker gefördert wurden und sie ihre Lernleistungen akkurater einschätzten als in einem Elaborationssetting. Dies stünde im Einklang mit bisherigen Befunden aus Studien (z. B. Kelemen, 2000; King et al., 1980), in denen berichtet wird, dass ein Abruf aus dem Gedächtnis zwischen Lern- und Testphase zu einer genaueren Selbsteinschätzung bezüglich der eigenen Lernleistung führte. Daher könnte von Lehrkräften und Lernenden ein Retrievalsetting

vorzuziehen sein, um metakognitive Prozesse der Evaluation zu fördern und das Lernen zu optimieren, da u. a. Entscheidungen bezüglich weiterer Lernschritte auf der Evaluation voriger Lernphasen basieren (Nelson & Dunlosky, 1991). Da beide Studien jedoch Limitationen besaßen, bleibt die Aussagekraft der Ergebnisse diesbezüglich begrenzt, weshalb weitere Studien zur Bedeutung der Textverfügbarkeit beim Concept Mapping für metakognitive Prozesse vonnöten sind. Ebenso besteht Bedarf nach weiteren Untersuchungen zu Parametern von Trainings im Concept Mapping, um die Vermittlung der Methode durch Lehrende und die Anwendung durch Lernende zu optimieren. Die Ergebnisse der beiden vorgestellten Studien widersprechen sich zwar bezüglich der Bedeutung eines extensiven Trainings für den Wissenserwerb in den Domänen des deklarativen, strukturellen und konzeptuellen Wissens, spiegeln aber den diesbezüglichen Forschungsstand wider und unterstreichen das Erfordernis weiterer empirischer Forschung.

8 Ausblick

Bereiche, die sich für zukünftige Forschungsprojekte anbieten, ergeben sich einerseits aufgrund der Ergebnisse, die das Projekt liefert, andererseits aufgrund der Limitationen der Studien, die in den jeweiligen Originalarbeiten diskutiert wurden (s. Kapitel 4 und 5). Im Folgenden werden die fünf wichtigsten Vorschläge aus beiden Studien für die Gestaltung zukünftiger Studien in diesem Bereich zusammengefasst und erläutert, weshalb auf fortlaufende Quellenverweise auf die Originalarbeiten in diesem Abschnitt verzichtet wird.

- (1) Aufgrund der Effektivitätsunterschiede der Kurzeinführung in Bezug auf den Wissenserwerb in den drei Wissensdomänen zwischen der ersten und zweiten Studie sollten künftige Studien jene spezifischen Trainingselemente ermitteln, die für Lernende hilfreich und essenziell für eine lernwirksame Anwendung des Concept Mappings sind. Eine systematische Variation der Gesamtzahl der Trainingssitzungen, der Sitzungsdauer sowie dem zeitlichen Abstand zwischen diesen einzelnen Sitzungen kann zur Klärung der Abhängigkeit des Trainingserfolgs von diesen Parametern führen. Des Weiteren wäre interessant herauszufinden, welche metakognitiven Prozesse auf welche Weise durch das Concept Mapping im Allgemeinen und durch ein Training im Speziellen gefördert werden. Dies könnte zur Klärung der Frage beitragen, welche Rolle Concept Mapping und dessen Vermittlung bei der Entwicklung metakognitiver Fertigkeiten spielt.

- (2) In diesem Zusammenhang könnte zusätzlich der Faktor Motivation stärker berücksichtigt und als potenzielle Kovariate erhoben werden. Die Motivation beeinflusst das Lernpotenzial bzw. -verhalten von Personen (z. B. Ryan & Deci, 2020), weshalb das Training darauf ausgelegt sein sollte, den Lernenden die Vorteile der Lernstrategie aufzuzeigen und motivationalen Problemen vorzubeugen. Ein solcher Ansatz wurde z. B. von Hübner et al. (2010) im Zusammenhang mit Lerntagebüchern implementiert, und festgestellt, dass er nicht nur in der Trainings-, sondern auch der Transferphase zu besseren Lernergebnissen führte.
- (3) Beim Vergleich der Kriterien für bedeutungsvolles Lernen mit den Anforderungen des Concept Mappings fällt auf, dass Concept Maps zwar das Wissen der Lernenden in Bezug auf ihr Vorwissen sowohl visualisieren als auch eine Analyse der Wissensstrukturen zulassen, es jedoch die Aufgabe der Lehrkraft ist, im Vorhinein zu eruiieren, welche Vorwissensbasis die Person bezüglich des Lerninhalts besitzt, da neue Inhalte mit bestehenden verknüpft werden müssen. Dies wurde mithilfe der Vorbefragung in beiden Studien umgesetzt. Des Weiteren muss die Lehrkraft dafür sorgen, dass diese Basis auch tatsächlich genutzt und das Vorwissen in die Concept Map integriert wird, so dass bedeutungsvoll gelernt werden kann (Novak & Cañas, 2008). Dies wurde in beiden Studien mit entsprechenden Aufgabenstellungen in der Lernphase zu erreichen versucht. Allerdings hatte sich in den Concept Maps der Teilnehmer:innen gezeigt, dass Organisations- und Elaborationspropositionen nur in sehr geringem Maße erstellt wurden (s. Kapitel 4, 3.4 und 5, 3.5.3). Wenn die Kriterien des bedeutungsvollen Lernens zur Betrachtung dieser Ergebnisse herangezogen werden, entsteht der Eindruck, dass die Teilnehmer:innen nur z. T. in der Lage waren, bedeutungsvoll zu lernen. Die fehlende Verknüpfung neuer Informationen mit dem Vorwissen mag demnach ein Grund für das Maß an Vergessen zwischen der Lern- zur Testphase gewesen sein. Für diese Annahme würden Befunde sprechen, die belegen, dass aktiv in das Vorwissen integrierte Informationen besser behalten werden (Novak, 2003; Okebukola, 1990). Dies wiederum unterstreicht die Notwendigkeit, Lernende aktiv zur Elaboration aufzufordern. Dementsprechend wichtig wäre es, herauszufinden, wie vorgeschaltete Trainingsmaßnahmen gestaltet und Aufgabenstellungen (auch in der Lernphase) formuliert werden müssen, um die erstrebenswerten kognitiven Prozesse zu induzieren und das Lernen zu optimieren.

Da Concept Mapping auch in anderen Wissenschafts- und Praxisbereichen vermittelt wird, hätte eine solche Optimierung des Trainings fächerübergreifende Konsequenzen. Als interdisziplinäre Lernmethode sollte Concept Mapping von den Themen der kritischen Lernphase unabhängig eingesetzt werden können und Concept Mapping-Fertigkeiten, die in einem Bereich erworben wurden, sollten auf andere Domänen übertragbar sein. Die hier vorgestellten Ergebnisse deuten zumindest darauf hin, dass dies gelingen kann, da die Inhalte der Trainingsphase (Thema Intelligenz) sich von denen der Lern- und Testphase (Thema Zellbiologie) unterschieden, aber ein lernwirksamer Einsatz der Methode in den beiden letzten Phasen beobachtet werden konnte (s. Kapitel 4 und 5).

- (4) Anhand der Ergebnisse der Gruppen im Test zum deklarativen Wissen aus Studie 1 liegt die Vermutung nahe, dass der Nutzen des Concept Mappings eher in der Förderung des Erwerbs von strukturellem und konzeptuellem Wissen liegt und damit den Anforderungen kognitiv anspruchsvollerer Leistungen im Vergleich zum Abruf von Fakten entspricht (L. W. Anderson et al., 2001; Krathwohl, 2002; s. Kapitel 2.1.2 und 2.2). Zur Steigerung des Erwerbs deklarativen Wissens wurde daher in Studie 2 das Concept Mapping in einem Retrievalsetting durchgeführt, da diese Bedingung sich in vorigen Studien als lernwirksam in diesem Bereich herausgestellt hatte (z. B. Moreira et al., 2019; Roediger & Karpicke, 2006b). Allerdings hatten genau diese Gruppen der Retrievalsettings in der zweiten Studie die schlechteren Ergebnisse im deklarativen Wissenstest, auch wenn diese Unterschiede statistisch nicht signifikant waren. Daher sollte in künftigen Studien geklärt werden, welche Rolle Concept Mapping beim Erwerb deklarativen Wissens spielt und wie die Retrievalsettings optimiert werden können.
- (5) Im direkten Zusammenhang mit Punkt (4) besteht auch Klärungsbedarf bezüglich der Frage, welches Maß an Vertrautheit der Lernenden mit dem Retrievalsetting für den Wissenserwerb nötig ist. Im Rahmen der zweiten Studie wurde der Teil der Proband:innen, der es aus den Trainingssitzungen gewohnt war, eine Concept Map mit dem verfügbaren Text zu erstellen, mit der neuen Situation konfrontiert, eine Concept Map ohne den Text zu erstellen. Diese Änderung eines gewohnten Ablaufs könnte zu Leistungseinbußen geführt haben, weshalb es interessant wäre, herauszufinden, welchen Einfluss die Implementierung solcher Phasen bereits im Rahmen des Trainings auf das Lernen hätte. Eine Alternierung der Textverfügbarkeit wurde bei Karpicke und Blunt (2011), Blunt und Karpicke (2014) sowie O'Day und Karpicke (2020) in der Lernphase umgesetzt, aber nicht in einer vorgeschalteten Trainingsphase – zumal

Letztere in den Studien nur aus einer Kurzeinführung in das Concept Mapping bestand. Des Weiteren wurden in diesen Studien kurze Texte von weniger als 300 Wörtern als Lernmaterial ausgehändigt, was unter den generellen Anforderungen für Studierende liegt und im Rahmen des hier vorgestellten Projektes durch einen Lerntext von ca. 2.000 Wörtern angepasst wurde (vgl. Kapitel 1 und 2.3.2 sowie 4, 2 und 5, 2). Weitere Studien könnten daher experimentelle Bedingungen schaffen, die den universitären ebenfalls näherkommen und die Bedeutung der Vertrautheit von Lernenden mit Retrievalsettings für die Lernphase untersuchen.

Die Entwicklung und Durchführung solcher Studien würde dazu beitragen, die Bedingungen für ein effizientes Training und einen lerneffektiven Einsatz von Concept Mapping zu ermitteln. Auf diese Weise könnte den Lernenden eine Methode an die Hand gegeben werden, mit der sie bedeutungsvoll lernen und ihr vernetztes Denken fördern können. Concept Mapping könnte in diesem Fall Lernenden dabei helfen, eine Wissensbasis aus gut strukturierten und vernetzten Konzepten zu Themengebieten in unterschiedlichen Disziplinen aufzubauen. Diese Basis würde ihnen wiederum ermöglichen, sowohl ein tieferes Verständnis des jeweiligen Fachgebiets zu erlangen als auch erfolgreicher am gesellschaftlichen und wissenschaftlichen Diskurs teilnehmen zu können.

Literaturverzeichnis

- Acton, W. H., Johnson, P. J., & Goldsmith, T. E. (1994). Structural knowledge assessment: Comparison of referent structures. *Journal of Educational Psychology*, 86(2), 303–311. <https://doi.org/10.1037/0022-0663.86.2.303>
- Adesope, O. O., & Nesbit, J. C. (2009). A systematic review of research on collaborative learning with concept maps. In P. L. Torres & R. d. C. V. Marriott (Eds.), *Handbook of research on collaborative learning using concept mapping* (pp. 238–255). IGI Global.
- Ajaja, O. P. (2011). Concept mapping as a study skill. *International Journal of Educational Sciences*, 3(1), 49–57. <https://doi.org/10.1080/09751122.2011.11890008>
- Ajaja, O. P. (2013). Which way do we go in biology teaching? Lecturing, concept mapping, cooperative learning or learning cycle? *Electronic Journal of Science Education*, 17(1).
- Allen, D., & Tanner, K. (2002). Approaches to cell biology teaching: Questions about questions. *Cell Biology Education*, 1(3), 63–67. <https://doi.org/10.1187/cbe.02-07-0021>
- Allen, D., & Tanner, K. (2003). Approaches to cell biology teaching: Mapping the journey - concept maps as signposts of developing knowledge structures. *Cell Biology Education*, 2(3), 133–136. <https://doi.org/10.1187/cbe.03-07-0033>
- Anderman, E. M., Sinatra, G. M., & Gray, D. L. (2012). The challenges of teaching and learning about science in the twenty-first century: Exploring the abilities and constraints of adolescent learners. *Studies in Science Education*, 48(1), 89–117. <https://doi.org/10.1080/03057267.2012.655038>
- Anderson, J. R. (1976). *Language, memory and thought*. Lawrence Erlbaum.
- Anderson, J. R. (1983). *The architecture of cognition* (Vol. 1114). Harvard University Press.
- Anderson, J. R. (2013). *Kognitive Psychologie* (J. Funke, Ed.) (7th ed.). Springer.
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Anderson, R. C. (1984). Some reflections on the acquisition of knowledge. *Educational Researcher*, 13(9), 5–10. <https://doi.org/10.3102/0013189X013009005>
- Andrews, K. E., Tressler, K. D., & Mintzes, J. J. (2008). Assessing environmental understanding: An application of the concept mapping strategy. *Environmental Education Research*, 14(5), 519–536. <https://doi.org/10.1080/13504620802278829>
- Anohina-Naumeca, A. (2015). Justifying the usage of concept mapping as a tool for the formative assessment of the structural knowledge of engineering students. *Knowledge Management & E-Learning*, 7(1), 56–72. <https://doi.org/10.34105/j.kmel.2015.07.005>

- Arnaudin, M. W., & Mintzes, J. J. (1985). Students' alternative conceptions of the human circulatory system: A cross-age study. *Science Education*, *69*(5), 721–733.
- Artelt, C. (1999). Lernstrategien und Lernerfolg: Eine handlungsnahe Studie. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *31*(2), 86–96.
- Artelt, C. (2000). *Strategisches Lernen*. Waxmann.
- Artelt, C. (2006). Lernstrategien in der Schule. In H. Mandl & H. F. Friedrich (Eds.), *Handbuch Lernstrategien* (pp. 337–351). Hogrefe.
- Assaraf, O. B.-Z., Dodick, J., & Tripto, J. (2013). High school students' understanding of the human body system. *Research in Science Education*, *43*(1), 33–56. <https://doi.org/10.1007/s11165-011-9245-2>
- Assaraf, O. B.-Z., & Orion, N. (2005). Development of system thinking skills in the context of earth system education. *Journal of Research in Science Teaching*, *42*(5), 518–560. <https://doi.org/10.1002/tea.20061>
- Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. Grune & Stratton.
- Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. Holt Rinehart.
- Ausubel, D. P. (2000). *The acquisition and retention of knowledge: A cognitive view*. Springer.
- Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational psychology: A cognitive view* (2nd ed.). Holt Rinehart & Winston.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. <https://doi.org/10.1146/annurev-psych-120710-100422>
- Bannert, M. (2003). Effekte metakognitiver Lernhilfen auf den Wissenserwerb in vernetzten Lernumgebungen. *Zeitschrift für Pädagogische Psychologie*, *17*(1), 13–25.
- Bannert, M. (2007). *Metakognition beim Lernen mit Hypermedien: Erfassung, Beschreibung und Vermittlung wirksamer metakognitiver Strategien und Regulationsaktivitäten*. Waxmann.
- Bannert, M., Hildebrand, M., & Mengelkamp, C. (2009). Effects of a metacognitive support device in learning environments. *Computers in Human Behavior*, *25*(4), 829–835. <https://doi.org/10.1016/j.chb.2008.07.002>
- Barak, J., Sheva, B., Gorodetsky, M., & Gurion, B. (1999). As 'process' as it can get: Students' understanding of biological processes. *International Journal of Science Education*, *21*(12), 1281–1292. <https://doi.org/10.1080/095006999290075>
- Bascones, J., & Novak, J. D. (1985). Alternative instructional systems and the development of problem-solving skills in physics. *European Journal of Science Education*, *7*(3), 253–261. <https://doi.org/10.1080/0140528850070304>

- Baumert, J. (1993). Lernstrategien, motivationale Orientierung und Selbstwirksamkeitsüberzeugungen im Kontext schulischen Lernens. *Unterrichtswissenschaft*, 21(4), 327–354.
- Baumert, J., & Köller, O. (1996). Lernstrategien und schulische Leistungen. In J. Möller & O. Köller (Eds.), *Pädagogische Psychologie Motivationspsychologie. Emotionen, Kognitionen und Schulleistung* (pp. 137–154). Beltz Psychologie-Verl.-Union.
- Bayram, S. (1995). *The effectiveness of concept and software mapping for representing student data and process schema in science* [Thesis, University of Pittsburgh, Pittsburgh, PA]. University of Pittsburgh.
- Becker, L. B., & Großschedl, J. (2019). Zerreißprobe für den Körper - Störungen im Proteintransport aufschlüsseln und dabei Systemebenen verbinden. *Unterricht Biologie*, 43(449), 24–29.
- Becker, L. B., Welter, V. D. E., Aschermann, E., & Großschedl, J. (2021). Comprehension-oriented learning of cell biology: Do different training conditions affect students' learning success differentially? *Education Sciences*, 11(8), 438. <https://doi.org/10.3390/educsci11080438>
- Becker, L. B., Welter, V. D. E., & Großschedl, J. (2021). Effects of strategy training and elaboration vs. retrieval settings on learning of cell biology using concept mapping. *Education Sciences*, 11(9), 530. <https://doi.org/10.3390/educsci11090530>
- Bentley, F., Kennedy, S., & Semsar, K. (2011). How not to lose your students with concept maps. *Journal of College Science Teaching*, 41(1), 61–68. <https://pdfs.semanticscholar.org/a11b/9d361c0e1055b1ebb37a7177efdd03b7aa78.pdf>
- Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction*, 17(5), 564–577.
- Besterfield-Sacre, M., Gerchak, J., Lyons, M., Shuman, L. J., & Wolfe, H. (2004). Scoring concept maps: An integrated rubric for assessing engineering education. *Journal of Engineering Education*, 19(2), 105–115.
- Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher & A. Koriat (Eds.), *Attention and performance: Vol. 14. Attention and Performance XVII: Cognitive regulation of performance, interaction of theory and application* (pp. 435–459). MIT Press.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. Longman.

- Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology, 106*(3), 849–858.
- Bonner, J. M., & Holliday, W. G. (2006). How college science students engage in note-taking strategies. *Journal of Research in Science Teaching, 43*(8), 786–818.
- Boo, H. K., & Hoh, Y. K. (2001). Using concept maps to enhance meaningful chemistry learning. *Journal of Science and Mathematics Education in Southeast Asia, 14*(2), 78–88.
- Bortz, J., & Döring, N. (2016). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler* (5th ed.). *Springer-Lehrbuch*. Springer.
- BouJaoude, S., & Attieh, M. (2008). The effect of using concept maps as study tools on achievement in chemistry. *Eurasia Journal of Mathematics, Science and Technology Education, 4*(3), 233–246. <https://doi.org/10.12973/ejmste/75345>
- Brame, C. J., & Biel, R. (2015). Test-enhanced learning: The potential for testing to promote greater learning in undergraduate science courses. *CBE-Life Sciences Education, 14*(2), 14:es4. <https://doi.org/10.1187/cbe.14-11-0208>
- Bramwell-Lalor, S., & Rainford, M. (2014). The effects of using concept mapping for improving advanced level biology students' lower- and higher-order cognitive skills. *International Journal of Science Education, 36*(5), 839–864. <https://doi.org/10.1080/09500693.2013.829255>
- Brandstädter, K., Harms, U., & Großschedl, J. (2012). Assessing System Thinking Through Different Concept-Mapping Practices. *International Journal of Science Education, 34*(14). <https://doi.org/10.1080/09500693.2012.716549>
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press.
- Brown, M. H., & Schwartz, R. S. (2009). Connecting photosynthesis and cellular respiration: Preservice teachers' conceptions. *Journal of Research in Science Teaching, 46*(7), 791–812. <https://doi.org/10.1002/tea.20287>
- Bruckermann, T., Aschermann, E., Bresges, A., & Schlüter, K. (2017). Metacognitive and multimedia support of experiments in inquiry learning for science teacher preparation. *International Journal of Science Education, 39*(6), 701–722. <https://doi.org/10.1080/09500693.2017.1301691>
- Brüning, L., & Saum, T. (2017). *Erfolgreich unterrichten durch Visualisieren: Die Kraft von Concept Maps & Co*. Neue Deutsche Schule Verlagsgesellschaft mbH.

- Buchner, A., & Brandt, M. (2017). Gedächtniskonzeptionen und Wissensrepräsentationen. In J. Müsseler & M. Rieger (Eds.), *Allgemeine Psychologie* (3rd ed., pp. 401–434). Springer. https://doi.org/10.1007/978-3-642-53898-8_12
- Budé, L., Imbos, T., van de Wiel, M. W., & Berger, M. P. (2011). The effect of distributed practice on students' conceptual understanding of statistics. *Higher Education*, *62*(1), 69–79. <https://doi.org/10.1007/s10734-010-9366-y>
- Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, *13*(4), 273–281. <https://doi.org/10.1037/1076-898X.13.4.273>
- Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 918–928. <https://doi.org/10.1037/0278-7393.34.4.918>
- Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition*, *36*(3), 604–616. <https://doi.org/10.3758/MC.36.3.604>
- Cadorin, L., Bagnasco, A., Rocco, G., & Sasso, L. (2014). An integrative review of the characteristics of meaningful learning in healthcare professionals to enlighten educational practices in health care. *Nursing Open*, *1*(1), 3–14. <https://doi.org/10.1002/nop2.3>
- Campbell, N. A., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Reece, J. B. (2021). *Biology: A global approach* (12th ed.). Pearson.
- Cañas, A. J., Coffey, J., Carnot, M. J., Feltovich, P., Hoffman, R. R., Feltovich, J., & Novak, J. D. (2003). *A summary of literature pertaining to the use of concept mapping techniques and technologies for education and performance support*. The Institute for Human and Machine Cognition. <https://www.ihmc.us/users/acanas/Publications/ConceptMapLitReview/IHMC%20Literature%20Review%20on%20Concept%20Mapping.pdf>
- Cañas, A. J., & Novak, J. D. (2014). Concept mapping using CmapTools to enhance meaningful learning. In A. Okada (Ed.), *Knowledge cartography: Software tools and mapping techniques* (2nd ed., pp. 23–45). Springer. https://doi.org/10.1007/978-1-4471-6470-8_2

- Cañas, A. J., Novak, J. D., & Reiska, P. (2012). Freedom vs. restriction of content and structure during concept mapping-possibilities and limitations for construction and assessment. In A. J. Cañas, J. D. Novak, & J. Vanhear (Eds.), *Proceedings of the Fifth International Conference on Concept Mapping* (Vol. 2, pp. 247–257). University of Malta. https://www.researchgate.net/profile/Alberto_Canas/publication/248392782_Freedom_vs_Restriction_of_Content_and_Structure_during_Concept_Mapping_-_Possibilities_and_Limitations_for_Construction_and_Assessment/links/00b7d51ded8390185c000000.pdf
- 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>
- Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & Cognition*, 34(2), 268–276. <https://doi.org/10.3758/BF03193405>
- Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of U.S. history facts. *Applied Cognitive Psychology*, 23(6), 760–771. <https://doi.org/10.1002/acp.1507>
- Cassata, A. E., Himangshu, S., & Iuli, R. J. (2004). What do you know? Assessing change in student conceptual understanding in science. In A. J. Cañas, J. D. Novak, & F. M. González García (Eds.), *Concept Maps: Theory, Methodology, Technology: Proceedings of the First International Conference on Concept Mapping* (pp. 127–130). Universidad Pública de Navarra.
- Cauvin, S., Moulin, C. J. A., Souchay, C., Kliegel, M., & Schnitzspahn, K. M. (2019). Prospective memory predictions in aging: Increased overconfidence in older adults. *Experimental Aging Research*, 45(5), 436–459. <https://doi.org/10.1080/0361073X.2019.1664471>
- Chang, K. E., Sung, Y. T., & Chen, S. F. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning*, 17(1), 21–33. <https://doi.org/10.1111/j.1365-2729.2001.00156.x>
- Chang, K. E., Sung, Y.-T., & Chen, I.-D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education*, 71(1), 5–23. <https://doi.org/10.1080/00220970209602054>

- Chen, C.-H., & Bradshaw, A. C. (2007). The effect of web-based question prompts on scaffolding knowledge integration and ill-structured problem solving. *Journal of Research on Technology in Education*, 39(4), 359–375. <https://doi.org/10.1080/15391523.2007.10782487>
- Chen, P., Chavez, O., Ong, D. C., & Gunderson, B. (2017). Strategic resource use for learning: A self-administered intervention that guides self-reflection on effective resource use enhances academic performance. *Psychological Science*, 28(6), 774–785. <https://doi.org/10.1177/0956797617696456>
- Chevron, M.-P. (2014). A metacognitive tool: Theoretical and operational analysis of skills exercised in structured concept maps. *Perspectives in Science*, 2(1-4), 46–54. <https://doi.org/10.1016/j.pisc.2014.07.001>
- Chi, M. T. H., & Roscoe, R. D. (2002). The processes and challenges of conceptual change. In M. Limón & L. Mason (Eds.), *Reconsidering Conceptual Change: Issues in Theory and Practice* (pp. 3–27). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47637-1_1
- Chi, M. T., Slotta, J. D., & Leeuw, N. de (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, 4(1), 27–43. [https://doi.org/10.1016/0959-4752\(94\)90017-5](https://doi.org/10.1016/0959-4752(94)90017-5)
- Chiou, C.-C. (2008). The effect of concept mapping on students' learning achievements and interests. *Innovations in Education and Teaching International*, 45(4), 375–387. <https://doi.org/10.1080/14703290802377240>
- Chiu, C.-H. (2004). Evaluating system-based strategies for managing conflict in collaborative concept mapping. *Journal of Computer Assisted Learning*, 20(2), 124–132. <https://doi.org/10.1111/j.1365-2729.2004.00072.x>
- Chularut, P., & DeBacker, T. K. (2004). The influence of concept mapping on achievement, self-regulation, and self-efficacy in students of English as a second language. *Contemporary Educational Psychology*, 29(3), 248–263. <https://doi.org/10.1016/j.cedpsych.2003.09.001>
- Cimolino, L., Kay, J., & Miller, A. (2003). *Incremental student modelling and reflection by verified concept-mapping*. https://cmapspublic3.ihmc.us/rid=1MM1V3QPH-1524SZS-B12B/vol5_cimolino.pdf

- Clark, R. E. (1989). When teaching kills learning: Research on mathemathantics. In Mandl, E. D. Corte, N. S. Bennet, & H. F. Friedrich (Eds.), *Learning and Instruction: European research in an international context* (Vol. 2, pp. 1–22). Pergamon Press. https://www.researchgate.net/publication/234744652_When_Teaching_Kills_Learning_Types_of_Mathemathantic_Effects
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 240–247. [https://doi.org/10.1016/S0022-5371\(69\)80069-1](https://doi.org/10.1016/S0022-5371(69)80069-1)
- Cook, D. L., & King, J. (1968). A study of the Hawthorne Effect in educational research. *Research in the Teaching of English*, 2(2), 93–98. <http://www.jstor.org/stable/40170508>
- Cox, R. (1999). Representation construction, externalised cognition and individual differences. *Learning and Instruction*, 9(4), 343–363. [https://doi.org/10.1016/S0959-4752\(98\)00051-6](https://doi.org/10.1016/S0959-4752(98)00051-6)
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12(11), 671–684. <https://doi.org/10.1037/h0043943>
- Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30(2), 116–127. <https://doi.org/10.1037/h0076829>
- Cronin, P. J., Dekkers, J., & Dunn, J. G. (1982). A procedure for using and evaluating concept maps. *Research in Science Education*, 12(1), 17–24. <https://doi.org/10.1007/BF02357009>
- Dahnke, H., Fuhrmann, A., & Steinhagen, K. (1998). Entwicklung und Einsatz von Computersimulation und Concept Mapping als Erhebungsinstrumente bei Vorstellungen zur Wärmephysik eines Hauses. *Zeitschrift für Didaktik der Naturwissenschaften*, 4(2), 67–80.
- Dansereau, D. (1978). The development of a learning strategies curriculum. In H. F. O’Neil (Ed.), *The Educational technology series. Learning strategies* (pp. 1–29). Academic Press. <https://doi.org/10.1016/B978-0-12-526650-5.50006-X>

- Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences*, 12(1), 91–142. https://doi.org/10.1207/S15327809JLS1201_4
- Debie, N., & van de Leemput, C. (2014). What does germane load mean? An empirical contribution to the cognitive load theory. *Frontiers in Psychology*, 5, 1099. <https://doi.org/10.3389/fpsyg.2014.01099>
- Del Fernández Fernández, M. M., & Jiménez Tejada, M. P. (2019). Difficulties learning about the cell. Expectations vs. reality. *Journal of Biological Education*, 53(3), 333–347. <https://doi.org/10.1080/00219266.2018.1469542>
- den Elzen-Rump, V., & Leutner, D. (2007). Naturwissenschaftliche Sachtexte verstehen – ein computerbasiertes Trainingsprogramm für Schüler der 10. Jahrgangsstufe zum selbstregulierten Lernen mit einer Mapping-Strategie. In M. Landmann & B. Schmitz (Eds.), *Pädagogische Psychologie. Selbstregulation erfolgreich fördern: Praxisnahe Trainingsprogramme für effektives Lernen* (pp. 251–268). Kohlhammer.
- den Elzen-Rump, V., Leutner, D., & Wirth, J. (2008). Lernstrategien im Unterrichtsalltag. In S. Kliemann (Ed.), *Schülergerechtes Arbeiten in der Sekundarstufe I: Diagnostizieren und Fördern* (pp. 101–111). Cornelsen Scriptor.
- DiCarlo, S. E. (2006). Cell biology should be taught as science is practised. *Nature Reviews. Molecular Cell Biology*, 7(4), 290–296. <https://doi.org/10.1038/nrm1856>
- DiSessa, A. A. (2001). *Changing minds: Computers, learning, and literacy*. MIT Press. <http://cognet.mit.edu/library/books/view?isbn=0262541327>
- Dobson, J. L., Linderholm, T., & Stroud, L. (2019). Retrieval practice and judgements of learning enhance transfer of physiology information. *Advances in Health Sciences Education*, 24(3), 525–537. <https://doi.org/10.1007/s10459-019-09881-w>
- Dowd, J. E., Duncan, T., & Reynolds, J. A. (2015). Concept maps for improved science reasoning and writing: Complexity isn't everything. *CBE-Life Sciences Education*, 14(4), ar39. <https://doi.org/10.1187/cbe.15-06-0138>
- Dreyfus, A., & Jungwirth, E. (1989). The pupil and the living cell: A taxonomy of dysfunctional ideas about an abstract idea. *Journal of Biological Education*, 23(1), 49–55. <https://doi.org/10.1080/00219266.1989.9655024>
- Driver, R. (1989). Students' conceptions and the learning of science. *International Journal of Science Education*, 11(5), 481–490. <https://doi.org/10.1080/0950069890110501>

- Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, 22(4), 271–280. <https://doi.org/10.1016/j.learninstruc.2011.08.003>
- Edelmann, W., & Wittmann, S. (2019). *Lernpsychologie* (8. Aufl.). Beltz.
- Egger-Rainer, A. (2019). Enhancing validity through cognitive interviewing. A methodological example using the Epilepsy Monitoring Unit Comfort Questionnaire. *Journal of Advanced Nursing*, 75(1), 224–233. <https://doi.org/10.1111/jan.13867>
- Eilam, B. (2012). System thinking and feeding relations: learning with a live ecosystem model. *Instructional Science*, 40(2), 213–239. <https://doi.org/10.1007/s11251-011-9175-4>
- Ellis, G. W., Rudnitsky, A., & Silverstein, B. (2004). Using concept maps to enhance understanding in engineering education. *International Journal of Engineering Education*, 20, 1012–1021.
- 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>
- Erdogan, Y. (2009). Paper-based and computer-based concept mappings: The effects on computer achievement, computer anxiety and computer attitude. *British Journal of Educational Technology*, 40(5), 821–836. <https://doi.org/10.1111/j.1467-8535.2008.00856.x>
- Eshuis, E. H., Vrugte, J., Anjewierden, A., & Jong, T. (2021). Expert examples and prompted reflection in learning with self-generated concept maps. *Journal of Computer Assisted Learning*, Article jcal.12615. Advance online publication. <https://doi.org/10.1111/jcal.12615>
- Feigenspan, K., & Rayder, S. (2017). Systeme und systemisches Denken in der Biologie und im Biologieunterricht. In H. Arndt (Ed.), *FAU Lehren und Lernen: Vol. 2. Systemisches Denken im Fachunterricht* (pp. 139–176). FAU University Press.
- Fischler, H., & Peuckert, J. (2000). Concept Mapping in Forschungszusammenhängen. In H. Fischler & J. Peuckert (Eds.), *Studien zum Physiklernen: Vol. 1. Concept mapping in fachdidaktischen Forschungsprojekten der Physik und Chemie* (pp. 1–23). Logos-Verl.
- Flavell, J. H. (1971). First discussant's comments: What is memory development the development of? *Human Development*, 14(4), 272–278. www.jstor.org/stable/26761846

- Flavell, J. H. (1984). Annahmen zum Begriff Metakognition sowie zur Entwicklung von Metakognition. In F. E. Weinert (Ed.), *Metakognition, Motivation und Lernen* (pp. 23–31). Kohlhammer.
- Flores, F., Tovar, M. E., & Gallegos, L. (2003). Representation of the cell and its processes in high school students: An integrated view. *International Journal of Science Education*, 25(2), 269–286. <https://doi.org/10.1080/09500690210126793>
- Fraune, K. (2013). *Modeling system thinking—assessment, structure validation and development* [Doctoral dissertation, Christian-Albrechts-Universität zu Kiel]. MACAU. https://macau.uni-kiel.de/receive/diss_mods_00015963
- Friedrich, H. F. (1992). Vermittlung von reduktiven Textverarbeitungsstrategien durch Selbstinstruktion. In H. Mandl & H. F. Friedrich (Eds.), *Lern- und Denkstrategien: Analyse und Intervention* (pp. 193–212). Hogrefe.
- Friedrich, H. F., & Mandl, H. (1992). Lern- und Denkstrategien – ein Problemaufriss. In H. Mandl & H. F. Friedrich (Eds.), *Lern- und Denkstrategien: Analyse und Intervention* (pp. 3–54). Hogrefe.
- Friedrich, H. F., & Mandl, H. (2006). Lernstrategien: Zur Strukturierung des Forschungsfeldes. In H. Mandl & H. F. Friedrich (Eds.), *Handbuch Lernstrategien* (pp. 1–23). Hogrefe.
- Garvin-Doxas, K., Klymkowsky, M., & Elrod, S. (2007). Building, using, and maximizing the impact of concept inventories in the biological sciences: Report on a National Science Foundation sponsored conference on the construction of concept inventories in the biological sciences. *CBE-Life Sciences Education*, 6(4), 277–282. <https://doi.org/10.1187/cbe.07-05-0031>
- Goldsmith, T. E., Johnson, P. J., & Acton, W. H. (1991). Assessing structural knowledge. *Journal of Educational Psychology*, 83(1), 88–96. <https://doi.org/10.1037/0022-0663.83.1.88>
- Gouli, E., Gogoulou, A., & Grigoriadou, M. (2003). A coherent and integrated framework using concept maps for various educational assessment functions. *Journal of Information Technology Education: Research*, 2, 215–240. <https://doi.org/10.28945/324>
- Gräsel, C. (1997). *Problemorientiertes Lernen: Strategieanwendung und Gestaltungsmöglichkeiten*. Hogrefe.
- Großschedl, J. (2010). *Einfluss ausgewählter instruktionaler Maßnahmen auf Struktur und Niveau zellbiologischen Wissens* [Doctoral dissertation, Christian-Albrechts-Universität zu Kiel]. Logos.

- Großschedl, J., & Harms, U. (2008). „Similarity Judgments Test“: Ein Verfahren zur Erfassung von Wissensstrukturen. In D. Krüger, A. Upmeyer zu Belzen, T. Riemeier, & K. Niebert (Eds.), *Erkenntnisweg Biologiedidaktik* (Vol. 7, pp. 85–100). Universitätsdruckerei Kassel.
- Großschedl, J., & Harms, U. (2013a). Assessing conceptual knowledge using similarity judgments. *Studies in Educational Evaluation*, 39(2), 71–81. <https://doi.org/10.1016/j.stueduc.2012.10.005>
- Großschedl, J., & Harms, U. (2013b). Effekte metakognitiver Prompts auf den Wissenserwerb beim Concept Mapping und Notizen Erstellen. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 375–395. http://archiv.ipn.uni-kiel.de/zfdn/pdf/19_Grossschedl.pdf
- Großschedl, J., & Harms, U. (2018). Metakognition - Denken aus der Vogelperspektive. In U. Spörhase-Eichmann & W. Ruppert (Eds.), *Biologie Methodik: Handbuch für die Sekundarstufe I und II* (4th ed., pp. 48–52). Cornelsen.
- Großschedl, J., & Tröbst, S. (2017). *Lernstrategietraining im Concept Mapping: Bedeutung für kognitive Belastung, kognitive Prozesse und Lernleistung (unveröffentlichter DFG-Projektantrag)*.
- Großschedl, J., & Tröbst, S. (2018). Biologie lernen durch Concept Mapping: Bedeutung eines Lernstrategietrainings für kognitive Belastung, kognitive Prozesse und Lernleistung – Kurzdarstellung des DFG-Projekts. *Zeitschrift für Didaktik der Biologie (ZDB) - Biologie Lehren und Lernen*, 22(1), 20–30. <https://doi.org/10.4119/UNIBI/zdb-v22-i1-340>
- Grotzer, T. A., Solis, S. L., Tutwiler, M. S., & Cuzzolino, M. P. (2017). A study of students' reasoning about probabilistic causality: Implications for understanding complex systems and for instructional design. *Instructional Science*, 45(1), 25–52. <https://doi.org/10.1007/s11251-016-9389-6>
- Gruber, H., & Stamouli, E. (2020). Intelligenz und Vorwissen. In E. Wild & J. Möller (Eds.), *Lehrbuch. Pädagogische Psychologie* (3rd ed., pp. 25–44). Springer. https://doi.org/10.1007/978-3-662-61403-7_2
- Gruber, T. (2018). *Gedächtnis* (2. Aufl.). *Basiswissen Psychologie*. Springer.
- Größ-Niehaus, T. (2010). *Zum Verständnis des Löslichkeitskonzeptes im Chemieunterricht - der Effekt von Methoden progressiver und kollaborativer Reflexion*. Logos.

- Gunstone, R. F. (1981). *Cognitive structure and performance after physics instruction: Paper presented at the Annual Meeting of the American Education Research Association (Los Angeles, CA, April 13-17, 1981)*. <https://files.eric.ed.gov/fulltext/ED202703.pdf>
- Hammann, M. (2019). Organisationsebenen biologischer Systeme unterscheiden und vernetzen: Empirische Befunde und Empfehlungen für die Praxis. In J. Groß, M. Hammann, P. Schmiemann, & J. Zabel (Eds.), *Biologiedidaktische Forschung: Erträge für die Praxis* (pp. 77–91). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-58443-9_5
- Hanczakowski, M., Zawadzka, K., Pasek, T., & Higham, P. A. (2013). Calibration of metacognitive judgments: Insights from the underconfidence-with-practice effect. *Journal of Memory and Language*, 69(3), 429–444. <https://doi.org/10.1016/j.jml.2013.05.003>
- Hardin, J., Bertoni, G., & Kleinsmith, L. J. (2015). *Always learning. Beckers Welt der Zelle* (W.-M. Weber, & W. M. Becker, Eds.). Pearson.
- Harris, K. R., Graham, S., Brindle, M., & Sandmel, K. (2009). Metacognition and children's writing. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of Metacognition in Education* (pp. 131–153). Routledge.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology: Vol. 52. Human mental workload* (Vol. 52, pp. 139–183). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Hashem, K., & Mioduser, D. (2013). Learning by Modeling (LbM): Understanding complex systems by articulating structures, behaviors, and functions. *International Journal of Advanced Computer Science and Applications*, 4(4). <https://doi.org/10.14569/IJACSA.2013.040414>
- Hasnain, A., & Bhamani, S. (2014). Exploring perceptions of university students pertaining to grades over knowledge and skills. *Journal of Education and Educational Development*, 1(2), 101–115.
- Hasnor, H. N., Ahmad, Z., & Nordin, N. (2013). The relationship between learning approaches and academic achievement among Intec students, Uitm Shah Alam. *Procedia - Social and Behavioral Sciences*, 90, 178–186. <https://doi.org/10.1016/j.sbspro.2013.07.080>
- Hasselhorn, M. (1992). Metakognition und Lernen. In G. Nold (Ed.), *Tübinger Beiträge zur Linguistik: Vol. 366. Lernbedingungen und Lernstrategien: Welche Rolle spielen kognitive Verstehensstrukturen?* (pp. 35–63). Narr.

- Hasselhorn, M., & Artelt, C. (2018). Metakognition. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (5th ed., pp. 520–525). Beltz.
- Hasselhorn, M., & Gold, A. (2013). *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren* (3rd ed.). *Standards Psychologie*. Kohlhammer.
<http://gbv.ebib.com/patron/FullRecord.aspx?p=1766654>
- Hasselhorn, M., & Gold, A. (2017). *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren* (4th ed.). *Standards Psychologie*. Kohlhammer.
<https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=4902108>
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
<http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10545559>
- Hattie, J., & Clarke, S. (2019). *Visible learning: Feedback*. Routledge.
- Hattie, J., Rogers, H. J., & Swaminathan, H. (2014). The role of meta-analysis in educational research. In A. D. Reid, E. P. Hart, & M. A. Peters (Eds.), *A companion to research in education* (Vol. 36, pp. 197–207). Springer. https://doi.org/10.1007/978-94-007-6809-3_26
- Haugwitz, M. (2009). *Kontextorientiertes Lernen und Concept Mapping im Fach Biologie : eine experimentelle Untersuchung zum Einfluss auf Interesse und Leistung unter Berücksichtigung von Moderationseffekten individueller Voraussetzungen beim kooperativen Lernen* [Doctoral dissertation, Universität Duisburg-Essen]. RIS.
http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-23401/Dissertation_Haugwitz.pdf
- Haugwitz, M., Nesbit, J. C., & Sandmann, A. (2010). Cognitive ability and the instructional efficacy of collaborative concept mapping. *Learning and Individual Differences*, 20(5), 536–543. <https://doi.org/10.1016/j.lindif.2010.04.004>
- Haugwitz, M., & Sandmann, A. (2009). Kooperatives Concept Mapping in Biologie: Effekte auf den Wissenserwerb und die Behaltensleistung. *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 89-107. http://archiv.ipn.uni-kiel.de/zfdn/pdf/15_Haugwitz.pdf
- Hauser, S., Nückles, M., & Renkl, A. (2006). Supporting concept mapping for learning from text. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *The International Conference of the Learning Sciences: Indiana University 2006. Proceedings of ICLS 2006* (Vol. 1, pp. 243–249). International Society of the Learning Sciences.

- Hay, D., Kinchin, I., & Lygo-Baker, S. (2008). Making learning visible: the role of concept mapping in higher education. *Studies in Higher Education*, 33(3), 295–311. <https://doi.org/10.1080/03075070802049251>
- Hegarty-Hazel, E., & Prosser, M. (1991a). Relationship between students' conceptual knowledge and study strategies – part 1: student learning in physics. *International Journal of Science Education*, 13(3), 303–312. <https://doi.org/10.1080/0950069910130308>
- Hegarty-Hazel, E., & Prosser, M. (1991b). Relationship between students' conceptual knowledge and study strategies – part 2: student learning in biology. *International Journal of Science Education*, 13(4), 421–429. <https://doi.org/10.1080/0950069910130405>
- Heinze-Fry, J. A., & Novak, J. D. (1990). Concept mapping brings long-term movement toward meaningful learning. *Science Education*, 74(4), 461–472. <https://doi.org/10.1002/sce.3730740406>
- Helmke, A. (2021). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* (8th ed.). Kallmeyer.
- Hesse, M. (2002). Nur geringes Wissen über Zellbiologie: Eine empirische Studie an Lehramtsstudenten. *Zeitschrift für Didaktik der Biologie*, 11, 21–33.
- Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. *Conceptual and Procedural Knowledge: The Case of Mathematics*, 2, 1–27.
- Hilbert, T., Nückles, M., & Matzel, S. (2008). Concept mapping for learning from text: evidence for a worked-out-map-effect. In G. Kanselaar, V. Jonker, P. A. Kirschner, & F. J. Prins (Eds.), *Proceedings of the Eighth international conference on International conference for the learning sciences - ICLS 2008: Vol. 1. International perspectives in the learning sciences: Creating a learning world* (pp. 358–365). International Society of the Learning Sciences. http://dl.acm.org/ft_gateway.cfm?id=1599855&type=pdf
- Hilbert, T., Nückles, M., Renkl, A., Minarik, C., Reich, A., & Ruhe, K. (2008). Concept Mapping zum Lernen aus Texten. *Zeitschrift für Pädagogische Psychologie*, 22(2), 119–125. <https://doi.org/10.1024/1010-0652.22.2.119>
- Hilbert, T., & Renkl, A. (2005). Individual Differences in Concept Mapping when Learning from Texts. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 27(27), 947–952. <https://escholarship.org/content/qt4081p84k/qt4081p84k.pdf>

- Hilbert, T., & Renkl, A. (2008). Concept mapping as a follow-up strategy to learning from texts: what characterizes good and poor mappers? *Instructional Science*, 36(1), 53–73. <https://doi.org/10.1007/s11251-007-9022-9>
- Hilbert, T., & Renkl, A. (2009). Learning how to use a computer-based concept-mapping tool: Self-explaining examples helps. *Computers in Human Behavior*, 25(2), 267–274. <https://doi.org/10.1016/j.chb.2008.12.006>
- Hilbert, T., Renkl, A., Kessler, S., & Reiss, K. (2008). Learning to prove in geometry: Learning from heuristic examples and how it can be supported. *Learning and Instruction*, 18(1), 54–65. <https://doi.org/10.1016/j.learninstruc.2006.10.008>
- Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding complex systems: Some core challenges. *Journal of the Learning Sciences*, 15(1), 53–61. https://doi.org/10.1207/s15327809jls1501_7
- Höfer, D., & Steffens, U. (2012). „Visible Learning for Teachers–Maximizing impact on learning“: Zusammenfassung der praxisorientierten Konsequenzen aus der Forschungsbilanz von John Hattie „Visible Learning“. https://visiblelearning.de/wp-content/uploads/2013/04/Hattie-2_Veroeff_Zsfa_2012_09_26.pdf
- Holovatch, Y., Kenna, R., & Thurner, S. (2017). Complex systems: Physics beyond physics. *European Journal of Physics*, 38(2), 23002. <https://doi.org/10.1088/1361-6404/aa5a87>
- Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, 77(1), 95–111. <https://doi.org/10.1002/sce.3730770107>
- Huang, K., Chen, C.-H., Wu, W.-S., & Chen, W.-Y. (2015). Interactivity of question prompts and feedback on secondary students’ science knowledge acquisition and cognitive load. *Journal of Educational Technology & Society*, 18(4), 159–171.
- Hübner, S., Nückles, M., & Renkl, A. (2010). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction*, 20(1), 18–29. <https://doi.org/10.1016/j.learninstruc.2008.12.001>
- Hundertmark, S. (2012). *Einblicke in kollaborative Lernprozesse: Eine Fallstudie zur reflektierenden Zusammenarbeit unterstützt durch die Methoden Concept Mapping und Lernbegleitbogen. Studien zum Physik- und Chemielernen: Vol. 140.* Logos.
- Ifenthaler, D., & Hanewald, R. (Eds.). (2014). *Digital knowledge maps in education: Technology-enhanced support for teachers and learners.* Springer. <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=1592676>

- Isaacson, R. M., & Fujita, F. (2006). Metacognitive knowledge monitoring and self-regulated learning: Academic success and reflections on learning. *Journal of Scholarship of Teaching and Learning*, 6(1), 39–55.
- Isaak, R., Hüfner, C., Basten, M., & Wilde, M. (2020). Der Einfluss von Elaborationsstrategien auf die Lernleistung im Biologieunterricht unter Berücksichtigung unterschiedlicher kognitiver Anforderungsbereiche. *Psychologie in Erziehung Und Unterricht*, 67(2), 124–137.
- Jack, G. U. (2013). Concept mapping and guided inquiry as effective techniques for teaching difficult concepts in chemistry: Effect on students' academic achievement. *Journal of Education and Practice*, 4(5), 9–15.
- Jacobs, B. (2002). *Aufgaben stellen und Feedback geben*. Medienzentrum der Philosophischen Fakultät der Universität Saarbrücken. <http://psydok.psycharchives.de/jspui/bitstream/20.500.11780/1024/1/feedback.pdf>
- Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *Journal of the Learning Sciences*, 15(1), 11–34. https://doi.org/10.1207/s15327809jls1501_4
- Jegede, O. J., Alaiyemola, F. F., & Okebukola, P. A. O. (1990). The effect of concept mapping on students' anxiety and achievement in biology. *Journal of Research in Science Teaching*, 27(10), 951–960. <https://doi.org/10.1002/tea.3660271004>
- Jersakova, R., Allen, R. J., Booth, J., Souchay, C., & O'Connor, A. R. (2017). Understanding metacognitive confidence: Insights from judgment-of-learning justifications. *Journal of Memory and Language*, 97, 187–207. <https://doi.org/10.1016/j.jml.2017.08.002>
- Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, 7(2), 75–83. <https://doi.org/10.1111/j.1365-2729.1991.tb00230.x>
- Jonassen, D. H. (2000). *Computers as mindtools for schools: Engaging critical thinking* (2. ed.). Merrill.
- Jonassen, D. H. (2003). Using cognitive tools to represent problems. *Journal of Research on Technology in Education*, 35(3), 362–381. <https://doi.org/10.1080/15391523.2003.10782391>
- Jonassen, D. H., & Marra, R. M. (1994). Concept mapping and other formalisms as mindtools for representing knowledge. *ALT-J*, 2(1), 50–56. <https://doi.org/10.1080/0968776940020107>

- Jonassen, D. H., Yacci, M., & Beissner, K. (1993). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge*. Lawrence Erlbaum. <http://gbv.eblib.com/patron/FullRecord.aspx?p=1099216>
- Jong, T. de, & Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*(2), 105–113. https://doi.org/10.1207/s15326985ep3102_2
- Jüngst, K. L. (1995). Studien zur didaktischen Nutzung von Concept Maps. *Unterrichtswissenschaft*, *23*(3), 229–250. <http://www.pedocs.de/volltexte/2013/8131>; <http://nbn-resolving.de/urn:nbn:de:0111-opus-81315>
- Jüngst, K. L., & Strittmatter, P. (1995). Wissensstrukturdarstellung: Theoretische Ansätze und praktische Relevanz. *Unterrichtswissenschaft*, *23*(3), 194–207. <http://www.pedocs.de/volltexte/2013/8129>; <http://nbn-resolving.de/urn:nbn:de:0111-opus-81297>
- 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>
- Kalyuga, S. (2009). Knowledge elaboration: A cognitive load perspective. *Learning and Instruction*, *19*(5), 402–410. <https://doi.org/10.1016/j.learninstruc.2009.02.003>
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4
- Kapuza, A. (2020). How concept maps with and without a list of concepts differ: The case of statistics. *Education Sciences*, *10*(4), 91. <https://doi.org/10.3390/educsci10040091>
- Karakuyu, Y. (2010). The effect of concept mapping on attitude and achievement in a physics course. *International Journal of Physical Sciences*, *5*(6), 724–737.
- Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General*, *138*(4), 469–486. <https://doi.org/10.1037/a0017341>
- Karpicke, J. D. (2012). Retrieval-based learning. *Current Directions in Psychological Science*, *21*(3), 157–163. <https://doi.org/10.1177/0963721412443552>
- Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (2nd ed., pp. 487–514). Academic Press. <https://doi.org/10.1016/B978-0-12-809324-5.21055-9>
- Karpicke, J. D., & Blunt, J. R. (2011a). Response to Comment on “Retrieval practice produces more learning than elaborative studying with concept mapping”. *Science*, *334*(6055), 453. <https://doi.org/10.1126/science.1204035>

- Karpicke, J. D., & Blunt, J. R. (2011b). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, *331*(6018), 772–775. <https://doi.org/10.1126/science.1199327>
- Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, *17*(4), 471–479. <https://doi.org/10.1080/09658210802647009>
- Karpicke, J. D., & Grimaldi, P. J. (2012). Retrieval-based learning: A perspective for enhancing meaningful learning. *Educational Psychology Review*, *24*(3), 401–418. <https://doi.org/10.1007/s10648-012-9202-2>
- Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, *319*(5865), 966–968. <https://doi.org/10.1126/science.1152408>
- Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language*, *67*(1), 17–29. <https://doi.org/10.1016/j.jml.2012.02.004>
- Kattmann, U. (1993). Das Lernen von Namen, Begriffen und Konzepten – Grundlagen biologischer Terminologie am Beispiel Zellenlehre. *Der Mathematische Und Naturwissenschaftliche Unterricht*, *46*, 275–285.
- Kelemen, W. L. (2000). Metamemory cues and monitoring accuracy: Judging what you know and what you will know. *Journal of Educational Psychology*, *92*(4), 800–810. <https://doi.org/10.1037/0022-0663.92.4.800>
- Kharatmal, M., & Nagarjuna, G. (2010). Introducing rigor in concept maps. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, M. Croitoru, S. Ferré, & D. Lukose (Eds.), *Lecture Notes in Computer Science. Conceptual structures: From information to intelligence* (Vol. 6208, pp. 199–202). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-14197-3_22
- Khodor, J., Halme, D. G., & Walker, G. C. (2004). A hierarchical biology concept framework: A tool for course design. *Cell Biology Education*, *3*(2), 111–121. <https://doi.org/10.1187/cbe.03-10-0014>
- Kinchin, I. M. (2000a). *The active use of concept mapping to promote meaningful learning in biological science* [Doctoral dissertation, University of Surrey]. RIS. <https://www.proquest.com/docview/2299779948/previewPDF/9026B6B46388422APQ/1?accountid=10218>

- Kinchin, I. M. (2000b). Concept mapping in biology. *Journal of Biological Education*, 34(2), 61–68. <https://doi.org/10.1080/00219266.2000.9655687>
- Kinchin, I. M. (2000c). Using concept maps to reveal understanding: A two-tier analysis. *School Science Review*, 81(296), 41–46.
- Kinchin, I. M. (2001). If concept mapping is so helpful to learning biology, why aren't we all doing it? *International Journal of Science Education*, 23(12), 1257–1269. <https://doi.org/10.1080/09500690010025058>
- Kinchin, I. M. (2011). Visualising knowledge structures in biology: discipline, curriculum and student understanding. *Journal of Biological Education*, 45(4), 183–189. <https://doi.org/10.1080/00219266.2011.598178>
- Kinchin, I. M. (2014). Concept Mapping as a learning tool in higher education: A critical analysis of recent reviews. *Journal of Continuing Higher Education*, 62(1), 39–49. <https://doi.org/10.1080/07377363.2014.872011>
- Kinchin, I. M., & Cabot, L. B. (2010). Reconsidering the dimensions of expertise: from linear stages towards dual processing. *London Review of Education*, 8(2), 153–166. <https://doi.org/10.1080/14748460.2010.487334>
- Kinchin, I. M., De-Leij, F. A. A. M., & Hay, D. B. (2005). The evolution of a collaborative concept mapping activity for undergraduate microbiology students. *Journal of Further and Higher Education*, 29(1), 1–14. <https://doi.org/10.1080/03098770500037655>
- Kinchin, I. M., Hay, D. B., & Adams, A. (2000). How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development. *Educational Research*, 42(1), 43–57. <https://doi.org/10.1080/001318800363908>
- King, J. F., Zechmeister, E. B., & Shaughnessy, J. J. (1980). Judgments of knowing: The influence of retrieval practice. *The American Journal of Psychology*, 93(2), 329. <https://doi.org/10.2307/1422236>
- Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, 8, Article 1997, 1–18. <https://doi.org/10.3389/fpsyg.2017.01997>

- Klieme, E., Baumert, J., Köller, O., & Bos, W. (2000). Mathematische und naturwissenschaftliche Grundbildung: Konzeptuelle Grundlagen und die Erfassung und Skalierung von Kompetenzen. In J. Baumert, W. Bos, & R. Lehmann (Eds.), *TIMSS/III Dritte Internationale Mathematik- und Naturwissenschaftsstudie — Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn: Band 1 Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit* (pp. 85–133). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-83411-9_4
- KMK (2004). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss: Beschluss vom 16.12.2004*. München, Neuwied. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Biologie.pdf
- KMK (2020). *Bildungsstandards im Fach Biologie für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.06.2020*. Berlin. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2020/2020_06_18-BildungsstandardsAHR_Biologie.pdf
- Knoll, A. R., Otani, H., Skeel, R. L., & van Horn, K. R. (2017). Learning style, judgements of learning, and learning of verbal and visual information. *British Journal of Psychology (London, England : 1953)*, *108*(3), 544–563. <https://doi.org/10.1111/bjop.12214>
- Konrad, K. (2006). Reflexion in interaktiven Lernumgebungen: Können (meta) kognitive prompts und concept maps reflexive Aktivitäten optimieren? *Psychologie in Erziehung Und Unterricht*, *53*(3), 188–201.
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*(4), 349–370. <https://doi.org/10.1037/0096-3445.126.4.349>
- Koriat, A., Ma'ayan, H., Sheffer, L., & Bjork, R. A. (2006). Exploring a mnemonic debiasing account of the underconfidence-with-practice effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 595–608. <https://doi.org/10.1037/0278-7393.32.3.595>
- Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory*, *17*(5), 493–501. <https://doi.org/10.1080/09658210902832915>
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

- Krause, U.-M., Stark, R., & Mandl, H. (2004). Förderung des computerbasierten Wissenserwerbs durch kooperatives Lernen und eine Feedbackmaßnahme. *Zeitschrift für Pädagogische Psychologie*, 18(2), 125–136. <https://doi.org/10.1024/1010-0652.18.2.125>
- Landmann, M., & Schmitz, B. (Eds.). (2007). *Pädagogische Psychologie. Selbstregulation erfolgreich fördern: Praxisnahe Trainingsprogramme für effektives Lernen*. Kohlhammer.
- Lasker, G. A. (2019). Connecting Systems Thinking and Service Learning in the Chemistry Classroom. *Journal of Chemical Education*, 96(12), 2710–2714. <https://doi.org/10.1021/acs.jchemed.9b00344>
- Lawless, C., Smee, P., & O'Shea, T. (1998). Using concept sorting and concept mapping in business and public administration, and in education: an overview. *Educational Research*, 40(2), 219–235. <https://doi.org/10.1080/0013188980400209>
- Leauby, B. A., Szabat, K. A., & Maas, J. D. (2010). Concept mapping - An empirical study in introductory financial accounting. *Accounting Education*, 19(3), 279–300. <https://doi.org/10.1080/09639280903412334>
- Leopold, C., den Elzen-Rump, V., & Leutner, D. (2006). Selbstreguliertes Lernen aus Sachtexten. In M. Prenzel & L. Allolio-Näcke (Eds.), *Untersuchungen zur Bildungsqualität von Schule: Abschlussbericht des DFG-Schwerpunktprogramms* (pp. 268–288). Waxmann.
- Leopold, C., & Leutner, D. (2002). Der Einsatz von Lernstrategien in einer konkreten Lernsituation bei Schülern unterschiedlicher Jahrgangsstufen. In M. Prenzel & J. Doll (Eds.), *Zeitschrift für Pädagogik: Beiheft 45. Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen*. (pp. 240–258). Beltz.
- Lewis, J., Leach, J., & Wood-Robinson, C. (2000a). Chromosomes: The missing link — young people's understanding of mitosis, meiosis, and fertilisation. *Journal of Biological Education*, 34(4), 189–199. <https://doi.org/10.1080/00219266.2000.9655717>
- Lewis, J., Leach, J., & Wood-Robinson, C. (2000b). What's in a cell? — Young people's understanding of the genetic relationship between cells, within an individual. *Journal of Biological Education*, 34(3), 129–132. <https://doi.org/10.1080/00219266.2000.9655702>
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Eds.), *Springer-Lehrbuch. Pädagogische Psychologie* (2nd ed., pp. 69–105). Springer Berlin Heidelberg.

- Lipowsky, F., & Hess, M. (2019). Warum es manchmal hilfreich sein kann, das Lernen schwerer zu machen: Kognitive Aktivierung und die Kraft des Vergleichens. In K. Schöppe & F. Schulz (Eds.), *KREApplus: Vol. 17. Kreativität & Bildung - nachhaltiges Lernen* (pp. 77–132). kopaed.
- Liu, X., & Hinchey, M. (1996). The internal consistency of a concept mapping scoring scheme and its effect on prediction validity. *International Journal of Science Education*, 18(8), 921–937. <https://doi.org/10.1080/0950069960180805>
- Lockl, K., & Schneider, W. (2017). Deklarativ-metakognitives Wissen. In M. A. Wirtz (Ed.), *Dorsch - Lexikon der Psychologie* (18th ed.). Hogrefe.
- Lumer, J., & Winter, K. (2019). Herausforderungen und Chancen einer sprachsensiblen Textarbeit im Biologieunterricht - ein Lehr-Lern-Konzept. In Y. Danilovich & G. Putjata (Eds.), *Edition Fachdidaktiken. Sprachliche Vielfalt im Unterricht: Fachdidaktische Perspektiven auf Lehre und Forschung im DaZ-Modul* (pp. 44–77). Springer Fachmedien Wiesbaden.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1–40. <https://doi.org/10.1017/S0261444812000365>
- Maker, C. J., Zimmerman, R., Gomez-Arizaga, M. P., Pease, R., & Burke, E. M. (2015). Developing Real-Life Problem Solving. In H. E. Vidergor & C. R. Harris (Eds.), *Applied Practice for Educators of Gifted and Able Learners* (pp. 131–168). SensePublishers. https://doi.org/10.1007/978-94-6300-004-8_8
- Mandl, H., & Fischer, F. (2000). Mapping-Techniken und Begriffsnetze in Lern- und Kooperationsprozessen. In H. Mandl & F. Fischer (Eds.), *Wissen sichtbar machen: Wissensmanagement mit Mapping-Techniken* (pp. 3–12). Hogrefe Verlag für Psychologie.
- Mandl, H., & Friedrich, H. F. (Eds.). (2006). *Handbuch Lernstrategien*. Hogrefe. http://hamburg.ciando.com/shop/book/index.cfm/fuseaction/show_book/bok_id/5634
- Markow, P. G., & Lonning, R. A. (1998). Usefulness of concept maps in college chemistry laboratories: Students' perceptions and effects on achievement. *Journal of Research in Science Teaching*, 35(9), 1015–1029. [https://doi.org/10.1002/\(SICI\)1098-2736\(199811\)35:9<1015::AID-TEA4>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1098-2736(199811)35:9<1015::AID-TEA4>3.0.CO;2-G)
- Martin, B. L., Mintzes, J. J., & Clavijo, I. E. (2000). Restructuring knowledge in biology: Cognitive processes and metacognitive reflections. *International Journal of Science Education*, 22(3), 303–323. <https://doi.org/10.1080/095006900289895>

- Martin, P.-Y., & Nicolaisen, T. (2015a). Einführung und Grundlagen. In P.-Y. Martin & T. Nicolaisen (Eds.), *Lernstrategien fördern: Modelle und Praxisszenarien* (pp. 9–69). Beltz Juventa.
- Martin, P.-Y., & Nicolaisen, T. (Eds.). (2015b). *Lernstrategien fördern: Modelle und Praxisszenarien*. Beltz Juventa.
- Martínez, G., Pérez, Á. L., Suero, M. I., & Pardo, P. J. (2013). The effectiveness of concept maps in teaching physics concepts applied to engineering education: Experimental comparison of the amount of learning achieved with and without concept maps. *Journal of Science Education and Technology*, 22(2), 204–214. <https://doi.org/10.1007/s10956-012-9386-8>
- Marton, F., & Säljö, R. (1997). Approaches to learning. In F. Marton, D. Hounsell, & N. J. Entwistle (Eds.), *The experience of learning: Implications for teaching and studying in higher education* (2nd ed., pp. 39–58). Scottish Academic Press.
- Mayer, R. E. (2002a). Rote versus meaningful learning. *Theory into Practice*, 41(4), 226–232. https://doi.org/10.1207/s15430421tip4104_4
- Mayer, R. E. (2002b). *Teaching for meaningful learning. The promise of educational psychology: Vol. 2*. Merrill.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6
- McCagg, E. C., & Dansereau, D. F. (1991). A Convergent Paradigm for Examining Knowledge Mapping as a Learning Strategy. *The Journal of Educational Research*, 84(6), 317–324. <https://doi.org/10.1080/00220671.1991.9941812>
- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36(4), 475–492. [https://doi.org/10.1002/\(SICI\)1098-2736\(199904\)36:4<475::AID-TEA5>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1098-2736(199904)36:4<475::AID-TEA5>3.0.CO;2-O)
- McCormick, R. (1997). Conceptual and procedural knowledge. *International Journal of Technology and Design Education*, 7(1), 141–159. <https://doi.org/10.1023/A:1008819912213>
- McGaghie, W. C., McCrimmon, D. R., Mitchell, G., & Thompson, J. A. (2004). Concept mapping in pulmonary physiology using pathfinder scaling. *Advances in Health Sciences Education : Theory and Practice*, 9(3), 225–240. <https://doi.org/10.1023/B:AHSE.0000038299.79574.e8>

- Mestre, J. P. (2002). *Transfer of learning: Issues and research agenda: Report of a workshop held at the National Science Foundation: Department of Physics*. National Science Foundation. <https://www.nsf.gov/pubs/2003/nsf03212/nsf03212.pdf>
- Metcalf, J. (2009). Metacognitive Judgments and Control of Study. *Current Directions in Psychological Science*, 18(3), 159–163. <https://doi.org/10.1111/j.1467-8721.2009.01628.x>
- Meyer, H. (2018). *Was ist guter Unterricht?* (13th ed.). Cornelsen.
- Michalak, M., & Müller, B. (2017). Durch Sprache zum systemischen Denken. In H. Arndt (Ed.), *FAU Lehren und Lernen: Vol. 2. Systemisches Denken im Fachunterricht* (pp. 111–138). FAU University Press.
- Mietzel, G. (2017). *Pädagogische Psychologie des Lernens und Lehrens* (9th ed.). Hogrefe. <https://elibrary.hogrefe.de/9783840924576/> <https://doi.org/10.1026/02457-000>
- Miller, N., Cañas, A., & Novak, J. D. (2008). Use of the CmapTools recorder to explore acquisition of skill in concept mapping. In A. J. Cañas, P. Reiska, M. Åhlberg, & J. D. Novak (Eds.), *Concept mapping: Connecting educators* (Vol. 2, pp. 674–681). Tallinn University.
- Mintzes, J. J., Cañas, A., Coffey, J., Gorman, J., Gurley, L., Hoffman, R., McGuire, S. Y., Miller, N., Moon, B., Trifone, J., & Wandersee, J. H. (2011). Comment on “Retrieval practice produces more learning than elaborative studying with concept mapping”. *Science (New York, N.Y.)*, 334(6055), 453; author reply 453. <https://doi.org/10.1126/science.1203698>
- Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (1997). Meaningful learning in science: The human constructivist perspective. In G. D. Phye (Ed.), *Handbook of academic learning : Educational psychology* (pp. 405–447). Academic Press. <https://doi.org/10.1016/B978-012554255-5/50014-4>
- Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (2001). Assessing understanding in biology. *Journal of Biological Education*, 35(3), 118–124. <https://doi.org/10.1080/00219266.2001.9655759>
- Montfort, D., Brown, S., & Pollock, D. (2009). An investigation of students’ conceptual understanding in related sophomore to graduate-level engineering and mechanics courses. *Journal of Engineering Education*, 98(2), 111–129. <https://doi.org/10.1002/j.2168-9830.2009.tb01011.x>
- Moors, A., & de Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, 132(2), 297–326. <https://doi.org/10.1037/0033-2909.132.2.297>

- Moreira, B. F. T., Pinto, T. S. S., Starling, D. S. V., & Jaeger, A. (2019). Retrieval practice in classroom settings: A review of applied research. *Frontiers in Education, 4*, Article 5. <https://doi.org/10.3389/feduc.2019.00005>
- Moreno, R. (2004). Decreasing cognitive load for novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Instructional Science, 32*(1/2), 99–113. <https://doi.org/10.1023/B:TRUC.0000021811.66966.1d>
- Morse, D., & Jutras, F. (2008). Implementing concept-based learning in a large undergraduate classroom. *CBE-Life Sciences Education, 7*(2), 243–253. <https://doi.org/10.1187/cbe.07-09-0071>
- Mory, E. H. (2004). Feedback research revisited. In D. H. Jonassen & M. Driscoll (Eds.), *Handbook of research on educational communications and technology* (2nd ed., pp. 745–783). Lawrence Erlbaum.
- Nawani, J., Rixius, J., & Neuhaus, B. J. (2016). Influence of using challenging tasks in biology classrooms on students' cognitive knowledge structure: an empirical video study. *International Journal of Science Education, 38*(12), 1882–1903. <https://doi.org/10.1080/09500693.2016.1213456>
- Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "Delayed-JOL Effect". *Psychological Science, 2*(4), 267–271. <https://doi.org/10.1111/j.1467-9280.1991.tb00147.x>
- Nerdel, C. (2017). *Grundlagen der Naturwissenschaftsdidaktik*. Springer Spektrum. <https://doi.org/10.1007/978-3-662-53158-7>
- Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413–448. <https://doi.org/10.3102/00346543076003413>
- Nesbit, J. C., & Adesope, O. O. (2013). Concept maps for learning. Learning through visual displays. In G. J. Schraw, M. T. McCrudden, & D. R. Robinson (Eds.), *Current perspectives on cognition, learning, and instruction. Learning through visual displays* (pp. 303–328). Information Age.
- Neubrand, C., Borzikowsky, C., & Harms, U. (2016). Adaptive prompts for learning Evolution with worked examples-Highlighting the students between the "novices" and the "experts" in a classroom. *International Journal of Environmental and Science Education, 11*(14), 6774–6795.
- Novak, J. D. (1977). *A theory of education*. Cornell University Press.

- Novak, J. D. (1990a). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching*, 27(10), 937–949. <https://doi.org/10.1002/tea.3660271003>
- Novak, J. D. (1990b). Concept maps and vee diagrams: Two metacognitive tools to facilitate meaningful learning. *Instructional Science*, 19(1), 29–52.
- Novak, J. D. (1998). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Lawrence Erlbaum Associates.
- Novak, J. D. (2002). Meaningful learning: The essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Science Education*, 86(4), 548–571. <https://doi.org/10.1002/sce.10032>
- Novak, J. D. (2003). The promise of new ideas and new technology for improving teaching and learning. *Cell Biology Education*, 2(2), 122–132. <https://doi.org/10.1187/cbe.02-11-0059>
- Novak, J. D. (2010). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Routledge.
- Novak, J. D., & Cañas, A. J. (2006). The origins of the concept mapping tool and the continuing evolution of the tool. *Information Visualization*, 5(3), 175–184. <https://doi.org/10.1057/palgrave.ivs.9500126>
- Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them*. Florida Institute for Human and Machine Cognition. <http://cmap.ihmc.us/docs/pdf/TheoryUnderlyingConceptMaps.pdf>
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press.
- Novak, J. D., Gowin, D. B., & Johansen, G. T. (1983). The use of concept mapping and knowledge vee mapping with junior high school science students. *Science Education*, 67(5), 625–645. <https://doi.org/10.1002/sce.3730670511>
- Nückles, M., Hübner, S., Dümer, S., & Renkl, A. (2010). Expertise reversal effects in writing-to-learn. *Instructional Science*, 38(3), 237–258. <https://doi.org/10.1007/s11251-009-9106-9>
- Nückles, M., Hübner, S., & Renkl, A. (2008). Short-term versus long-term effects of cognitive and metacognitive prompts in writing-to-learn. In P. A. Kirschner, F. J. Prins, V. Jonker, & G. Kanselaar (Eds.), *Proceedings of the 8th international conference on International conference for the learning sciences - Vol. 2* (pp. 124–131). International Society of the Learning Sciences.

- O'Day, G. M., & Karpicke, J. D. (2020). Comparing and combining retrieval practice and concept mapping. *Journal of Educational Psychology*. Advance online publication. <https://doi.org/10.1037/edu0000486>
- O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review*, *14*(1), 71–86. <https://doi.org/10.1023/A:1013132527007>
- Okebukola, P. A. (1990). Attaining meaningful learning of concepts in genetics and ecology: An examination of the potency of the concept-mapping technique. *Journal of Research in Science Teaching*, *27*(5), 493–504. <https://doi.org/10.1002/tea.3660270508>
- Okebukola, P. A., & Jegede, O. J. (1988). Cognitive preference and learning mode as determinants of meaningful learning through concept mapping. *Science Education*, *72*(4), 489–500. <https://doi.org/10.1002/sce.3730720408>
- Paas, F. G. W. C. (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>
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford psychology series: Vol. 9. Oxford University Press; Clarendon Press.
- Parker, R., Guillemard, L., Goetz, E., & Galarza, A. (1996). Using semantic map tests to assess subject matter comprehension. *Diagnostique*, *22*(1), 39–62. <https://doi.org/10.1177/073724779602200103>
- Pearsall, N. R., Skipper, J. E. J., & Mintzes, J. J. (1997). Knowledge restructuring in the life sciences: A longitudinal study of conceptual change in biology. *Science Education*, *81*(2), 193–215. [https://doi.org/10.1002/\(SICI\)1098-237X\(199704\)81:2<193::AID-SCE5>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1098-237X(199704)81:2<193::AID-SCE5>3.0.CO;2-A)
- Perkins, D. N., & Grotzer, T. A. (2005). Dimensions of causal understanding: The role of complex causal models in students' understanding of science. *Studies in Science Education*, *41*(1), 117–165. <https://doi.org/10.1080/03057260508560216>
- Peters, J. H. (2012). *Angstbewältigung und Erinnerung: Eine funktionale Sicht des Gedächtnisses*. Imprint VS Verlag für Sozialwissenschaften. <http://gbv.ebib.com/patron/FullRecord.aspx?p=994066> <https://doi.org/10.1007/978-3-531-19526-1>
- Peuckert, J. (2000). Schulische Wissensdiagnose mittels Concept Mapping. *Vorträge / Physikertagung, Deutsche Physikalische Gesellschaft, Fachausschuss Didaktik Der Physik, 1999*, 215–220.

- Philipp, M. (2021). Prinzipien zur Förderung hierarchiehoher Prozesse. In M. Philipp (Ed.), *Schreiben lernen, schreibend lernen: Prinzipien des Aufbaus und der Nutzung von Schreibkompetenz* (pp. 61–115). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-33253-2_4
- Pintrich, P. R. (2000). The Role of Goal Orientation in Self-Regulated Learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 451–502). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50043-3>
- Pintrich, P. R., Smith, D., Garcia, T., & McKeachie, W. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. National Center for Research to Improve Postsecondary Teaching and Learning, Ann Arbor, MI. <https://files.eric.ed.gov/fulltext/ED338122.pdf>
- Plate, R. (2010). Assessing individuals' understanding of nonlinear causal structures in complex systems. *System Dynamics Review*, 26(1), 19–33. <https://doi.org/10.1002/sdr.432>
- Pressley, M., Forrest-Pressley, D. L., Elliott-Faust, D., & Miller, G. (1985). Children's use of cognitive strategies, how to teach strategies, and what to do if they can't be taught. In C. J. Brainerd & M. Pressley (Eds.), *Springer Series in Cognitive Development. Cognitive learning and memory in children* (pp. 1–47). Springer New York. https://doi.org/10.1007/978-1-4613-9544-7_1
- Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, 27(1), 91–109. https://doi.org/10.1207/s15326985ep2701_7
- Proctor, R. W., & van Zandt, T. (2018). *Human factors in simple and complex systems* (3rd ed.). CRC Press Taylor & Francis Group.
- Prosser, M. (1987). Is prior knowledge of subject matter more important to the development and use of meaningful learning skills? *Programmed Learning & Educational Technology*, 24(4), 280–285. <https://doi.org/10.1080/0033039870240403>
- Pruett, J. L., & Weigel, E. G. (2020). Concept map assessment reveals short-term community-engaged fieldwork enhances sustainability knowledge. *CBE-Life Sciences Education*, 19(3), ar38. <https://doi.org/10.1187/cbe.20-02-0031>
- Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, 60(4), 437–447.

- Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing* (pp. 227–270). MIT Press.
- Quinn, H. J., Mintzes, J. J., & Laws, R. A. (2003). Successive concept mapping. *Journal of College Science Teaching*, 33(3), 12–16.
- Raia, F. (2005). Students' understanding of complex dynamic systems. *Journal of Geoscience Education*, 53(3), 297–308. <https://doi.org/10.5408/1089-9995-53.3.297>
- Raia, F. (2008). Causality in complex dynamic systems: A challenge in earth systems science education. *Journal of Geoscience Education*, 56(1), 81–94. <https://doi.org/10.5408/1089-9995-56.1.81>
- Reader, W., & Hammond, N. (1994). Computer-based tools to support learning from hypertext: Concept mapping tools and beyond. *Computers & Education*, 22, 99–106. [https://doi.org/10.1016/0360-1315\(94\)90078-7](https://doi.org/10.1016/0360-1315(94)90078-7)
- Reiner, M., & Eilam, B. (2001). Conceptual classroom environment - a system view of learning. *International Journal of Science Education*, 23(6), 551–568. <https://doi.org/10.1080/095006901300172458>
- Renkl, A. (2015). Wissenserwerb. In E. Wild & J. Möller (Eds.), *Springer-Lehrbuch. Pädagogische Psychologie* (2nd ed., pp. 3–24). Springer Berlin Heidelberg.
- Renkl, A. (2020). Wissenserwerb. In E. Wild & J. Möller (Eds.), *Lehrbuch. Pädagogische Psychologie* (3rd ed., pp. 3–24). Springer. https://doi.org/10.1007/978-3-662-61403-7_1
- Renkl, A., & Nückles, M. (2006). Lernstrategien der externen Visualisierung. In H. Mandl & H. F. Friedrich (Eds.), *Handbuch Lernstrategien* (pp. 135–147). Hogrefe.
- Ritchhart, R., Turner, T., & Hadar, L. (2009). Uncovering students' thinking about thinking using concept maps. *Metacognition and Learning*, 4(2), 145–159. <https://doi.org/10.1007/s11409-009-9040-x>
- Ritter, R., Wehner, A., Lohaus, G., & Krämer, P. (2019). Pre-service teachers' beliefs about inclusive education before and after multi-compared to mono-professional co-teaching: An exploratory study. *Frontiers in Education*, 4, Article 101. <https://doi.org/10.3389/educ.2019.00101>
- Roediger, H. L., & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science*, 1(3), 181–210. <https://doi.org/10.1111/j.1745-6916.2006.00012.x>

- 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>
- Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: vol. 55. Cognition in education* (Vol. 55, pp. 1–36). Elsevier. <https://doi.org/10.1016/B978-0-12-387691-1.00001-6>
- Roessger, K. M., Daley, B. J., & Hafez, D. A. (2018). Effects of teaching concept mapping using practice, feedback, and relational framing. *Learning and Instruction*, *54*, 11–21. <https://doi.org/10.1016/j.learninstruc.2018.01.011>
- Romero, C., Cazorla, M., & Buzón, O. (2017). Meaningful learning using concept maps as a learning strategy. *Journal of Technology and Science Education*, *7*(3), 313–332. <http://files.eric.ed.gov/fulltext/EJ1156111.pdf>
- Roth, W.-M., & Roychoudhury, A. (1993). The concept map as a tool for the collaborative construction of knowledge: A microanalysis of high school physics students. *Journal of Research in Science Teaching*, *30*(5), 503–534. <https://doi.org/10.1002/tea.3660300508>
- Royer, J. M., Cisero, C. A., & Carlo, M. S. (1993). Techniques and procedures for assessing cognitive skills. *Review of Educational Research*, *63*(2), 201–243.
- Ruiz-Gallardo, J.-R., & Reavey, D. (2019). Learning science concepts by teaching peers in a cooperative environment: A longitudinal study of preservice teachers. *Journal of the Learning Sciences*, *28*(1), 73–107. <https://doi.org/10.1080/10508406.2018.1506988>
- Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching*, *38*(2), 260–278. [https://doi.org/10.1002/1098-2736\(200102\)38:2<260::AID-TEA1005>3.0.CO;2-F](https://doi.org/10.1002/1098-2736(200102)38:2<260::AID-TEA1005>3.0.CO;2-F)
- Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, *33*(6), 569–600. [https://doi.org/10.1002/\(SICI\)1098-2736\(199608\)33:6<569::AID-TEA1>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1098-2736(199608)33:6<569::AID-TEA1>3.0.CO;2-M)
- Ruiz-Primo, M. A., & Shavelson, R. (1997). *Concept-map based assessment: On possible sources of sampling variability* (No. 142). Los Angeles. Center for Research on Evaluation, Standards, and Student Testing. <https://files.eric.ed.gov/fulltext/ED422403.pdf>

- Ruiz-Primo, M. A., Shavelson, R. J., Li, M., & Schultz, S. E. (2001). On the validity of cognitive interpretations of scores from alternative concept-mapping techniques. *Educational Assessment*, 7(2), 99–141. https://doi.org/10.1207/S15326977EA0702_2
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860.
- Rye, J. A., & Rubba, P. A. (2002). Scoring concept maps: An expert map-based scheme weighted for relationships. *School Science and Mathematics*, 102(1), 33–44. <https://doi.org/10.1111/j.1949-8594.2002.tb18194.x>
- Ryssel, J., Sommer, S., Fürstenau, B., & Kunath, J. (2008). The effect of different concept-mapping techniques on promoting students' learning processes in the field of business. In A. J. Cañas, P. Reiska, M. Åhlberg, & J. D. Novak (Eds.), *Proceedings of the Third International Conference on Concept Mapping* (pp. 238–241). Vali Press. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.472.7384&rep=rep1&type=pdf>
- Sadava, D. E., Hillis, D. M., & Heller, H. C. (2019). *Purves Biologie* (10th ed.).
- Salmon, D., & Kelly, M. (2014). *Using concept mapping to foster adaptive expertise: Enhancing teacher metacognitive learning to improve student academic performance*. Peter Lang.
- Schmid, R. F., & Telaro, G. (1990). Concept mapping as an instructional strategy for high school biology. *The Journal of Educational Research*, 84(2), 78–85. <https://doi.org/10.1080/00220671.1990.10885996>
- Schneider, W., Schlagmüller, M., & Ennemoser, M. (2007). *Lesegeschwindigkeits- und verständnistest für die Klassen 6-12*. Hogrefe.
- Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, 19(4), 469–508. <https://doi.org/10.1007/s10648-007-9053-4>
- Schreiber, D. A., & Abegg, G. L. (1991). *Scoring student-generated concept maps in introductory college chemistry: Paper presented at the annual meeting of the National Association for the Research in Science Teaching, Lake Geneva, WI*.
- Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review*, 30(2), 431–455. <https://doi.org/10.1007/s10648-017-9403-9>

- Schwendimann, B. (2015). Concept Mapping. In R. F. Gunstone (Ed.), *Springer reference. Encyclopedia of science education* (pp. 198–202). Springer. https://doi.org/10.1007/978-94-007-2150-0_409
- Schwendimann, B. (2021). Concept Mapping. In R. F. Gunstone (Ed.), *Encyclopedia of Science Education* (pp. 1–5). Springer Netherlands. https://doi.org/10.1007/978-94-007-6165-0_409-5
- Serbessa, D. D. (2006). Tension between traditional and modern teaching-learning approaches in Ethiopian primary schools. *Journal of International Cooperation in Education*, 9(1), 123–140.
- Shavelson, R. J. (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology*, 63(3), 225–234. <https://doi.org/10.1037/h0032652>
- Shavelson, R. J., Ruiz-Primo, M. A., & Wiley, E. W. (2005). Windows into the mind. *Higher Education*, 49(4), 413–430. <https://doi.org/10.1007/s10734-004-9448-9>
- Shute, V., & Towle, B. (2003). Adaptive e-learning. *Educational Psychologist*, 38(2), 105–114. https://doi.org/10.1207/S15326985EP3802_5
- Songer, C. J., & Mintzes, J. J. (1994). Understanding cellular respiration: An analysis of conceptual change in college biology. *Journal of Research in Science Teaching*, 31(6), 621–637. <https://doi.org/10.1002/tea.3660310605>
- Stangl, W. (2006). Mnemotechnik. In H. Mandl & H. F. Friedrich (Eds.), *Handbuch Lernstrategien* (pp. 89–100). Hogrefe.
- Stanton, J. D., Sebesta, A. J., & Dunlosky, J. (2021). Fostering metacognition to support student learning and performance. *CBE-Life Sciences Education*, 20(2), fe3. <https://doi.org/10.1187/cbe.20-12-0289>
- Stark, R., Tyroller, M., Krause, U.-M., & Mandl, H. (2008). Effekte einer metakognitiven Promptingmaßnahme beim situierten, beispielbasierten Lernen im Bereich Korrelationsrechnung. *Zeitschrift für Pädagogische Psychologie*, 22(1), 59–71. <https://doi.org/10.1024/1010-0652.22.1.59>
- Stasz, C., Shavelson, R. J., Cox, D. L., & Moore, C. A. (1976). Field independence and the structuring of knowledge in a social studies minicourse. *Journal of Educational Psychology*, 68(5), 550–558. <https://doi.org/10.1037/0022-0663.68.5.550>
- Steiner, G. (2006). Wiederholungsstrategien. In H. Mandl & H. F. Friedrich (Eds.), *Handbuch Lernstrategien* (pp. 101–113). Hogrefe.

- Stracke, I. (2004). *Einsatz computerbasierter concept maps zur Wissensdiagnose in der Chemie: Empirische Untersuchungen am Beispiel des chemischen Gleichgewichts* [Doctoral dissertation, Christian-Albrechts-Universität zu Kiel]. *Internationale Hochschulschriften: Vol. 428*. Waxmann.
- Streblov, L., & Schiefele, U. (2006). Lernstrategien im Studium. In H. Mandl & H. F. Friedrich (Eds.), *Handbuch Lernstrategien* (pp. 352–364). Hogrefe.
- Sturm, A., & Weder, M. (2016). *Schreibkompetenz, Schreibmotivation, Schreibförderung. Grundlagen und Modelle zum Schreiben als soziale Praxis*. Klett Kallmeyer.
- Sumfleth, E., Neuroth, J., & Leutner, D. (2010). Concept Mapping - eine Lernstrategie muss man lernen. Concept Mapping - Learning Strategy is Something You Must Learn. *CHEMKON*, 17(2), 66–70. <https://doi.org/10.1002/ckon.201010114>
- Suskie, L. (2018). *Assessing student learning: A common sense guide* (3rd ed.). John Wiley & Sons. <https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=5215462>
- Suwono, H., Prasetyo, T. I., Lestari, U., Lukiati, B., Fachrunnisa, R., Kusairi, S., Saefi, M., Fauzi, A., & Atho'illah, M. F. (2021). Cell Biology Diagnostic Test (CBD-Test) portrays pre-service teacher misconceptions about biology cell. *Journal of Biological Education*, 55(1), 82–105. <https://doi.org/10.1080/00219266.2019.1643765>
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816819.003>
- Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296. <https://doi.org/10.1023/A:1022193728205>
- Tauber, S. K., Dunlosky, J., Rawson, K. A., Wahlheim, C. N., & Jacoby, L. L. (2013). Self-regulated learning of a natural category: Do people interleave or block exemplars during study? *Psychonomic Bulletin & Review*, 20(2), 356–363. <https://doi.org/10.3758/s13423-012-0319-6>
- Thiede, K. W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review*, 6(4), 662–667. <https://doi.org/10.3758/BF03212976>
- Thiede, K. W., Anderson, M. C. M., & Theriault, D. (2003). 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>

- Tibell, L. A. E., & Rundgren, C.-J. (2010). Educational challenges of molecular life science: Characteristics and implications for education and research. *CBE—Life Sciences Education*, 9(1), 25–33. <https://doi.org/10.1187/cbe.08-09-0055>
- Toppino, T. C., & Cohen, M. S. (2009). The testing effect and the retention interval: Questions and answers. *Experimental Psychology*, 56(4), 252–257. <https://doi.org/10.1027/1618-3169.56.4.252>
- Toppino, T. C., Heslin, K. A., Curley, T. M., Jackiewicz, M. K., Flowers, C. S., & Phelan, H.-A. (2021). Why do learners ignore expected feedback in making metacognitive decisions about retrieval practice? *Memory & Cognition*. Advance online publication. <https://doi.org/10.3758/s13421-021-01171-4>
- Toppino, T. C., LaVan, M. H., & Iaconelli, R. T. (2018). Metacognitive control in self-regulated learning: Conditions affecting the choice of restudying versus retrieval practice. *Memory & Cognition*, 46(7), 1164–1177. <https://doi.org/10.3758/s13421-018-0828-2>
- Tripto, J., Assaraf, O. B., Snapir, Z., & Amit, M. (2017). How is the body’s systemic nature manifested amongst high school biology students? *Instructional Science*, 45(1), 73–98.
- Tsai, C.-C. (2002). Nested epistemologies: science teachers’ beliefs of teaching, learning and science. *International Journal of Science Education*, 24(8), 771–783.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). Academic Press.
- Turns, J., Atman, C. J., & Adams, R. (2000). Concept maps for engineering education: a cognitively motivated tool supporting varied assessment functions. *IEEE Transactions on Education*, 43(2), 164–173. <https://doi.org/10.1109/13.848069>
- Ufer, S., & Neumann, K. (2018). Measuring competencies. In F. Fischer (Ed.), *International handbook of the learning sciences* (pp. 433–443). Routledge.
- Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Reece, J. B. (2019). *Campbell Biologie* (11th ed.). Pearson Studium.
- van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271–296. <https://doi.org/10.1007/s10648-010-9127-6>
- Verhoeff, R. P. (2003). *Towards systems thinking in cell biology education* [Doctoral dissertation, Universiteit Utrecht]. RIS. <https://dspace.library.uu.nl/bitstream/handle/1874/239/full.pdf?sequence=1&isAllowed=y>

- Verhoeff, R. P., Waarlo, A. J., & Boersma, K. T. (2008). Systems modelling and the development of coherent understanding of cell biology. *International Journal of Science Education*, 30(4), 543–568. <https://doi.org/10.1080/09500690701237780>
- Wadouh, J. (2007). *Vernetzung und kumulatives Lernen im Biologieunterricht der Gymnasialklasse 9* [Doctoral dissertation, Universität Duisburg-Essen]. https://duepublico.uni-duisburg-essen.de/servlets/DocumentServlet/Document/20492/Dissertation_komplett_5_bib_end.pdf
- Wadouh, J., Liu, N., Sandmann, A., & Neuhaus, B. J. (2014). The effect of knowledge linking levels in biology lessons upon students' knowledge structure. *International Journal of Science and Mathematics Education*, 12(1), 25–47. <https://doi.org/10.1007/s10763-012-9390-8>
- Walker, J. M. T., & King, P. H. (2003). Concept mapping as a form of student assessment and instruction in the domain of Bioengineering. *Journal of Engineering Education*, 92(2), 167–178. <https://doi.org/10.1002/j.2168-9830.2003.tb00755.x>
- Watson, M. K., Pelkey, J., Noyes, C. R., & Rodgers, M. O. (2016). Assessing conceptual knowledge using three concept map scoring methods. *Journal of Engineering Education*, 105(1), 118–146. <https://doi.org/10.1002/jee.20111>
- Webb, P., & Bolt, G. (1990). Food chain to food web: A natural progression? *Journal of Biological Education*, 24(3), 187–190. <https://doi.org/10.1080/00219266.1990.9655139>
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching: A project of the American Educational Research Association* (3rd ed., pp. 315–327). Macmillan.
- Wernke, S. (2013). *Aufgabenspezifische Erfassung von Lernstrategien mit Fragebögen: Eine empirische Untersuchung mit Kindern im Grundschulalter. Pädagogische Psychologie und Entwicklungspsychologie: Vol. 88. Waxmann.* <https://elibrary.utb.de/doi/book/10.31244/9783830978428>
- Wild, K. P. (2005). Individuelle Lernstrategien von Studierenden: Konsequenzen für die Hochschuldidaktik und die Hochschullehre. *Beiträge Zur Lehrerinnen- Und Lehrerbildung*, 23(2), 191–206.
- Wild, K. P., & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15(4), 185–200. <https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/3182>

- Wolters, C. A., Pintrich, P. R., & Karabenick, S. A. (2005). Assessing academic self-regulated learning. In K. A. Moore & L. H. Lippman (Eds.), *The Search Institute Series on Developmentally Attentive Community and Society. What Do Children Need to Flourish?* (Vol. 3, pp. 251–270). Springer US. https://doi.org/10.1007/0-387-23823-9_16
- Yang, C., Huang, T. S.-T., & Shanks, D. R. (2018). Perceptual fluency affects judgments of learning: The font size effect. *Journal of Memory and Language*, *99*, 99–110. <https://doi.org/10.1016/j.jml.2017.11.005>
- Yang, C., Potts, R., & Shanks, D. R. (2017). Metacognitive unawareness of the errorful generation benefit and its effects on self-regulated learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(7), 1073–1092. <https://doi.org/10.1037/xlm0000363>
- Yoon, S. A. (2008). An evolutionary approach to harnessing complex systems thinking in the science and technology classroom. *International Journal of Science Education*, *30*(1), 1–32. <https://doi.org/10.1080/09500690601101672>
- Yoon, S. A., Anderson, E., Koehler-Yom, J., Evans, C., Park, M., Sheldon, J., Schoenfeld, I., Wendel, D., Scheintaub, H., & Klopfer, E. (2017). Teaching about complex systems is no simple matter: Building effective professional development for computer-supported complex systems instruction. *Instructional Science*, *45*(1), 99–121. <https://doi.org/10.1007/s11251-016-9388-7>
- Yu, K.-C., Lin, K.-Y., & Fan, S.-C. (2015). An exploratory study on the application of conceptual knowledge and critical thinking to technological issues. *International Journal of Technology and Design Education*, *25*(3), 339–361. <https://doi.org/10.1007/s10798-014-9289-5>
- Zaromb, F. M., Karpicke, J. D., & Roediger, H. L. (2010). Comprehension as a basis for metacognitive judgments: Effects of effort after meaning on recall and metacognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(2), 552–557. <https://doi.org/10.1037/a0018277>
- Zaromb, F. M., & Roediger, H. L. (2010). The testing effect in free recall is associated with enhanced organizational processes. *Memory & Cognition*, *38*(8), 995–1008. <https://doi.org/10.3758/MC.38.8.995>

Danksagung

An dieser Stelle möchte ich allen Personen danken, die mich auf dem Weg der Promotion unterstützt haben und ohne die dieser Abschluss nicht möglich gewesen wäre.

Mein größter Dank gilt Prof. Dr. Jörg Großschedl, der mir diese Promotion ermöglicht hat. Mit seiner Betreuung hat er mein Interesse an der empirischen Bildungsforschung während des Studiums geweckt und meinen Blick auf die Forschung während der Promotionszeit an der Universität zu Köln erweitert. Für all seine Unterstützung über die Jahre und seine Motivierung, die eigenen Ergebnisse im kleinen wie im großen Kreis zu präsentieren und zu diskutieren, werde ich immer dankbar sein. Die Gespräche über Wissenschaftliches und Fachfremdes mit ihm sowie der Austausch mit der Arbeitsgruppe und anderen Forscher:innen auf Tagungen und Konferenzen haben mir nicht nur sehr viel Freude bereitet, sondern mich auch persönlich sowie beruflich weiterentwickelt.

Prof. Dr. Till Bruckermann danke ich für seine Unterstützung und Betreuung sowie seine Bereitschaft, die Zweitbegutachtung im Rahmen meines Promotionsvorhabens zu übernehmen.

Prof.' Dr.' Ellen Aschermann danke ich für die Betreuung, ihre Hilfe und Unterstützung bei der Studienteilnehmer:innenakquise sowie die netten und lehrreichen Gespräche in ihrem Büro, die seit 2020 pandemiebedingt online stattfinden mussten.

Ebenso danke ich Prof. Dr. Frank Schäbitz, der sich entgegenkommenderweise bereit erklärt hat, den Prüfungsvorsitz in der Disputation zu übernehmen.

Mein Dank geht auch an Dr. Fabian Seredszus für seine Bereitschaft, den Prüfungsbeisitz in der Disputation zu übernehmen. Darüber hinaus danke ich ihm für seine Hilfe in der Zeit und seine poetischen Beiträge auf den Weihnachtstreffen, die ein persönlicher Höhepunkt für mich waren und mir in Erinnerung bleiben werden.

Zudem danke ich Dr.' Gaby Schwager-Büschges für ihre Unterstützung auf allen Ebenen während der Promotionsphase. Ihr Engagement für die Graduiertenschule und all ihre Mitglieder habe ich sehr geschätzt und die Gespräche mit ihr sowie die veranstalteten Treffen und Tagungen haben mir die Zeit immens veredelt.

Prof.' Dr.' Daniela Schmeinck danke ich für ihre organisatorische Hilfe und die Bereitstellung digitaler Endgeräte.

Ein Dank geht auch an alle Kolleg:innen der Arbeitsgruppe. Besonders danke ich Virginia Welter, die mich während der Promotion begleitet sowie mit Prof. Dr. Jörg Großschedl, Prof. Dr. Ellen Aschermann und mir die Publikationen gemeistert hat. Die Gespräche mit ihr habe ich sehr geschätzt und bin ihr dankbar für ihre Unterstützung. Jannis Müller, der einen SHK-Orden verdient hat, danke ich für alles, wie auch meinen dienstältesten Promotions-Kolleginnen Sina Lenski und Helena Aptyka. Vielen Dank für die schöne Zeit und die Gespräche über Fachliches und weniger Fachliches.

Ferner danke ich den Argusaugenpaaren von Mirlinda Mustafa, Helena Aptyka, Benjamin Johann und Eike Dinort, die sich bereiterklärt haben, die Arbeit Korrektur lesen.

Mein Dank gilt außerdem all jenen Studierenden, die durch ihre Teilnahme an den Studien einen wesentlichen Teil zu dieser Arbeit beigetragen haben.

Zuletzt möchte ich mich bei meiner Familie und Freund:innen, die mich privat kontinuierlich unterstützt und mit Freude und Motivation versorgt haben, bedanken, insbesondere Alexandra und Christian Werner sowie Eike Dinort. Meinen Eltern danke ich für die selbstlose, geduldige und liebevolle Unterstützung. Meiner Freundin Anna, die mit mir alles durchgemacht hat, danke ich von ganzem Herzen.

Erklärung zur Verfügbarkeit von Primärdaten

Die Primärdaten, welche den vorgestellten Ergebnissen zugrunde liegen, wurden im Rahmen des korrespondierenden DFG-geförderten Forschungsprojektes (GR 47632-1) erhoben und archiviert. Die Archivierung fand durch die projektleitende Person am Institut für Biologiedidaktik der Universität zu Köln statt. Auf entsprechende Anfrage können diese Primärdaten – mit Ausnahme personenbezogener Daten der Versuchspersonen – zwecks Analyse oder Replikation als digitale Datei zur Verfügung gestellt werden. Eine derartige Anfrage ist per E-Mail an den folgenden Ansprechpartner zu richten:

Projekt 1 (Kapitel 4): Prof. Dr. Jörg Großschedl
 E-Mail: j.grossschedl@uni-koeln.de

Projekt 2 (Kapitel 5): Prof. Dr. Jörg Großschedl
 E-Mail: j.grossschedl@uni-koeln.de

Erklärung zur Dissertation

gemäß der Promotionsordnung vom 12. März 2020

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Teilpublikationen:

Becker, L. B., Welter, V. D. E., Großschedl, J. (2021). Effects of strategy training and elaboration vs. retrieval settings on learning of cell biology using concept mapping. *Education Sciences*, 11(9), 530. doi. 10.3390/educsci11090530

Becker, L. B., Welter, V. D. E., Aschermann, E., & Großschedl, J. (2021). Comprehension-oriented learning of cell biology: Do different training conditions affect students' learning success differentially? *Education Sciences*, 11, 438. doi: 10.3390/educsci11080438

Köln, den 03.01.2022



Lukas Bernhard Becker