



Bewährte Lernstrategien in neuem Licht:

Einfluss von Lernstrategien auf kognitive und affektive Aspekte
beim Lernen in der Biologie

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Mirlinda Mustafa, M. Ed.

aus Mihaliq

| | |
|-----------------------|-----------------------------|
| Berichtersteller: | Prof. Dr. Jörg Großschedl |
| (Gutachter) | Prof. Dr. Daniela Schmeinck |
| Datum der Disputation | 06.09.2024 |

Zusammenfassung

Das Lernen in der Biologie ist sowohl in der Schule als auch in der Universität einer Vielzahl von Herausforderungen ausgesetzt. Diese Herausforderungen ergeben sich einerseits aus der Komplexität des biologischen Wissens und andererseits aus den veränderten Lernbedingungen, die aus der fortschreitenden Entwicklung digitaler Lerngelegenheiten und zunehmend heterogener Lerngruppen resultieren. Um diese Herausforderungen bewältigen zu können, sind gezielte Unterstützungsmaßnahmen erforderlich, die durch den Einsatz geeigneter Lernstrategien umgesetzt werden können. Es bleibt jedoch fraglich, welche Lernstrategien das Lernen in der Biologie am effektivsten fördern und den veränderten Lernbedingungen am wirksamsten begegnen.

Um dies zu untersuchen, wurden die Studien I - III initiiert. Im Rahmen dieser Studien wurde die Wirksamkeit der eingesetzten Lernstrategien hinsichtlich kognitiver Aspekte (Wissens- und Kompetenzzuwachs sowie kognitive Belastung) und affektiver Aspekte (Motivation) betrachtet. Ziel war es zu ermitteln, ob die Lernstrategien, die in empirischen Studien erprobt wurden (bewährte Lernstrategien) oder die modifizierten Lernstrategien das Lernen in der Biologie besser unterstützen. Zwei Bedingungen wurden in den Studien jeweils untersucht: Lernende in der Kontrollbedingung lernten mit bewährten Lernstrategien, während in der Experimentalbedingung mit modifizierten Lernstrategien gelernt wurde.

In Studie I wurden im Rahmen einer experimentellen Untersuchung Rückmeldungen, die Lernenden Informationen zu ihrem Lernprozesse gaben, als Lernstrategie untersucht. Hierbei nahmen $N = 62$ Lernende (Biologie-Lehramtsstudierende) teil. Sie arbeiteten innerhalb computerbasierter Lernmodulen mit interaktiven Lernvideos, die forschungsmethodische Inhalte vermittelten. Die Lernenden der Kontrollbedingung erhielten korrektive Rückmeldungen („Die Antwort ist richtig/falsch“), während Lernende der Experimentalbedingung unterstützende Rückmeldungen erhielten, die sie zu inhaltsrelevanten Video-Sequenzen zurückleiteten. Ziel war es, die Wirksamkeit der beiden Formen von Rückmeldungen auf die Forschungskompetenz, das forschungsmethodische Wissen und die intrinsische Motivation zu untersuchen. Die Ergebnisse der Studie zeigten, dass in beiden Bedingungen unabhängig von der Form der Rückmeldung ein Zuwachs im forschungsmethodischen Wissen und in der kognitiven Komponente der Forschungskompetenz vorlag. Hinsichtlich der affektiv-motivationalen Komponente der Forschungskompetenz und der intrinsischen Motivation lag jedoch weder ein Zuwachs noch ein Unterschied zwischen den Formen der Rückmeldung vor.

In Studie II wurde in einer quasi-experimentellen Untersuchung die Lernstrategie der Scaffolds untersucht, welche als Lernhilfe fungieren bis Lernende die Aufgaben selbstständig bewältigen können. Hierbei erhielten $N = 105$ Lernende (Schüler:innen der siebten Klasse) beim Bearbeiten experimentbezogener Aufgaben zum Thema chemische Reaktionen entweder instruktionale Unterstützung durch die Lehrkraft (Kontrollbedingung) oder gestufte Lernhilfen, die Lernenden schrittweise bei der Bearbeitung von Aufgaben verhelfen sollten (Experimentalbedingung). Ziel war es, zu untersuchen, welche Form der Scaffolds den Zuwachs an konzeptuellem Wissen begünstigen. Die Ergebnisse zeigten, dass Lernende der Experimentalbedingung einen signifikant höheren Zuwachs an konzeptuellem Wissen aufweisen im Vergleich zu Lernenden der Kontrollbedingung. Keine Unterschiede konnten bezüglich der Anzahl der verwendeten fachlich inadäquaten Vorstellungen zwischen den Bedingungen festgestellt werden.

In Studie III wurde mittels einer quasi-experimentellen Untersuchung die Lernstrategie der Concept Maps, die als graphische Darstellung von Konzepten und ihren Beziehungen zueinander fungieren, untersucht. $n = 129$ Lernende (Schüler:innen der neunten Klasse) lernten entweder über abrufbasiertes Concept Mapping (Experimentalbedingung), bei dem das Lernmaterial bei der Konstruktion der Concept Maps nicht verfügbar war oder über materialbasiertes Concept Mapping (Kontrollbedingung), bei dem das Lernmaterial während der Konstruktion verfügbar war. Untersucht wurden hierbei die kognitive Belastung, die Concept Map-Qualität, die Elaborationsprozesse und die Lernperformanz. Lernenden beim abrufbasierten Concept Mapping erstellten Concept Maps mit geringerer Concept Map-Qualität, zeigten jedoch eine höhere Lernperformanz und höhere Elaborationsprozesse sowie eine höhere intrinsische kognitive Belastung im Vergleich zu Lernenden des materialbasierten Concept Mappings.

Die gewonnenen Ergebnisse ermöglichen es, die Wirksamkeit der untersuchten Lernstrategien einzuschätzen und indizieren, dass sowohl bewährte als auch modifizierte Lernstrategien, Einfluss auf kognitive Aspekte nehmen. Hinsichtlich affektiver Aspekte konnte jedoch kein Einfluss identifiziert werden, sodass dies stärker in zukünftigen Studien berücksichtigt werden sollte. Insgesamt zeigen die Ergebnisse der Studien I-III, dass Rückmeldungen, Scaffolds und Concept Maps wirkungsvolle Lernstrategien darstellen, die sich sowohl in schul- als auch universitätsbezogenen Lerngelegenheiten einsetzen lassen.

Abstract

Learning in biology is faced with a variety of challenges both at school and at university. These arise on the one hand from the complexity of biological knowledge and on the other hand from the changed learning conditions resulting from the progressive development of digital learning opportunities and increasingly heterogeneous learning groups. In order to deal with these challenges, targeted support measures are required, which can be implemented through the use of suitable learning strategies. However, the question remains, which learning strategies are most effective in promoting learning in biology and which are most effective in dealing with the changing learning conditions.

Studies I-III were initiated to investigate this. In these studies, the effectiveness of the learning strategies used was analysed with regard to cognitive aspects (knowledge and competence growth as well as cognitive load) and affective aspects (motivation). The aim was to determine whether the learning strategies tested in empirical studies (established learning strategies) or the modified learning strategies better support learning in biology. Two conditions were used in each of the studies: Learners in the control condition learnt with established learning strategies, while learners in the experimental condition learnt with modified learning strategies.

In Study I, feedback that provides learners with information about their learning process was investigated as a learning strategy in an experimental study. $N = 62$ learners (biology pre-service teachers) participated. They worked within computer-based learning modules with interactive learning videos that conveyed research methodological content. Learners in the control condition received corrective feedback ("The answer is correct/incorrect"), while learners in the experimental condition received supportive feedback that guided them back to content-relevant video sequences. The aim was to investigate the effectiveness of the two forms of feedback on research competence, research methodological knowledge and intrinsic motivation. The results of the study showed that regardless of the form of feedback, both conditions led to an increase in research methodological knowledge and in the cognitive component of research competence. Regarding the affective-motivational component of research competence and intrinsic motivation, however, neither an increase nor a difference between the forms of feedback could be identified.

In Study II, scaffolds, which act as learning aids until learners can complete the tasks independently, were investigated as a learning strategy as part of a quasi-experimental study. Here, $N = 105$ learners (seventh-grade students) received either instructional support from

the teacher (control condition) or incremental scaffolds designed to help learners work through tasks step by step (experimental condition) when working on experiment-related tasks on the topic of chemical reactions. The aim was to investigate which form of scaffolds promotes the growth of conceptual knowledge. The results indicate that learners in the experimental condition show a significantly higher increase in conceptual knowledge compared to learners in the control condition. No differences were found between the conditions in terms of the number of misconceptions used.

In Study III, the learning strategy of concept maps, which function as a graphical representation of concepts and their relationships to each other, was investigated in a quasi-experimental study. $n = 129$ learners (ninth grade students) learnt either through retrieval-based concept mapping (experimental condition), in which the learning material was not available during the construction of the concept maps, or through study-based concept mapping (control condition), in which the learning material was available during the construction. The cognitive load, concept map quality, elaboration processes and learning performance were analysed. Learners in retrieval-based concept mapping created concept maps with lower concept map quality but showed higher learning performance and higher collaboration processes as well as a higher intrinsic cognitive load compared to learners in study-based concept mapping.

The results obtained make it possible to assess the effectiveness of the learning strategies investigated and indicate that both established and modified learning strategies have an influence on cognitive aspects. Regarding affective aspects, however, no influence could be identified, so that this should be given greater consideration in future studies. Overall, the results of studies I-III indicate that feedback, scaffolds and concept maps are effective learning strategies that can be used in both school and university-related learning environments.

Inhaltsverzeichnis

| | |
|---|-----|
| 1. Einleitung | 1 |
| 2. Theoretischer Hintergrund | 4 |
| 2.1 Wissen und Lernen im Fach Biologie | 4 |
| 2.1.1 Arten von Wissen | 5 |
| 2.1.2 Lernen im Fach Biologie und Herausforderungen | 7 |
| 2.2 Lernen mit Lernstrategien | 9 |
| 2.2.1 Kognitive Aspekte bei der Nutzung von Lernstrategien | 12 |
| 2.2.2 Affektive Aspekte bei der Nutzung von Lernstrategien..... | 17 |
| 3. Fragestellung | 19 |
| 4. Einordnung der Manuskripte | 21 |
| 4.1 Rückmeldesysteme in interaktiven Lernvideos für angehende Biologielehrkräfte: Einfluss auf die Forschungskompetenz, das forschungsmethodische Wissen und die intrinsische Motivation | 22 |
| 4.1.1 Lernstrategie Rückmeldungen: Theoretischer Zugang, Hypothesen, Studiendesign | 22 |
| 4.1.2 Eigenanteil..... | 25 |
| 4.1.3 Manuskript der Studie I | 26 |
| 4.2 Fostering Learning with Incremental Scaffolds During Chemical Experimentation: A Study on Junior High School Students Working in Peer-Groups | 40 |
| 4.2.1 Lernstrategie Scaffolds: Theoretischer Zugang, Hypothesen, Studiendesign | 40 |
| 4.2.2 Eigenanteil..... | 43 |
| 4.2.3 Manuskript der Studie II | 44 |
| 4.3 Concept Mapping – Increased Potential as a Retrieval-based Task | 63 |
| 4.3.1 Lernstrategie Concept Maps: Theoretischer Zugang, Hypothesen, Studiendesign | 63 |
| 4.3.2 Eigenanteil..... | 67 |
| 4.3.3 Manuskript der Studie III..... | 68 |
| 5. Diskussion | 95 |
| 5.1 Inhaltliche Diskussion | 95 |
| 5.1.1 Inhaltliche Diskussion zu kognitiven Aspekten | 97 |
| 5.1.2 Inhaltliche Diskussion zu affektiven Aspekten | 102 |
| 5.2 Methodische Diskussion und Limitationen | 103 |
| 5.3 Implikationen und Ausblick | 105 |
| 6. Fazit | 108 |
| Literaturverzeichnis | 109 |

1. Einleitung

Das Ziel von schulischer und von universitärer Lehre ist es, das Lernen zu ermöglichen und zu fördern. Wie die Lernenden bestmöglich unterstützt werden können, spielt hierbei eine zentrale Rolle (Hattie, 2023; Hattie & Donoghue, 2016). Gezielte Unterstützungsmaßnahmen stellen Lernstrategien dar, welche als methodische Techniken den Lernprozess effizienter gestalten sollen (Misra, 2021; Streblow & Schiefele, 2006; Weinstein & Mayer, 1986). Die zentrale Aufgabe der Bildungsarbeit von Lehrenden ist es, diese Lernstrategien zu vermitteln und ihren Einsatz im Lernprozess zu unterstützen (Hattie & Donoghue, 2016; Misra, 2021). Durch den Einsatz von Lernstrategien wird es Lernenden ermöglicht, nicht nur ihre Leistungen zu verbessern, sondern auch langfristige Fähigkeiten zu entwickeln, von denen sie ihr gesamtes Bildungs- sowie Berufsleben profitieren (Friedrich & Mandl, 2006; Zakrajsek, 2023). Demzufolge gilt die Nutzung von Lernstrategien als Prädiktor für den Erfolg in der Schule (Artelt et al., 2001; Hüfner et al., 2020) und im Studium (Binder et al., 2021).

Welche Lernstrategien das Lernen effektiv unterstützen können, bleibt fraglich, zumal die Lernbedingungen kontinuierlich Veränderungen ausgesetzt sind (Szempruch, 2022). Diese Veränderungen sind vormals geprägt durch die fortschreitende Digitalisierung, die nun auch in den Schulen Einzug genommen hat (Herrmann, 2013; Thorpe, 2022), da unter anderem zunehmend mit iPads und Computern im Unterricht gearbeitet wird (BMBF, 2023). Verstärkt wurden diese Veränderungen vor allem durch die COVID-19-Pandemie, durch welche das Arbeiten mit digitalen Medien unausweichlich war (Vincent-Lancrin, 2022). Die digitalen Veränderungen eröffnen zwar vielfach neue Möglichkeiten, bergen jedoch auch Herausforderungen, da beispielsweise neue Informationsquellen aus dem Internet mit traditionellen Informationsquellen (z. B. Büchern) konkurrieren (Bryce & Blown, 2023). Neben der fortschreitenden Digitalisierung ist auch die zunehmende Heterogenität der Klassen ein wesentlicher Faktor für die veränderten Lernbedingungen (Fend, 2013; Whalley, 2022). Die zunehmende Heterogenität resultiert beispielsweise aus Maßnahmen zur Inklusion, durch welche alle Lernenden, unabhängig von ihren individuellen Fähigkeiten und Einschränkungen, am Lerngeschehen miteinbezogen werden sollen (Bratz & Gaida, 2024). Angesichts der veränderten Lernbedingungen ist es unerlässlich, Lernstrategien einzusetzen, die den Veränderungen begegnen und das Lernen effizient unterstützen.

Besonders in einer Domäne wie der Biologie, welche aufgrund der Komplexität des biologischen Wissens selbst (Harms et al., 2004), viele Herausforderungen birgt, ist die Auswahl

von geeigneten Lernstrategien eine anspruchsvolle Aufgabe der Lehrenden. Um diese Aufgabe zu bewältigen, muss eine Vielzahl von Faktoren und Einflüssen berücksichtigt werden. So sollten Lernstrategien gewählt werden, die das Vorwissen der Lernenden aktivieren (Friedrich & Mandl, 2006) und die begrenzten Kapazitäten des Arbeitsgedächtnisses nicht überfordern und somit auf die kognitive Belastung der Lernenden Rücksicht nehmen (Sweller, 2010; Sweller et al., 2019). Weiterhin sollte die Motivation und das Interesse der Lernenden berücksichtigt werden, die durch den geeigneten Einsatz von Lernstrategien gefördert werden können und so das Lernen unterstützen (Hattie, 2023; Hattie & Donoghue, 2016; Higgins et al., 2021; Misra, 2021; Zoelch et al., 2019).

Angesichts der veränderten Lernbedingungen und den unverändert hohen An- sowie Herausforderungen des fachspezifischen Lernens in der Biologie ist es erforderlich, bewährte Lernstrategien, die bereits in vergangenen empirischen Studien erprobt worden sind, unter aktuellen Lernbedingungen zu untersuchen. Um den aktuellen Lernbedingungen effizienter zu begegnen, können Lernstrategien zudem modifiziert werden. Folglich kann die Frage aufgeworfen werden, ob sich bewährte Lernstrategien oder modifizierte Lernstrategien besser für das Lernen in der Biologie eignen.

Im Rahmen der vorliegenden Arbeit wird diese Frage adressiert, indem interessensbasiert drei bewährte Lernstrategien berücksichtigt werden: 1. Rückmeldungen (englisch: feedback), welche Lernenden, ihre Verbesserungspotentiale aufzeigen und es ihnen ermöglichen, ihren Lernfortschritt einzuschätzen (Hattie & Timperley, 2007; Lipnevich & Panadero, 2021; Mandouit & Hattie, 2023), 2. Scaffolds, welche Lernenden temporär zur Verfügung gestellt werden und mithilfe derer schrittweise zum selbstständigen Arbeiten angeleitet werden kann (Kleinert et al., 2022; van de Pol et al., 2010), 3. Concept Maps, welche die graphische Darstellung des Wissens durch die Darstellung der Konzepte und ihrer Beziehungen zu einander ermöglichen (Blunt & Karpicke, 2014; O'Day & Karpicke, 2020).

Für die Einschätzung der Wirksamkeit verschiedener Formen von Lernstrategien werden im theoretischen Hintergrund die Spezifika des Lernens in der Biologie und die damit einhergehenden Herausforderungen erörtert. Anschließend werden Lernstrategien vorgestellt, welche auf kognitiver und auf affektiver Ebene beim Lernen in der Biologie unterstützen sollen (**Kapitel 2**). Die Forschungsfrage wird in **Kapitel 3** genannt und im Rahmen der in **Kapitel 4** eingebundenen Manuskripte adressiert. In **Kapitel 5** wird sowohl eine inhaltliche als auch eine methodische Diskussion der Ergebnisse gegeben, welche durch einen Ausblick und

durch Implikationen erweitert werden. Das Fazit (**Kapitel 6**) gibt eine abschließende Einschätzung über den Beitrag der vorliegenden Arbeit zum wissenschaftlichen Diskurs.

2. Theoretischer Hintergrund

Mit dem nachfolgenden theoretischen Hintergrund soll eine Grundlage geschaffen werden für die durchgeführten Studien. Hierzu soll zunächst hergeleitet werden, inwiefern Lernstrategien für das Lernen in der Biologie bedeutsam sind und welchen Einfluss sie auf kognitive und affektive Aspekte des Lernens haben.

2.1 Wissen und Lernen im Fach Biologie

Die Studie des *Programme for International Student Assessment* (PISA) löste in Deutschland den sogenannten „PISA-Schock“ aus (Schaffner & Schiefele, 2002; Stanat et al., 2003; Tillmann, 2015). Der Grund hierfür waren die Ergebnisse der deutschen Schüler:innen in den Bereichen Lesekompetenz sowie mathematische und naturwissenschaftliche Grundbildung, welche entgegen der Erwartungen, unter dem Durchschnitt der *Organisation für wirtschaftliche Zusammenarbeit und Entwicklung* (OECD) lagen (Artelt et al., 2001). Als Reaktion auf die Ergebnisse wurden weitreichende bildungspolitische Veränderungen veranlasst im schulischen und universitären Bereich.

Die bildungspolitischen Veränderungen im universitären Bereich zielten darauf ab, die empirische Bildungsforschung zu verstärken, damit Schwierigkeiten im Bildungssystem besser erkannt und geeignetere Lösungen entwickelt und erprobt werden konnten (Tillmann, 2015; Schaffner & Schiefele, 2002). Zu diesem Zeitpunkt verfügten die Universitäten jedoch weder über genügend Fachkompetenzen zur Durchführung der Forschungsprojekte noch über ausreichende finanzielle Mittel. Zur Abhilfe stellte deshalb das Bundesministerium für Bildung und Forschung Mittel bereit (z. B. Rahmenprogramm zur Förderung der empirischen Bildungsforschung, Pähl, 2005; Tillmann, 2015).

Die bildungspolitischen Veränderungen im schulischen Bereich dienten dem Ziel, die Entwicklung eines kompetenzorientierten Unterrichts zu unterstützen, weshalb die Kultusministerkonferenz (KMK) Bildungsstandards definierte (KMK, 2004). Mit den Bildungsstandards wurde festgelegt, welche fachlichen und überfachlichen Kompetenzen die Lernenden bis zu einem bestimmten Zeitpunkt ihrer schulischen Laufbahn erwerben sollen. Kompetenzen bezeichnen hierbei das Wissen und Können im jeweiligen Fach, um Probleme zu lösen (KMK, 2024) sowie die damit verknüpften motivationalen Bestrebungen, die Probleme zu bewältigen (Weinert, 2001). Der KMK-Beschluss wurde auf Landesebene durch Kernlehrpläne aufgegriffen und konkretisiert. Innerhalb dieser Kernlehrpläne wird für das

Bundesland Nordrhein-Westfalen (hier exemplarisch für die Sekundarstufe II Gymnasium/Gesamtschule) für das Fach Biologie unterschieden zwischen den Kompetenzbereichen „Umgang mit Fachwissen“, „Erkenntnisgewinnung“, „Kommunikation“ und „Bewertung“ (MSW NRW, 2013). Die Kompetenzen sind nicht nur an die Kompetenzbereiche geknüpft, sondern auch an die Fachinhalte, die über Inhaltsfelder strukturiert sind. Für das Inhaltsfeld „Ökologie“ sieht der Kernlehrplan beispielsweise die Vermittlung des Energiehaushaltes von Ökosystemen sowie die Wirkung abiotischer und biotischer Faktoren auf Individuen vor (MSW NRW, 2013). Aus der Perspektive der Lernenden stellt das Erlernen dieses Wissens jedoch eine Herausforderung dar, die in der Natur des biologischen Wissens selbst begründet liegt (Harms et al., 2004; siehe Weitzel, 2014). Um die Herausforderung der Lernenden zu antizipieren, ist es erforderlich die Charakteristika biologischen Wissens zu eruieren. Grundlage hierfür bietet die Kenntnis der verschiedenen Arten von Wissen, die durch das Lernen im Fach Biologie adressiert und unterstützt werden soll.

2.1.1 Arten von Wissen

Wissen in der Literatur wird im Konsens als ein komplexes und multidimensionales Konstrukt dargestellt (Binder et al., 2019). Hierbei wird unterschieden zwischen Informationen (Daten, Fakten) und Wissen, wobei Wissen als Ergebnis der Informationsverarbeitung betrachtet wird (Weitzel, 2014). Bereits Bloom (1956) versuchte die verschiedenen Dimensionen beziehungsweise Arten von Wissen zu klassifizieren. Diese Klassifizierung wurde erweitert von Krathwohl (2002), der zwischen „factual knowledge“ (hier: Deklaratives Wissen), „conceptual knowledge“ (Konzeptuelles Wissen), „procedural knowledge“ (Prozedurales Wissen) und „metacognitive knowledge“ (Metakognitives Wissen) unterschied (Adams, 2015; Gogus, 2012).

Das deklarative Wissen (auch *knowing that*) bezieht sich auf das Wissen, das die Lernenden benötigen, um in einem Fachbereich kundig zu sein. Es umfasst die Kenntnis spezifischer Fakten, Prinzipien, Konzepte (Alexander & Judy, 1988) und somit die Kenntnis der Terminologie und das Wissen über bestimmte Details und Elemente (Krathwohl, 2002) in einem Fachbereich. Ein Beispiel für das deklarative Wissen stellt das Verständnis der grundlegenden Konzepte der endothermen und exothermen Reaktionen (bei endothermen Reaktionen wird Wärmeenergie aus der Umgebung aufgenommen; bei exothermen Reaktionen wird Wärmeenergie an die Umgebung abgegeben: Bannwarth et al., 2019) dar.

Während es beim deklarativen Wissen um das Verständnis der einzelnen Konzepte geht, geht es beim konzeptuellen Wissen um Zusammenhänge zwischen den Konzepten (Rittle-Johnson, 2006). Diese Art von Wissen, konzeptuelles Wissen, bezeichnet das Wissen über Wechselbeziehungen zwischen den Grundelementen innerhalb einer größeren Struktur. Es umfasst die Kenntnis von Klassifizierungen und Kategorien, die Kenntnis von Grundsätzen und Verallgemeinerungen sowie die Kenntnis von Theorien, Modellen und Strukturen (Krathwohl, 2002). Der Erwerb von konzeptuellem Wissen kann Lernenden helfen ein tieferes Verständnis von Konzepten zu erlangen (Rittle-Johnson, 2017). Konzeptuelles Wissen kann jedoch auch eine Herausforderung darstellen (Ekinci & Şen, 2020), da es hierbei nicht nur um die Wiedergabe von Konzepten geht, sondern auch um das Verständnis von Konzepten (Montford et al., 2009). Konzeptuelles Wissen über das *Ökosystem* erfordert beispielsweise das Wissen, dass biotische Faktoren lebende Organismen sind, während abiotische Faktoren Umwelteinflüsse sind (Sadava et al., 2019). Um folglich die Wechselwirkungen zwischen abiotischen und biotischen Faktoren innerhalb eines Ökosystems zu verstehen, müssen die damit assoziierten Konzepte (z. B. Ökosystem, biotische Faktoren, Produzenten, Konsumenten, Destruenten, abiotische Faktoren, Temperatur, Nährstoffe usw.) in Zusammenhang gesetzt werden (siehe Lenski, 2023).

Während konzeptuelles Wissen als vielschichtiges Konstrukt zum Wissen über Zusammenhänge der Konzepte gilt, ist mit prozeduralem Wissen, das Wissen über (Forschungs-) Methoden (Braithwaite & Sprague, 2021) und damit das *knowing how* gemeint (Alexander & Judy, 1988). Zum prozeduralen Wissen gehört das Wissen zu fachspezifischen Fertigkeiten, Techniken sowie das Wissen zu Anwendungskriterien für die Methoden (Krathwohl, 2002). Es wird angenommen, dass deklaratives Wissen durch Übung schrittweise in prozedurales Wissen umgewandelt werden kann, ohne dass bewusste Aufmerksamkeitsressourcen erforderlich sind (Abernethy et al., 2007). So soll durch den Einsatz eines Trainings (beispielsweise zur Lernstrategie der Concept Maps), der Erwerb prozeduralen Wissens zu Concept Maps erreicht werden (Becker, 2022).

Nach der Überarbeitung von Krathwohl (2002) wird auch das metakognitive Wissen mitberücksichtigt, welches übergeordnete Prozesse der Kognition behandelt (Zoelch et al., 2019). Metakognitives Wissen bezeichnet das Wissen über Kognition im Allgemeinen sowie das Bewusstsein und Wissen über die eigene Kognition (Krathwohl, 2002). Dazu gehören Wissen zu Strategien und das Wissen über kognitive Aufgaben (Krathwohl, 2002). Metakognitives Wissen ist beispielsweise beim Einsatz von Rückmeldungen erforderlich, da Lernende

hierbei ihre eigenen Leistungen einschätzen (Luo & Lio, 2023) sowie ihren Lernprozess überwachen und reflektieren (Clark & Dumas, 2016).

2.1.2 Lernen im Fach Biologie und Herausforderungen

Das Lernen im Allgemeinen und auch das Lernen in der Biologie im Speziellen ist verschiedenen Einflüssen ausgesetzt und birgt vielfach Herausforderungen. Lernen, aus kognitionspsychologischer Sicht, meint die Veränderung der Wissensstruktur (Mayer, 1998), welche abbildet wie organisiert und kohärent das Wissen im Fach ist (Wadouh et al., 2013). Während Novizen im Fach Biologie über eine Ansammlung isolierter Fakten verfügen, ist das Wissen von Experten kohärent verbunden (siehe de Jong & Ferguson-Hessler, 1996). Hierbei stellt sich die Frage, wie Lernen in der Biologie effizient unterstützt werden kann und welchen Einflüssen es unterliegt.

Innerhalb dieses Diskurses zum Lernen ist Ausubels (1963, 1968) Assimilationstheorie unumgänglich. Für das Lernen spricht er mit seinem bekannten Zitat: „the most important single factor influencing learning is what the student already knows“ (Ausubel, 1968, S. 6) dem bisherigen Wissen beziehungsweise dem Vorwissen die höchste Relevanz zu. Je nachdem wie erfolgreich das erlernte Wissen in das Vorwissen assimiliert wird (Bryce & Blown, 2023), erfolgt *rote learning* (Auswendiglernen) oder *meaningful learning* (bedeutungsvolles Lernen, Ausubel, 1963, 1968). Beim Auswendiglernen wird das Vorwissen mit dem neuen Wissen nur begrenzt verbunden, da es primär um die wortwörtliche Memorierung isolierter Fakten geht. Das bedeutungsvolle Lernen hingegen verläuft über kognitive Elaboration, so dass das neue Wissen in die bestehende, vernetzte Wissensstruktur integriert wird und somit eine stärkere Verbindung zwischen neuem Wissen und Vorwissen ermöglicht wird (Schroeder et al., 2017). Für bedeutungsvolles Lernen wird vorausgesetzt, dass (1.) das Material konzeptuell klar ist und in Verbindung mit dem Vorwissen gebracht werden kann. Zudem müssen (2.) die Lernenden ein gewisses Vorwissen besitzen. Auch müssen Lernende (3.) die Motivation haben, neues Wissen in die bisherige Wissensstruktur zu integrieren und sich somit bewusst dazu entscheiden, bedeutungsvoll zu lernen (Novak & Cañas, 2008).

Die Verknüpfung von neuem Wissen und Vorwissen und damit der Prozess des bedeutungsvollen Lernens ist zentrales Ziel der Bildung und so auch für das Lernen in der Biologie von hoher Relevanz. Hierbei ist eine Herausforderung beim Lernen durch den Umstand gegeben, dass unter aktuellen Lernbedingungen Informationsquellen (wie Internet, Social Media), welche teilweise fachlich inadäquates Wissen vermitteln, mit traditionellen Wissensquellen

(wie Büchern) konkurrieren (Bryce & Blown, 2023). In Konsequenz bringen Lernende nicht ausschließlich fachlich adäquates Wissen mit, sondern auch Vorstellungen aus ihrem Alltag (KMK, 2004), die vom wissenschaftlichen Konsensus abweichen können und das Lernen in der Biologie beeinträchtigen (Weitzel, 2014). Mittels bedeutungsvollen Lernens sollen fachlich inadäquate Vorstellungen (englisch: *misconceptions*) eliminiert werden, indem sie mit wissenschaftlichen Vorstellungen verglichen und anschließend aussortiert werden (Novak & Cañas, 2008). Da fachlich inadäquate Vorstellungen meist zeitlich beständig sind (Smith III et al., 1994) und sogar neben den wissenschaftlichen Vorstellungen koexistieren können (Bryce & Blown, 2023), ist zentrales Anliegen des Faches Biologie, welche aufgrund der Komplexität der biologischen Inhalte anfällig für fachlich inadäquate Vorstellungen ist (Tsui & Treagust, 2009), jene Vorstellungen zu berücksichtigen.

Biologisches Wissen umfasst die Entstehung, Entwicklung, die verschiedenen Arten und Merkmale von Lebewesen sowie ihre Interaktionen untereinander und mit ihrer Umwelt (KMK, 2004; Sadava et al., 2019; Welter et al., 2020). Biologische Konzepte stellen für Lernende meist eine Herausforderung dar, da sie meist zu abstrakt sind und somit nicht mit dem bloßen Auge greifbar und in ihrer Lebenswelt verortbar sind (Wai & Khine, 2020). Herausfordernd ist zudem, dass biologisches Wissen auf verschiedenen Organisationsebenen verortet wird (Sadava et al., 2019). Unterschieden wird dabei zwischen „Biosphäre, Ökosystem, Lebensgemeinschaft (Biozönose), Organismus, Organsysteme und Organe, Gewebe, Zellen Organellen und Moleküle“ (Hammann, 2019, S. 78). Da lebendige Systeme aus unterschiedlichen Elementen bestehen, welche miteinander interagieren (KMK, 2004), bedarf es einer Verknüpfung der Konzepte verschiedener Organisationsebenen, um biologische Phänomene hinreichend erschließen zu können (Hammann, 2019). Dies ist herausfordernd für Lernende, sodass in Konsequenz biologisches Wissen meist unverknüpft erworben wird (Wadouh et al., 2013).

Die Komplexität biologischer Inhalte kann einerseits auf die zahlreichen Verknüpfungen zwischen den Organisationsebenen attribuiert werden und ist darüber hinaus auch auf die Komplexität zurückzuführen, welche aus der interdisziplinären Vernetzung des Wissens resultiert (siehe KMK, 2004). So nutzen die Fächer Chemie und Physik vergleichbare Konzepte wie das Fach Biologie. Ein Beispiel hierfür bietet das intradisziplinäre Konzept der Energie (Opitz et al., 2017). Es wird angenommen, dass der Erwerb von physikalischem und chemischen Wissen für das umfassende Verständnis der Biologie erforderlich ist (siehe

Blundell, 2020) und Lernenden ein umfangreiches naturwissenschaftliches Grundwissen ermöglicht (KMK, 2004).

Jenes biologisch-naturwissenschaftliches Wissen beinhaltet zudem die Kenntnis von Denk- und Arbeitsweisen der Biologie (MSW NRW, 2013). Für die Biologie als Wissenschaft und damit als akademisches Fach ist die Kenntnis und der systematische Einsatz wissenschaftlicher Methoden (z. B. Experimentieren, Beobachten) unerlässlich (Welter et al., 2020). Wissenschaftliche Methoden, wie das Experimentieren, stellen Lernende jedoch vielfach vor Herausforderungen (Baur, 2023), sodass sie Schwierigkeiten zeigen bei der Hypothesenbildung, beim Planen von Experimenten sowie bei der Analyse und der Bewertung von Daten (Hammann et al., 2006; Kranz et al., 2022). Die Komplexität der wissenschaftlichen Methode des Experimentierens wird auch von Lehrern erkannt, die den Einsatz wissenschaftlicher Methoden im Unterricht eher meiden (Gropengießer, 2013). Für das Verständnis wie wissenschaftliche Erkenntnisse gewonnen werden, ist der Einsatz wissenschaftlicher Methoden jedoch unverzichtbar (KMK, 2004; Mayer, 2013; siehe Stiller et al., 2020). Die Fähigkeit, Wissen und Erkenntnisse selbstständig und methodisch zu erarbeiten, ist eine zentrale berufliche Anforderung an Lehrkräfte (Fichten, 2010; Hofer, 2013). Kenntnisse über Forschungsmethoden sollten daher bereits in der Ausbildung von Lehrkräften vermittelt werden (Fichten, 2010). Die geringe Motivation angehender Lehrkräfte sich mit Forschungsmethoden zu beschäftigen (Earley, 2014) und die Komplexität des forschungsmethodischen Wissens stellen jedoch eine Hürde dar (Balloo, 2019; Riewerts et al., 2018). Beim Lernen in der Biologie werden zwar Kompetenzen und Wissen im Bereich der Forschung sowohl bei jüngeren als auch von fortgeschrittenen Lernenden vorausgesetzt, von diesen jedoch als Herausforderung empfunden.

Insgesamt ist somit für das Lernen in der Biologie neben der Umsetzung curricularer Anforderungen (siehe Kapitel 2.1) auch der Erwerb von verschiedenen Arten von Wissen erforderlich (siehe Kapitel 2.1.1). Biologisches Wissen stellt jedoch aufgrund seiner Natur vielfach eine Herausforderung dar (siehe Kapitel 2.1.2). Um diese An- und Herausforderungen zu antizipieren und das Lernen im Fach Biologie effizient zu gestalten, wird der Einsatz adäquater Lernstrategien erforderlich (siehe Kapitel 2.2).

2.2 Lernen mit Lernstrategien

Der Begriff „Lernstrategie“ ist nicht eindeutig festgelegt. Dies liegt unter anderem daran, dass Forschende, Lernstrategien im Kontext ihrer eigenen Studien definieren (Purpura,

1998) und somit in der Literatur zahlreiche Definitionen miteinander konkurrieren (Streblov & Schiefele, 2006). Hinzu kommt, dass der Begriff der Lernstrategien nicht eindeutig abgegrenzt wird von beispielsweise den Begriffen der Lernstile und Lernorientierungen (Looß, 2007). Diese begriffliche Pluralität sowie Ungenauigkeit erschwert es, den Begriff der Lernstrategien zu definieren. Um jedoch einen einheitlichen Rahmen für die vorliegende Arbeit zu schaffen, ist eine Definition von Lernstrategien unerlässlich.

Für die Definition von Lernstrategien wurden ursprünglich kognitive und affektive Aspekte getrennt betrachtet (siehe Wild & Schiefele, 1994). Eine kombinatorische Betrachtung dieser beiden Aspekte im Zusammenhang mit Lernstrategien gewinnt jedoch zunehmend an Bedeutung. So definieren Friedrich und Mandl (2006), in Anknüpfung an Weinstein und Mayer (1986), Lernstrategien als „jene Verhaltensweisen und Gedanken, die Lernende aktivieren, um ihre Motivation und den Prozess des Wissenserwerbs zu beeinflussen und zu steuern“ (Friedrich & Mandl, 2006, S. 1). Diese Definition dient der vorliegenden Arbeit als Grundlage, da sie den Einfluss von Lernstrategien auf kognitive und affektive Aspekte akzentuiert.

Nebst konkurrierenden Definitionen um den Begriff der Lernstrategien, konkurrieren auch zahlreiche Modelle, welche die Klassifikation von Lernstrategien behandeln (Streblov & Schiefele, 2006). Ein Klassifizierungsmodell, welches von Weinstein und Mayer (1986) entwickelt und von zahlreichen Autoren (beispielsweise Friedrich & Mandl, 2006, Wild & Schiefele, 1994) antizipiert wurde, unterteilt Lernstrategien in ressourcenbezogene, metakognitive und kognitive Lernstrategien.

Ressourcenbezogene Lernstrategien sind Aktivitäten, welche Ressourcen zugänglich machen, die Lernen selbst unterstützen oder das Lernen vor externen Einflüssen schützen (Wild & Schiefele, 1994). Hierbei werden interne und externe Ressourcen unterschieden. Die Bereitstellung interner Ressourcen umfasst die Steuerung der eigenen Anstrengung und des eigenen zeitlichen Aufwandes sowie das Investieren von Konzentration und Aufmerksamkeit. Externe Ressourcen können zur Förderung des Lernerfolgs verwendet werden, indem zusätzliche Literatur konsultiert oder die Möglichkeiten von Arbeitsgruppen einbezogen wird (Schiefele & Wild, 1994).

Metakognitive Lernstrategien werden eingesetzt zur Kontrolle des Lernprozesses. Sie umfassen Prozesse, wie die Planung der Lernschritte, die Überprüfung des eigenen Lernfortschrittes und die Regulierung des eigenen Lernverhaltens abhängig vom Lernfortschritt (Misra, 2021).

Kognitive Lernstrategien umfassen Prozesse, die der Aufnahme, Verarbeitung und Speicherung von Informationen dienen (Winn et al., 2019). Hierbei können drei Komponenten unterschieden werden: (1.) Elaborationsstrategien beziehen sich auf Prozesse, durch die neues Wissen in die bestehende Wissensstruktur integriert wird (z. B. neu gelerntes Material wird mit bekannten Alltagsbeispielen verknüpft, de Boer et al., 2013). (2.) Organisationsstrategien umfassen Prozesse, bei der Informationen in eine leichter zu verarbeitende Form überführt werden (z. B. Lerntext wird mittels Diagramms veranschaulicht, de Boer et al., 2013). (3.) Wiederholungsstrategien bezeichnen Prozesse, die darauf abzielen, einzelne Fakten durch Wiederholung im Langzeitgedächtnis zu verankern (z. B. Wörter oder Textpassage werden wiederholt gelesen, de Boer et al., 2013).

Neben der Unterteilung von kognitiven Lernstrategien in Elaborations-, Organisations- und Wiederholungsstrategien, kann entsprechend der Klassifizierung von Marton und Säljö (1976, 1984) bei kognitiven Lernstrategien zwischen Oberflächen- und Tiefenstrategien unterschieden werden, wobei der Grad der Informationsverarbeitung maßgeblich ist (*surface-level processing*, *deep-level processing*, siehe auch Craik & Lockhart, 1972). Es wird davon ausgegangen, dass Informationen, die mit mehr Tiefe verarbeitet werden, besser behalten und gespeichert werden können. So wird durch Tiefenstrategien (auch: Elaborations- und Organisationsstrategien) das Verstehen eines Sachverhaltes intendiert, indem beispielsweise vorhandenes Wissen in neue Wissensstrukturen integriert wird. Bei Oberflächenstrategien (auch: Wiederholungsstrategien) hingegen werden die Lerninhalte auf einem niedrigeren Niveau der Integration bearbeitet. Die einzelnen Lerninhalte bleiben getrennt voneinander und werden lediglich wörtlich wiedergegeben (Artelt, 1999).

Eine weitere Möglichkeit der Klassifizierung stellt die von Weinstein und Mayer (1986) eingeführten kognitiven Lernstrategien (Elaborations-, Organisations- und Wiederholungsstrategien) als Enkodierungsstrategien, den Abrufstrategien gegenüber. Während Enkodierungsstrategien vorrangig der Speicherung von Informationen dienen, kommen Abrufstrategien während des Erinnerns von Informationen zum Einsatz (Zoelch et al., 2019). Abrufstrategien galten lange Zeit nicht als eigentliche Lernstrategien, sondern wurden lediglich zur Leistungsprüfung eingesetzt (Karpicke & Blunt, 2011). In dieser Arbeit werden Abrufstrategien jedoch als Lernstrategien mitberücksichtigt.

2.2.1 Kognitive Aspekte bei der Nutzung von Lernstrategien

Um kognitive Aspekte bei der Nutzung von Lernstrategien zu erläutern, bedarf es einer thematischen Einführung in den Themenbereich der *Kognition* und damit dem Gedächtnis sowie den darin ablaufenden Prozessen. Das Gedächtnis kann verstanden werden als aktives und veränderbares Verarbeitungssystem (Zimbardo et al., 2008), in welchem drei wesentlichen Prozesse ablaufen, nämlich die Enkodierung, die Speicherung und der Abruf von Informationen (Horstmann & Dreisbach, 2012). Während es bei der Enkodierung um die Umwandlung von Informationen (z. B. von Sinnesreizen) in eine für das menschliche Gedächtnis verfügbare Form geht, bezieht sich die Speicherung auf die dauerhafte Lagerung der Informationen im Gedächtnis und der Abruf auf die erfolgreiche Bereitstellung der abgespeicherten Informationen (Horstmann & Dreisbach, 2012; Zoelch et al., 2019). Innerhalb dieser kognitionspsychologischen Auslegung wird Lernen verstanden als ein Prozess, der die Wissensstruktur im Gedächtnis verändert (Mayer, 1998), indem Informationen in das Gedächtnis aufgenommen und langfristig gespeichert werden (Brand & Markowitsch, 2006). Um ein umfassenderes Verständnis der Lernprozesse und kognitiven Prozessen zu gewährleisten, ist die Veranschaulichung der Struktur und Funktion des menschlichen Gedächtnisses zuträglich.

Einen Versuch dieser Veranschaulichung unternahmen Atkinson und Shiffrin (1968) mit dem Mehrspeicher-Modell. Wie der Titel des Modells bereits indiziert, wird das Gedächtnis als Speicher betrachtet. Drei Komponenten werden unterschieden: das sensorische Gedächtnis, das Kurzzeitgedächtnis und das Langzeitgedächtnis. Zunächst werden Informationen als Umweltreize über die Sinnesorgane im sensorischen Gedächtnis gespeichert (Atkinson & Shiffrin, 1968). Je nach Sinnesmodalität wird im sensorischen Gedächtnis unterschieden zwischen dem ikonischen Gedächtnis, in welchem visuelle Informationen verarbeitet werden und dem echoischen Gedächtnis, in welchem akustische Informationen verarbeitet werden. Informationen werden innerhalb des sensorischen Gedächtnisses für wenige Sekunden (maximal vier Sekunden) aufrechterhalten, bis sie zerfallen und vergessen werden (Gruber, 2018). Ob der Reiz übermittelt wird aus dem sensorischen Gedächtnis in das Kurzzeitgedächtnis, hängt zusammen mit der Relevanz, die dem Reiz (beispielsweise auf emotionaler Ebene) zugeschrieben wird (Brand & Markowitsch, 2006). Erst wenn den Informationen erhöhte Aufmerksamkeit gewidmet wird, beispielsweise durch repetitives Aufsagen (Wiederholungsstrategie, englisch: *rehearsal*), können die Informationen bewusst aufrechterhalten und in das Kurzzeitgedächtnis übermittelt werden (siehe Zoelch et al., 2019). Die

Informationen werden dann im Kurzzeitgedächtnis aufgenommen und dort für wenige Sekunden gehalten und verarbeitet. Das Kurzzeitgedächtnis weist eine begrenzte Kapazität auf und es wird angenommen, dass lediglich sieben (+/- 2) Informationseinheiten dort verarbeitet werden können (Miller, 1956). Werden die Informationen über eine längere Zeit im Kurzzeitgedächtnis gehalten, durch beispielsweise Verknüpfung mit Vorwissen (Elaborationsstrategie), steigt die Wahrscheinlichkeit, dass diese Informationen ins Langzeitgedächtnis getragen und dort für wenige Minuten bis ein Leben lang gespeichert werden (Gruber, 2018). Die Dauer der Speicherung im Langzeitgedächtnis hängt, Annahmen zu Folge, von der Tiefe der Verarbeitung der Information ab (*surface- vs. deep-level processing*: Craik & Lockhart, 1972 [siehe Kapitel 2.2]; Zoelch et al., 2019). Das Langzeitgedächtnis ist gekennzeichnet durch eine unbegrenzte Dauer und Kapazität der Speicherung (Gruber, 2018). Innerhalb des Langzeitgedächtnisses wird unterschieden zwischen dem deklarativem und nicht-deklarativem Gedächtnis. Das deklarative Gedächtnis, in dem verbalisierbare Fakten gespeichert werden, unterteilt sich in episodisches und semantisches Gedächtnis. Während im episodischen Gedächtnis situationsabhängige, autobiographische Ereignisse gespeichert werden (z.B. persönliche Erfahrung: Schulabschluss), wird im semantischen Gedächtnis situationsunabhängiges Wissen, Wissen zu komplexen Beziehungen und Begriffsbedeutungen (z. B. Definition von Ökosystem) gespeichert (siehe Zoelch et al., 2019). Das nicht-deklarative Gedächtnis bezieht sich auf die Fähigkeit, sich an handlungsbezogene Phänomene zu erinnern, die nur unzureichend verbalisiert werden können. Hier werden motorische oder automatisiert ablaufende kognitive Fähigkeiten und damit prozedurales Wissen gespeichert (Zoelch et al., 2019). Mit Bezugnahme auf das obige Beispiel (siehe Kapitel 2.1.1) wird das deklarative Wissen, welches im Rahmen einer theoretischen Einführung in die Lernstrategie der Concept Maps (verbalisiert) erworben wurde, im deklarativen Gedächtnis gespeichert, während die konkrete Anwendung der Fertigkeit (hier: das Erstellen einer Concept Map), im nicht-deklarativem Gedächtnis gespeichert wird. Durch Übung oder Training wird der Übergang vom deklarativen in das prozedurale Wissen möglich, wobei ursprünglich deklaratives und damit verbalisierbares Wissen, in prozedurales und damit nicht-verbalisierbares Wissen überführt wird (siehe Anderson, 2014).

Auf die Frage wie verbalisiertes Wissen im deklarativen Gedächtnis repräsentiert wird, existieren verschiedene Annahmen. Eine Annahme ist, dass die Repräsentation über Propositionen geschieht (Zoelch et al., 2019). Propositionen stellen die kleinsten Wissenseinheiten dar, die als Aussagen eigenständig sind (z. B. Alge ist Produzent, siehe Collins & Quillian, 1969). Durch Vorwissen werden zwischen den Wissenseinheiten Verbindungen hergestellt, die

ihrerseits ein Netzwerk (propositionales Netzwerk bzw. semantisches Netzwerk) bilden, in denen zusammenhängende Aussagen gespeichert sind (siehe Zoelch et al., 2019). Beim Abruf von Informationen werden auf Basis von Hinweisreizen (z. B. Wort: Produzent), gelernte Informationen aus dem Langzeitgedächtnis im Kurzzeitgedächtnis verfügbar gemacht (siehe Carpenter, 2009). Meist erfolgt die Wiedergabe nicht wortwörtlich, sondern entsprechend der Repräsentation im propositionalen Netzwerk (Zoelch et al., 2019).

In dem ursprünglichen Modell von Atkinson und Schiffrin (1968) werden die Beziehungen zwischen den Speichern (sensorisches Gedächtnis, Kurzzeitgedächtnis, Langzeitgedächtnis) nur unzureichend erklärt. Besonders das Kurzzeitgedächtnis, als Schnittstelle zwischen sensorischem Gedächtnis und Langzeitgedächtnis erfordert erhöhte Aufmerksamkeit. Baddeley und Hitch (1974) erweiterten die Annahmen zum Kurzzeitgedächtnis, welches in ihrem Modell *Arbeitsgedächtnis* (englisch: *working memory*) betitelt wird. Im Arbeitsgedächtnis unterscheiden sie zwischen der phonologischen Schleife, in welcher phonologische Informationen gespeichert sind, dem visuell-räumlichen Notizblock, in welchem visuell-räumliche Informationen gespeichert sind und der zentralen Exekutive, welche die übergeordnete Kontrollinstanz darstellt. Phonologische Informationen, das heißt Informationen, die sich auf akustische und sinntragende (semantische) Merkmale der Sprache beziehen, werden umgewandelt in einen phonologischen Code und in der phonologischen Schleife verarbeitet. Visuelle und räumliche Informationen, wie beispielsweise Form und Farbe eines Umweltreizes, werden im visuell-räumlichen Notizblock verarbeitet. Die zentrale Exekutive kontrolliert Enkodier- und Abrufstrategien und vermittelt somit zwischen den Informationen aus sensorischem Gedächtnis und Langzeitgedächtnis (Baddeley & Hitch 2007).

Nach Baddeley und Hitchs (1974) Modell ist das Arbeitsgedächtnis als Schnittstelle zwischen neu eingehenden Informationen und dem Abruf von Wissen aus dem Langzeitgedächtnis von großer Bedeutung, da hier bewusste Denk- und Lernprozesse stattfinden. Neue Informationen werden mit vorhandenem Vorwissen verknüpft oder interpretiert auf der Grundlage des Vorwissens. Für den Lernprozess ist es jedoch unumgänglich auch die begrenzte Kapazität des Arbeitsgedächtnisses zu berücksichtigen.

Die Bedeutung der begrenzten Kapazität des Arbeitsgedächtnisses für das Lernen wird behandelt in der Theorie zum *cognitive load* (Sweller et al., 1998, Sweller, 2010). In der Theorie wird die Ansicht vertreten, dass Lernende beim Lernen kognitive Belastung erfahren. Die Quelle der Belastung kann dabei unterschiedlich sein und so werden drei Quellen der Belastung unterschieden: intrinsische kognitive Belastung (*intrinsic cognitive load*),

extrinsische kognitive Belastung (*extraneous cognitive load*) und lernbezogene kognitive Belastung (*germane cognitive load*).

Die intrinsische kognitive Belastung bezeichnet die Belastung, die durch die inhärente Komplexität der Lerninhalte gegeben ist. Lerninhalte gelten dann als komplex, wenn die einzelnen Lernelemente nicht isoliert voneinander, sondern nur im Zusammenhang zueinander gelernt werden können, sprich wenn die Elementinteraktivität hoch ist (Sweller, 2010). Ein typisches Beispiel, was in diesem Kontext angeführt wird, ist das Lernen von Vokabeln. Vokabellisten weisen eine geringe Elementinteraktivität auf, da jede Vokabel einzeln gelernt werden kann. Der Erwerb des Wissens zum Ökosystem weist jedoch eine hohe Elementinteraktivität auf, da hierbei das Verständnis von Zusammenhängen erforderlich wird (siehe Sweller, 2010). Nebst der Komplexität der Lerninhalte, kann auch die Expertise der Lernenden und damit das Ausmaß des Vorwissens in einem Fachbereich, ihre intrinsische kognitive Belastung beeinflussen (Chen et al., 2017). So wäre ein Lernender, der zum ersten Mal konfrontiert wird mit dem Thema *Ökosystem*, überfordert von der Komplexität und den Wechselbeziehungen zwischen den Komponenten eines Ökosystems und würde somit eine hohe intrinsische kognitive Belastung erfahren. Ein Lernender, der jedoch über eine gewisse Expertise verfügt und damit ein gewisses Verständnis von Ökosystemen zeigt, kann sein Vorwissen aus dem Langzeitgedächtnis in das Arbeitsgedächtnis überführen, wo es weniger kognitive Kapazität beansprucht im Vergleich zu völlig neuen Informationen. Folglich würde dieser Lernende eine geringere intrinsische kognitive Belastung wahrnehmen.

Im Vergleich zur intrinsischen kognitiven Belastung, die durch die Lerninhalte selbst gegeben und meist nicht veränderbar ist, ist die extrinsische kognitive Belastung geknüpft an die Darstellung und Gestaltung der Lerninhalte und somit veränderbar. Eine höhere extrinsische kognitive Belastung wird erfahren, wenn das Lernmaterial redundante Informationen enthält, die nicht zum Lernen beitragen, sondern unnütz kognitive Kapazitäten in Anspruch nehmen (*redundancy effect*, Sweller et al., 1998). Die extrinsische kognitive Belastung kann auch durch Lernmaterial erhöht werden, welches verschiedene Informationsquellen enthält, zwischen denen aufmerksamkeitsstechnisch gewechselt werden muss, um die Lerninhalte zu erfassen (*split attention*, Sweller et al., 2019). Um die extrinsische kognitive Belastung zu reduzieren, ist es wichtig Lernmaterial zu gestalten, in welchem beispielsweise zusammengehörende Informationen integriert dargestellt werden (z. B. Abbildung und betreffender Text nah beieinander platziert) oder nur für das Lernen wichtige Informationen präsentiert werden (Sweller et al., 1998).

Es wird angenommen, dass die extrinsische und intrinsische kognitive Belastung additiv sind, sodass beispielsweise eine hohe Komplexität des Lerninhalts, die für Lernende eine hohe intrinsische kognitive Belastung darstellen, mit einer niedrigeren extrinsisch kognitiven Belastung einhergeht. Folglich kann die kognitive Kapazität nur für das Verständnis des Lerninhalts investiert werden und nicht für das Verständnis der Lerninhaltsdarstellung (Sweller et al., 1998). Wenn die Komplexität der Lerninhalte niedrig ist, ist eine hohe extrinsische Belastung durch ungünstig gestaltete Lernmaterialien nicht schädlich (Sweller et al., 1998; Sweller et al., 2019). Insgesamt wird empfohlen, dass Lehrende durch eine entsprechende Gestaltung des Lernmaterials sicherstellen, dass die extrinsische kognitive Belastung reduziert wird (Sweller et al., 1998).

Unter lernbezogener kognitiver Belastung wird die Kapazität verstanden, die nur genutzt werden kann, wenn die Kapazität im Arbeitsgedächtnis nicht bereits belegt ist durch extrinsische und intrinsische kognitive Belastung (Zander, 2010). Die lernbezogene kognitive Belastung ist keine Belastung im herkömmlichen Sinne, da sie sich nicht negativ auswirkt auf die Lernenden. Vielmehr beschreibt die lernbezogene kognitive Belastung die Anstrengung, die Lernende aufbringen, um Schemata im Langzeitgedächtnis zu konstruieren und sich aktiv auseinanderzusetzen mit dem Lerninhalt und diesen zu verstehen (Sweller, 2010). So kann beispielsweise das Abrufen von Informationen aus dem Langzeitgedächtnis (Abrufstrategie) die lernbezogene kognitive Belastung begünstigen, da Konzepte des Lerninhalts verknüpft und verstanden werden (Hultberg et al., 2018). Die lernbezogene kognitive Belastung hängt von der Bereitschaft und der Motivation der Lernenden ab, kognitive Ressourcen in das Lernen zu investieren (Schnotz & Kürschner, 2007).

Aus den bisherigen Erläuterungen kann geschlussfolgert werden, dass Lernen einen domänenspezifischen, mehrstufigen Prozess beschreibt, der dazu beiträgt, die Wissensstruktur durch Enkodierung, Speicherung und Abruf von Informationen zu verändern. Kognitive Lernprozesse werden grundlegend beeinflusst durch das Vorwissen, die Motivation und maßgeblich durch den Einsatz von Lernstrategien (Binder et al., 2021). Lernstrategien stellen bewusste und intentionale Aktivitäten dar (Alexander et al., 1998), welche im Arbeitsgedächtnis ausgeführt, aber im Langzeitgedächtnis aufbewahrt werden und der Bewältigung von Lernaufgaben dienen. Mit dem Ziel von Bildung effiziente Weisen des Lernens (*learning how to learn*) zu vermitteln, geht einher, Lernende mit effektiven Lernstrategien auszustatten (Hattie & Donoghue, 2016) und damit die verfügbaren kognitiven Ressourcen gezielt auszuschöpfen (Zoelch et al., 2019).

2.2.2 Affektive Aspekte bei der Nutzung von Lernstrategien

Für eine umfängliche Betrachtung der Wirksamkeit von Lernstrategien beim Lernen in der Biologie ist die Betrachtung von lediglich kognitiven Aspekten unzureichend. Denn wie bereits durch die Definition von Friedrich und Mandl (2006) kenntlich gemacht, werden für Lernstrategien nicht nur kognitive sondern auch affektive Aspekte, wie Interesse und Motivation, relevant. Betreffende affektive Aspekte werden im Folgenden theoretisch hergeleitet.

Interesse welches als Grundvoraussetzung für erfolgreiches Lernen gilt (Virtbauer et al., 2019), kann als eine Einstellung oder Neigung zu bestimmten Themen, (Lern-) Objekten oder Handlungen verstanden werden (Grassinger et al., 2019). Stellt dieses Interesse an einem Thema einen einmaligen Zustand bei einer bestimmten Lerngelegenheit dar, ist es *situativ*, wohingegen *individuelles Interesse* als stabiler (meist unveränderlicher) Aspekt der Persönlichkeit verstanden wird (Personen-Gegenstands-Theorie, Krapp, 1992). Wenn Lernende wiederholt *situatives Interesse* an einem Lernobjekt erfahren, entwickeln sie *individuelles Interesse* (Grassinger et al., 2019). Das Erleben von Interesse an einem Gegenstand führt dazu, dass positive Affekte im Umgang mit dem Gegenstand erlebt werden und Lernende das Bedürfnis verspüren, mehr über diesen Gegenstand zu lernen (Grassinger et al., 2019), was die Möglichkeit des anschließenden Übens erhöht (Svicher et al. 2022).

Es wird allgemein davon ausgegangen, dass ein starkes Interesse die Lernmotivation fördert und somit auch zu einem effektiven Lernergebnis führt (Krapp, 1998; Plass & Kalyuga, 2019; Virtbauer et al., 2019). Motivation bezeichnet einen „psychische(n) Prozess, der die Initiierung, Ausrichtung und Aufrechterhaltung, aber auch die Steuerung, Qualität und Bewertung zielgerichteten Handelns beeinflusst“ (Grassinger et al., 2019, S. 208). Wenn die Motivation von äußeren Zuständen abhängt (Belohnung, Bestrafungen), dann ist sie extrinsisch, während Motivation aus innerem Antrieb als intrinsisch bezeichnet wird (Virtbauer et al., 2019). Als ausschlaggebende Einflussfaktoren der Motivation werden in der Selbstbestimmungstheorie die psychologischen Grundbedürfnisse der Autonomie, der Kompetenz und der sozialen Eingebundenheit aufgeführt (Deci & Ryan, 1993). Während sich das Bedürfnis nach Autonomie auf den Willen bezieht, eigenes Verhalten zu steuern und unabhängig Entscheidungen zu treffen, bezeichnet Kompetenz den Wunsch, sich kompetent und effizient beim Bearbeiten von Aufgaben zu erleben. Soziale Eingebundenheit meint das Bedürfnis, Beziehungen zu anderen Menschen aufzubauen und sich mit anderen verbunden zu fühlen (Deci & Ryan, 1993). Für die intrinsische Motivation scheinen jedoch insbesondere

das Bedürfnis nach Autonomie und Kompetenz von Bedeutung zu sein (Ryan & Deci, 2002; Wilde et al., 2009). Weiterhin wird in der *Cognitive Evaluation Theory*, welche die Wirkung extrinsischer Belohnungen auf die intrinsische Motivation behandelt, von der Kontrollwahrnehmung und der Informationsfunktion gesprochen. Die Kontrollwahrnehmung bezeichnet den Umstand, dass die extrinsische Belohnung als kontrollierend wahrgenommen wird und das Gefühl der Autonomie eingeschränkt ist, wodurch sich die intrinsische Motivation verringern kann. Die Informationsfunktion bezeichnet den Umstand, dass extrinsische Belohnung als informierend wahrgenommen wird, wodurch das Gefühl von Kompetenz und schließlich die intrinsische Motivation erhöht werden kann. Für das Erleben intrinsischer Motivation ist somit ausschlaggebend, ob Lernende Interesse und Vergnügen empfinden an der Ausübung einer Aufgabe, ob sie Kompetenz erleben, ob sie Wahlfreiheiten in ihren Entscheidungen haben und damit autonom handeln können und ob sie frei von extrinsischem Druck oder Anspannung bei der Ausübung agieren (Wilde et al., 2009).

Affektive Aspekte, darunter Motivation und Interesse werden als entscheidend eingeschätzt für die Qualität der Lernprozesse und die Leistungen der Lernenden (Binder et al., 2021; Krapp, 1998). Somit ist es für das Lernen beispielsweise erforderlich, die Relevanz des Themas kenntlich zu machen und die Inhalte bedeutsam für Lernende zu gestalten. Dies begünstigt die Motivation der Lernenden und ihre Bereitschaft kognitive Kapazitäten in das Lernen zu investieren (Hattie & Donoghue, 2016). Fehlendes Interesse oder Motivation in der Biologie kann nämlich dazu führen, dass Lernende nicht gewillt sind, sich mit biologischen Inhalten zu beschäftigen und den Aufwand meiden (Husna et al., 2023). Für das Lernen in der Biologie ist folglich die Förderung positiver Affekte fundamental (Wai & Khine, 2020), welches besonders durch die geeignete Implementierung von Lernstrategien begünstigt werden kann.

3. Fragestellung

Vielfach wird in empirischen Studien, darunter Primär- und Metastudien, die Wirksamkeit von verschiedenen Lernstrategien untersucht. In diesen werden vormalig der Zuwachs des Wissens und damit kognitive Aspekte des Lernens angestrebt. Jedoch können auch gezielt affektive Aspekte beeinflusst werden (Wild, 2005). Besonders Rückmeldungen wird ein hohes Potential zugeschrieben, um affektive Aspekte zu beeinflussen, wie auch Hattie (2008) in seiner umfangreichen Metaanalyse akzentuiert. Demnach können Rückmeldungen bei Lernenden zu erhöhten Bemühungen und einer erhöhten Motivation beitragen und damit das Lernen positiv beeinflussen (Hattie & Clarke, 2018). Bei der Betrachtung kognitiver Aspekte kann die Metastudie von Dunlosky et al. (2013), welche von Donoghue und Hattie (2021) repliziert wurde, hinzugezogen werden. In dieser Metastudie werden verschiedene Lernstrategien auf ihre Wirksamkeit geprüft. Die Ergebnisse indizieren, dass besonders Abrufstrategien (*practice testing*) eine hohe Wirksamkeit hinsichtlich der Lernperformanz aufzeigen, während häufig verwendete Lernstrategien, wie Markieren (*underlining/highlighting*) und Wiederholtes Lesen (*re-reading*) eine geringe Wirksamkeit zeigen. Während diese Metastudien eine kleine Auswahl von Lernstrategien, auch im Vergleich zueinander, überprüfen, gibt es zahlreiche Metastudien, die die Wirksamkeit einzelner Lernstrategien bewerten, so beispielsweise zu den Lernstrategien *Rückmeldungen* (z. B. Hattie & Timperley, 2007), zu *Scaffolds* (z. B. Zuo et al., 2023) oder zu *Concept Maps* (z. B. Nesbit & Adescope, 2006; Schroeder & Nesbit, 2018). Angesichts der vorliegenden Metastudien und zahlreichen Primärstudien fungieren zuletzt genannte Lernstrategien (Rückmeldungen, Scaffolds, Concept Maps), welche den Fokus der vorliegenden Arbeit darstellen, als bewährte Lernstrategien. Da jedoch ein stetiger Wandel der Lernbedingungen zu beobachten ist, welche aus dem Einfluss von Faktoren wie der Digitalisierung und zunehmend heterogenen Lerngruppen resultiert (siehe Kapitel 1), besteht das Erfordernis zur (Weiter-) Entwicklung von Lernstrategien. So scheinen bereits bewährte Lernstrategien zwar wertvoll, jedoch nicht hinreichend zu sein, um den veränderten Anforderungen gerecht zu werden. Folglich sollten bewährte Lernstrategien innerhalb aktueller Lerngruppen erprobt und modifiziert werden, um den wandelnden Lernbedingungen effektiver zu begegnen. Zudem ist die Wirksamkeit der Lernstrategien von der Aufgabe und der Domäne abhängig, in welcher sie Einsatz findet (de Boer et al., 2013). Für zukünftige empirische Studien stellt sich somit unweigerlich der Anspruch, Lernstrategien hinsichtlich verschiedener Aufgaben und auch hinsichtlich bestimmter Domänen (hier für die Domäne der Biologie), sowohl in Bezug auf kognitive als auch affektive Aspekte zu

erproben. Durch den Einsatz unterschiedlicher Lernstrategien in Felduntersuchungen mit unterschiedlichen Kontexten (Looß, 2007), kann das Lernen in authentischen Lernumgebungen (Schule, Universität) weitgehend erfasst werden (siehe Unterbrunder, 2007).

Mittels (quasi-) experimenteller Feld-Untersuchungen mit jeweils zwei Bedingungen wird überprüft, welche der zwei jeweils untersuchten Formen einer Lernstrategien (bewährte vs. modifizierte Lernstrategie), Lernende in der Biologie hinsichtlich kognitiver und affektiver Aspekte unterstützt. Folgende übergeordnete Forschungsfrage liegt der vorliegenden Arbeit zugrunde:

Unterscheiden sich Lernende der Biologie hinsichtlich kognitiver und affektiver Aspekte, je nachdem, welche Form einer bestimmten Lernstrategie angewendet wird?

4. Einordnung der Manuskripte

Den Herausforderungen des Lernens in der Biologie sollte durch den angemessenen Einsatz von Lernstrategien begegnet werden. Für diesen Zweck werden drei kognitive Lernstrategien (für eine Übersicht: siehe Tabelle 1) untersucht. Diese Lernstrategien werden entsprechend des gegebenen theoretischen Hintergrundes und den damit einhergehenden Forschungsdesiderata auf kognitive und affektive Aspekte beim Lernen in der Biologie untersucht.

Tabelle 1

Übersicht der Studien – Lernstrategien, abhängige und unabhängige Variablen, Hypothesenkennzeichnungen

| Studie | Lernstrategie | Unabhängige Variable | Abhängige Variable | Hypothese |
|---------------|----------------------|---|--|------------------|
| I | Rückmeldungen | Unterstützende Rückmeldungen | Forschungskompetenz | 1.1 |
| | | vs. | Forschungsmethodisches Wissen | 1.2 |
| | | Korrektive Rückmeldungen | Intrinsische Motivation bei der Bearbeitung der Lernmodule | 1.3 |
| II | Scaffolds | Gestufte Lernhilfen | Konzeptuelles Wissen | 2 |
| | | vs. Instruktionale Unterstützung durch Lehrkraft | | |
| III | Concept Maps | Abrufbasiertes Concept Mapping | Lernperformanz (deklaratives Wissen) | 3.1 |
| | | vs. | Kognitive Belastung | 3.2 |
| | | Materialbasiertes Concept Mapping | Concept Map-Qualität | 3.3 |
| | | | Elaborationsprozesse beim Lernen | 3.4 |

4.1 Rückmeldesysteme in interaktiven Lernvideos für angehende Biologielehrkräfte: Einfluss auf die Forschungskompetenz, das forschungsmethodische Wissen und die intrinsische Motivation

4.1.1 Lernstrategie Rückmeldungen: Theoretischer Zugang, Hypothesen, Studiendesign

Theoretischer Zugang

Rückmeldungen gelten als wirksames Instrument, um Lernende optimal im Lernprozess zu unterstützen (Hattie & Timperley, 2007; Lipnevich & Panadero, 2021). So zeigen Rückmeldungen den Lernenden auf, wie erfolgreich sich ihr Lernprozess entwickelt hat und wo gegebenenfalls Verbesserungsbedarf ist (Hattie & Clarke, 2018; Hattie & Timperley, 2007). Rückmeldungen ermöglichen es, die Spanne zwischen dem, was bereits gelernt wurde und dem, was noch zu lernen ist, aufzuzeigen (Hattie & Timperley, 2007). Auf affektiver Ebene können Rückmeldungen die intrinsische Motivation steigern (Kuklick & Lindner, 2021), indem sie Lernende dazu animieren, sich höhere Lernziele zu setzen (Burgers et al., 2015). Folglich kann auch die wahrgenommene Kompetenz erhöht und die Lernleistung unterstützt werden (Burgers et al., 2015; Kuklick & Lindner, 2021). Obwohl Rückmeldungen für den Lernprozess Vorteile bieten, können sie von Lehrenden nicht nach jeder Lerneinheit, meist aus Gründen der Zeitökonomie, bereitgestellt werden. Digitale Möglichkeiten leisten hier jedoch Abhilfe. Computerbasierte Lernsysteme können zur Unterstützung von individuellen Rückmeldungen zum Lernfortschritt genutzt werden (Hillmayr et al., 2020; Sosa et al., 2011, van der Kleij et al., 2015).

Während der vergangenen Jahre entwickelte sich eine große Auswahl an computerbasierten Lernsystemen (Hillmayr et al., 2020; van der Kleij et al., 2015). Innerhalb dieser Lernsysteme können Videos integriert werden, mit Hilfe derer die zu lernenden Inhalte präsentiert werden (Weidlich & Spannagel, 2014). Beim Lernen mit Videos steuern Lernende ihren eigenen Lernprozess und pausieren so entsprechend ihres Bedarfs das Video, um sich Notizen zu machen oder relevante Sequenzen erneut anzuschauen (Noetel et al., 2021). Durch diese eigenständige Kontrolle können Lernende das Ausmaß an kognitiver Belastung beim Lernen selbst steuern und sind so intrinsisch motivierter, sich mit den Inhalten zu beschäftigen (Noetel et al., 2021). Meist werden die Potentiale von Lernvideos jedoch nicht ausgeschöpft, sodass die kurzzeitig präsentierten Inhalte im Video von den Lernenden kognitiv nur oberflächlich verarbeitet werden (Weidlich & Spannagel, 2014). Um eine tiefere kognitive Verarbeitung zu erreichen, können innerhalb computerbasierter Lernsysteme interaktive

Elemente, wie Fragen zu den Video-Inhalten, in die Videos implementiert werden (Reinhold et al., 2020; Weidlich & Spannagel, 2014). Zusätzlich integrierte Rückmeldungen zu den gegebenen Antworten auf die Fragen könnten insbesondere beim Lernen komplexer Inhalte unterstützen (Hillmayr et al., 2020; van der Kleij et al., 2015). Drei Möglichkeiten Rückmeldungen auf Antworten der Lernenden zu geben, werden in empirischen Untersuchungen kontrastiert:

(1.) Bei Rückmeldungen zu *knowledge of results* (deutsch: korrektive Rückmeldungen) erhalten Lernende lediglich die Information über die korrekte oder nicht-korrekte Beantwortung der Frage (z. B. „Falsche Antwort“ / „Richtige Antwort“).

(2.) Bei Rückmeldungen zu *knowledge of correct response* wird Lernenden die richtige Antwort zur Frage vorgegeben (z. B. „Die richtige Antwort wäre: Antwort B“).

(3.) Bei *elaborated feedback* (hier: unterstützende Rückmeldungen) können verschiedene Formen von weiterführenden Unterstützungsmaßnahmen gegeben werden, damit sich die Lernenden die richtige Antwort auf die Frage selbst erarbeiten können (Taxipulati & Lu, 2021).

Grundsätzlich wird angenommen, dass die Form der eingesetzten Rückmeldung den Lernerfolg (Mertens et al., 2022) und die Motivation (Großmann et al., 2020) beeinflusst. Demnach werden Lernende, die mehr Unterstützung erhalten (beispielsweise mit der 3. Form der Rückmeldung), stärker sowohl hinsichtlich kognitiver als auch affektiver Aspekte unterstützt, als Lernende, die weniger Unterstützung (beispielsweise mit der 1. Form der Rückmeldung) erhalten (siehe Hillmayr et al., 2020; siehe Kuklick & Lindner, 2021; siehe van der Kleij et al., 2015). Die Unterstützung naturwissenschaftlichen Lernens mithilfe von computerbasierten Lernsystemen bietet zwar ein hohes Potential, ist jedoch nicht hinreichend untersucht (Hillmayr et al., 2020). So erfordern die kognitiven und affektiven Aspekte bei der Implementierung verschiedener Formen von Rückmeldungen (Kuklick & Lindner, 2021) bei fortgeschrittenen Lernenden weiterer Untersuchung (Mertens et al., 2022).

Hypothesen

Die nachfolgende Studie adressiert dieses Desiderat und untersucht die Wirksamkeit unterstützender Rückmeldungen im Vergleich zu korrektiven Rückmeldungen hinsichtlich kognitiver und affektiver Aspekte bei weiter fortgeschrittenen Lernenden (Studierenden).

Innerhalb biologiedidaktischer Forschungsmodule sollen forschungsmethodisches Wissen (siehe Kapitel 2.1.2) und weiterführend auch die Forschungskompetenz, also die Fähigkeit Forschung auszuüben (Wulf et al., 2020), vermittelt werden. In ihrer universitären Ausbildung erhalten Studierende im Laufe des Semesters jedoch meist wenige Rückmeldungen zu ihren Leistungen (Streblow & Schiefele, 2006). Da der Erwerb forschungsmethodischen Wissens jedoch komplex ist und für Lernende eine Herausforderung darstellt (Balloo, 2019), kann die Implementierung von Unterstützungsmaßnahmen über Rückmeldungen hilfreich sein.

Im Rahmen der nachfolgenden Studie werden Lernende durch den Einsatz interaktiver Lernvideos mit integrierten Rückmeldungen unterstützt. Bei dieser Studie werden Lernende bei falscher Beantwortung der Frage zurück zur inhaltsrelevanten Video-Sequenz geleitet (unterstützende Rückmeldung) oder sie erhalten eine korrektive Rückmeldung. Folgende Überlegungen liegen der Studie zugrunde:

Es wird angenommen, dass der Einsatz unterschiedlicher Formen von Rückmeldungen innerhalb der Lernvideos, Einfluss auf den Erwerb forschungsmethodischen Wissens und der Forschungskompetenz haben könnte. Unterstützende Rückmeldungen enthalten, im Vergleich zu korrektiven Rückmeldungen, mehr Information zur Beantwortung der implementierten Fragen (siehe Taxipulati & Lu, 2021). So könnten mittels unterstützender Rückmeldungen die kurzzeitig präsentierten Inhalte des Videos (siehe Merkt et al., 2011) tiefer verarbeitet werden. Folgende Hypothesen können formuliert werden:

Hypothese 1.1: Lernende, die mit Hilfe von unterstützenden Rückmeldungen lernen, zeigen einen höheren Zuwachs an Forschungskompetenz als Lernenden, die mit Hilfe von korrektiven Rückmeldungen lernen.

Hypothese 1.2: Lernende, die mit Hilfe von unterstützenden Rückmeldungen lernen, zeigen einen höheren Zuwachs an forschungsmethodischem Wissen als Lernenden, die mit Hilfe von korrektiven Rückmeldungen lernen.

Zudem kann angelehnt an die Selbstbestimmungstheorie (Deci & Ryan, 1993; siehe auch Kapitel 2.2.2), der Einsatz von Rückmeldungen innerhalb interaktiver Lernvideos einerseits das Gefühl von Autonomie und damit die Möglichkeit den Lernprozess selbstregulativ zu gestalten, beeinflussen. Andererseits können Rückmeldungen auch das Gefühl der Kompetenz beeinflussen (Hillmayr et al., 2020). Entsprechend kann angenommen werden, dass auch die intrinsische Motivation, dessen Prädiktoren die Kompetenz- und

Autonomiewahrnehmung darstellen (siehe Kapitel 2.2.2), durch den Einsatz verschiedener Formen von Rückmeldungen beeinflusst wird. Folgende Hypothese kann formuliert werden:

Hypothese 1.3: Lernende, die mit Hilfe von unterstützenden Rückmeldungen lernen, zeigen eine höhere intrinsischen Motivation bei der Bearbeitung der Lernmodule als Lernenden, die mit Hilfe von korrektiven Rückmeldungen lernen.

Studiendesign

In der vorliegenden experimentellen Interventionsstudie lernten $N = 62$ Biologie-Lehramtsstudierende der Universität zu Köln mittels computerbasierten Lernmodulen, welche interaktive Lernvideos mit integrierten Rückmeldungen enthielten. Das Studiendesign beinhaltete drei Phasen. (1. Phase:) In dem Prä-Test wurden zu Semesterbeginn demographische Daten (z. B. Alter, Geschlecht), die Forschungskompetenz und das forschungsmethodische Wissen erhoben. (2. Phase:) In der Interventionsphase lernten die Studierenden forschungsmethodische Inhalte über computerbasierte Lernmodule. Studierende der Experimentalbedingung ($n = 31$) erhielten für die Beantwortung der Fragen innerhalb der interaktiven Lernvideos, unterstützende Rückmeldungen und Studierende der Kontrollbedingung ($n = 31$) erhielten korrektive Rückmeldungen. (3.Phase:) In dem Post-Test wurde die Forschungskompetenz, das forschungsmethodische Wissen und die intrinsische Motivation hinsichtlich der Bearbeitung der Lernmodule erhoben.

4.1.2 Eigenanteil

Studie I der vorliegenden Dissertation wurde von der Autorin geplant und durchgeführt. Zudem führte die Autorin die Datenanalyse und –interpretation durch, erstellte das Manuskript und reichte dieses ein. Bei der Erstellung des Manuskripts wurde sie von Dr. Nadine Großmann unterstützt. Prof. Dr. Großschedl überarbeitete das Manuskript und übernahm die Betreuung.

4.1.3 Manuskript der Studie I

Mirlinda Mustafa, Nadine Großmann & Jörg Großschedl

Institut für Biologiedidaktik, Universität zu Köln, 50931 Köln

Rückmeldesysteme in interaktiven Lernvideos für angehende Biologielehrkräfte: Einfluss auf die Forschungskompetenz, das forschungsmethodische Wissen und die intrinsische Motivation

Zusammenfassung

Forschungsmethodisch ausgerichtete Lehrveranstaltungen im Lehramtsstudium sind darauf ausgelegt, die Entwicklung von Forschungskompetenz und forschungsmethodischem Wissen bei Studierenden zu unterstützen. Die Komplexität der Lerninhalte und die niedrige Motivation der Studierenden machen i. d. R. zusätzliche Unterstützungsmaßnahmen notwendig, wie sie z. B. in videobasierten Lernmodulen durch integrierte Rückmeldesysteme bereitgestellt werden. In einem biologiedidaktischen Forschungsmodul ($N = 62$ Lehramtsstudierende) wurden forschungsmethodische Inhalte aus einer theoretischen Perspektive mittels videobasierter Lernmodule vermittelt, die auf die Videoinhalte bezogene Fragen enthielten. Die Experimentalgruppe erhielt unterstützende Rückmeldungen zur Fragenbeantwortung, die Kontrollgruppe korrektive Rückmeldungen. Die Rückmeldungen adressierten den gesamten Forschungsprozess. Als abhängige Variablen wurden die Forschungskompetenz, das forschungsmethodische Wissen sowie die intrinsische Motivation erhoben. Die Forschungskompetenz wurde in einer kognitiven und einer affektiv-motivationalen Komponente erhoben: Erstere beschreibt die notwendigen Fähigkeiten und Fertigkeiten, um eigenständig zu forschen; Letztere bezieht sich auf das Interesse an Forschung und ihren wahrgenommenen Nutzen. Die Ergebnisse zeigen einen Zuwachs im forschungsmethodischen Wissen und der kognitiven Komponente der Forschungskompetenz unabhängig von der Gruppenzugehörigkeit. In der affektiv-motivationalen Komponente und der intrinsischen Motivation konnten weder zeitliche Veränderungen noch Unterschiede zwischen den Gruppen festgestellt werden. Die Befunde deuten darauf hin, dass Rückmeldungen in computerbasierten Lernsystemen weiter untersucht und Lehramtsstudierende auf die Relevanz von Forschung aufmerksam gemacht werden müssen, um die untersuchten Variablen positiv zu beeinflussen.

Abstract

Research method-oriented courses in initial teacher education are designed to support the development of research competence and research methodological knowledge among students. The complexity of the learning content and the low motivation of students make additional support measures necessary, such as those provided in video-based learning modules through integrated feedback systems. In a research module in biology education ($N = 62$ pre-service teachers), research methodology content was taught from a theoretical perspective using video-based learning modules. The video-based learning modules contained questions related to the video content. The experimental group received supportive feedback for answering the questions, while the control group received corrective feedback. The feedback

addressed the entire research process. Research competence, research methodological knowledge and intrinsic motivation were surveyed as dependent variables. Research competence was assessed in a cognitive and an affective-motivational component: the first describes the skills and abilities required to conduct independent research; the latter refers to the interest in research and its perceived benefits. The results show an increase in research methodological knowledge and the cognitive component of research competence, regardless of group affiliation. Neither temporal changes nor differences between the groups were found in the affective-motivational component and intrinsic motivation. The findings suggest that feedback in computer-based learning systems needs to be investigated further and that pre-service teachers need to be made aware of the relevance of research in order to positively influence the variables investigated.

1. Einleitung

Forschungskompetenz wird im Rahmen der Lehramtsausbildung eine zentrale Bedeutung zugeschrieben (Fichten, 2010) und setzt sich aus einer affektiv-motivationalen und kognitiven Komponente zusammen (Besa et al., 2023; Thiel & Böttcher 2014). Nach Besa et al. (2023) weisen Lehramtsstudierende im Vergleich zu Studierenden anderer Studiengänge Schwächen hinsichtlich der kognitiven Komponente auf. Gleichzeitig schätzen sie die Bedeutung von Forschung, als Aspekt der affektiv-motivationalen Komponente, für ihr späteres Berufsleben geringer ein als Studierende anderer Studiengänge (Besa et al., 2023). Dem steht die Bedeutung von Forschungskompetenz für die spätere professionelle Praxis gegenüber: Forschungskompetenz wird eine entscheidende Rolle zugeschrieben, um die eigene Lehrtätigkeit reflektieren und im Unterricht evidenzbasiert handeln zu können (Fichten, 2010). Die Lehrkraftbildungsstandards erkennen diese besondere Rolle von Forschungskompetenz an (KMK, 2024, S. 23; vgl. auch Thiel & Böttcher, 2014).

Um die Forschungskompetenz von angehenden Biologielehrkräften zu fördern, wurden in dieser Studie videobasierte Lernmodule eingesetzt (vgl. auch Gussen et al., 2023a). Frühere Studien zeigen, dass videobasierte Lernmodule motivationale Vorteile haben und sich günstig auf den Wissenserwerb von Studierenden auswirken können (Noetel et al., 2021). Sie unterstützen raum- und zeitunabhängiges Arbeiten und tragen so zu einer Flexibilisierung der Organisation des Studiums bei. Außerdem ermöglichen sie individuelle inhaltliche Schwerpunktsetzungen entsprechend der Lernbedürfnisse der Studierenden (Noetel et al., 2021). Videobasierte Lernmodule sollten interaktive Elemente in Form von z. B. Fragen enthalten (Weidlich & Spannagel, 2014), die den Lernenden eine Rückmeldung geben und

dadurch zum Lernerfolg beitragen können. Mertens et al. (2022) konnten zeigen, dass hierbei die Art der Rückmeldung von entscheidender Bedeutung ist. Während korrektive Rückmeldungen (richtig/falsch) als wenig effektiv eingeschätzt werden, zeigen Rückmeldungen, die weiterführende Unterstützung leisten, eine hohe Lernwirksamkeit (Mertens et al., 2021; van der Kleij et al., 2015). Zugleich können sich verschiedene Arten von Rückmeldung auf die Motivation und das Interesse der Studierenden auswirken (Reeve, 2015), wobei unterstützende Rückmeldungen die intrinsische Motivation im Vergleich zu korrektiven Rückmeldungen stärker begünstigen (Großmann et al., 2020) und das Interesse an der Beschäftigung mit den Inhalten steigern können (Kuklick & Lindner, 2021).

In der vorliegenden Studie wurden grundlegende forschungsmethodische Inhalte aus einer theoretischen Perspektive mittels videobasierter Lernmodule vermittelt, wobei alle Schritte eines biologiedidaktischen Forschungsprozesses berücksichtigt wurden. In den Lernmodulen wurden zwei Arten von Rückmeldungen (unterstützend/korrektiv) zu den gegebenen Antworten und den zugehörigen forschungsmethodischen Inhalten gegeben, und ihre Auswirkung auf die Forschungskompetenz, das forschungsmethodische Wissen sowie die intrinsische Motivation untersucht.

2. Theoretischer Hintergrund

2.1 Forschungskompetenz

Forschungskompetenz stellt die Fähigkeit dar, selbstständig Forschung zu betreiben (Wulf et al., 2020) und untergliedert sich in eine kognitive und eine affektiv-motivationale Komponente (Wessels et al., 2018). Die kognitive Komponente beschreibt die notwendigen Fähigkeiten und Fertigkeiten zum Umgang mit und zur Durchführung von Forschung. Häufig wird darunter der Umgang mit forschungsmethodischem Wissen und Forschungsprozesswissen verstanden (Wulf et al., 2020) sowie die Fähigkeit, Wissen aus der Forschung zu recherchieren, zu verstehen und zu beurteilen (Böttcher-Oschmann et al., 2019). Damit einhergehend differenzieren Thiel und Böttcher (2014) hinsichtlich der kognitiven Komponente zwischen dem fachlichen Wissen über Forschung, der Recherche-, Methoden-, Reflexions- und Kommunikationskompetenz. Die affektiv-motivationale Komponente bezieht sich dagegen auf das gefühls- und wertbezogene Interesse an Forschung und ihren wahrgenommenen Nutzen (Wessels et al., 2021). Ihre Ausprägung gilt als entscheidend dafür, ob die Fähigkeiten und Fertigkeiten der kognitiven Komponente zur Anwendung kommen (Wulf et al., 2020).

2.2 Gestaltung forschungsmethodischer Lehrveranstaltungen mit videobasierten Lernmodulen

Für die Vermittlung von Inhalten in der universitären Lehre werden Videos vorwiegend in Form von Vorlesungsvideos eingesetzt (Weidlich & Spannagel, 2014). Videos als Kombination auditiver und visueller Darstellungen ermöglichen es, komplexe Inhalte in multimedialer Form zu vermitteln (Noetel et al., 2021). Das Lernen mit Videos erfolgt jedoch häufig nur oberflächlich, da die Inhalte meist kurzzeitig präsentiert und nicht adäquat elaboriert werden können (Merkt et al., 2011). Aus diesem Grund schlagen Weidlich und Spannagel (2014) die Implementation interaktiver Elemente (z. B. Quizfragen) vor, um ein bedeutsames und aktives Lernen zu fördern. Hypermediale Tutoring-Systeme (z. B. H5P) ermöglichen zudem eine Rückmeldung über den Lernfortschritt (Hillmayr et al., 2020).

Im Lernprozess spielen Rückmeldungen eine wichtige Rolle, da sie die Lücke zwischen dem aktuellen Wissens- und Kompetenzstand und dem Lernziel schließen (Hattie & Timperley, 2007). So zeigen Metanalysen insgesamt positive Effekte von Rückmeldungen auf computerbasiertes Lernen (Mertens et al., 2022; van der Kleij et al., 2015), wobei hauptsächlich drei Arten von Rückmeldungen entsprechend ihrer Komplexität (steigend von 1 nach 3) unterschieden werden: Bei (1) *knowledge of results* (hier: korrektive Rückmeldungen) werden Informationen über die Richtigkeit der Antwort gegeben (z. B. „Deine Antwort ist richtig/falsch“), bei (2) *knowledge of correct response* wird die richtige Antwort genannt (z. B. „Antwort C ist richtig.“) und bei (3) *elaborated feedback* (hier: unterstützende Rückmeldungen) werden weiterführende Informationen und Instruktionen gegeben, um zur richtigen Antwort zu gelangen (Taxipulati & Lu, 2021). Während Studien zu korrektiven Rückmeldungen konsistente Ergebnisse zeigen, fallen die Befunde zu anderen Arten von Rückmeldungen heterogen aus (Mertens et al., 2022). Es wird angenommen, dass korrektive Rückmeldungen nicht so lernwirksam sind wie unterstützende Rückmeldungen (van der Kleij et al., 2015).

Neben der Lernwirksamkeit wirken sich unterschiedliche Arten von Rückmeldung auch auf die Motivation aus, sich mit Inhalten auseinanderzusetzen. Wisniewski et al. (2020) gehen beispielsweise davon aus, dass Rückmeldungen umso effektiver sind, je mehr lernrelevante Informationen sie enthalten. Dies steht in Einklang mit den Annahmen der Selbstbestimmungstheorie (Ryan & Deci, 2017). Demnach können sich Rückmeldungen, die weitere

Hinweise für den Lernprozess beinhalten, positiver auf die Motivation und den Lernerfolg auswirken als korrektive Rückmeldungen (Großmann et al., 2020; Reeve, 2015).

Der bisherige Forschungsstand gibt keine konsistente Antwort auf die Frage, wie sich korrektive und unterstützende Rückmeldungen auf die Forschungskompetenz, das forschungsmethodische Wissen und die intrinsische Motivation von erfahrenen Lernenden (hier: Studierenden) auswirken (Mertens et al., 2022). Dieser Frage nimmt sich die vorliegende Studie an.

2.3 Hypothesen

Vor dem Hintergrund der bisherigen Ausführungen wird erwartet, dass sich unterstützende Rückmeldungen (im Vergleich zu korrektiven Rückmeldungen) positiv auswirken auf ...

... die Forschungskompetenz (**H1**).

... das forschungsmethodische Wissen (**H2**).

... die intrinsische Motivation (**H3**).

3. Methode

Die Untersuchung wurde innerhalb eines biologiedidaktischen Forschungsmoduls durchgeführt. An der Studie nahmen $N = 62$ Masterstudierende ($M = 25,40$ Jahre, $SD = 3,53$ Jahre; 74,2 % weiblich, 22,6 % männlich, 1,6 % divers, 1,6 % keine Angabe) teil.

3.1 Studiendesign und -ablauf

Die Teilnehmer:innen nahmen zu Beginn (Prätest) und gegen Ende (Posttest) des Semesters an einer Online-Umfrage teil. Im Prätest wurden personenbezogene Daten (Alter, Geschlecht), die Forschungskompetenz und das forschungsmethodische Wissen der Studierenden erhoben. Letztere wurden zudem im Posttest zusammen mit der intrinsischen Motivation der Studierenden bei der Bearbeitung der Lernmodule erfasst.

Zu Beginn des Semesters wurden die Studierenden randomisiert einer Gruppe (Experimentalgruppe, = EG; Kontrollgruppe = KG) zugeteilt. Beide Gruppen lernten mittels videobasierter Lernmodule, die sich aus einer theoretischen Perspektive auf grundlegende forschungsmethodische Inhalte des gesamten Forschungsprozesses bezogen (u. a. Verfahren

der Datenerhebung und -analyse, Interpretation und Bewertung fachdidaktischer Forschungsarbeiten). Die videobasierten Lernmodule wurden als *H5P-Elemente (interactive videos)* in die digitale Lernplattform *Ilias* eingebaut und enthielten Fragen, die sich auf die Video-Inhalte bezogen. Bei korrekter Beantwortung einer Frage erhielten die Studierenden beider Gruppen die Rückmeldung „richtige Antwort“. Wurde die Frage (z. B. „In welchem Bereich sollte die Itemschwierigkeit liegen?“) falsch beantwortet, erhielten Studierende aus der KG eine korrektive Rückmeldung („falsche Antwort“), wohingegen Studierende der EG eine unterstützende Rückmeldung erhielten, indem sie zur Sequenz im Video geleitet wurden, die eine richtige Beantwortung der Frage ermöglichte. Die Fragen konnten in beiden Gruppen beliebig oft wiederholt werden.

3.2 Messinstrumente

Die kognitive Komponente der Forschungskompetenz wurde anhand von fünf Subskalen erhoben und konnte über eine fünfstufige Rating-Skala von „trifft überhaupt nicht zu“ (1) bis „trifft genau zu“ (5) beantwortet werden (Böttcher & Thiel, 2017). Die affektiv-motivationale Komponente der Forschungskompetenz wurde über drei Subskalen (Wessels et al., 2018) erfasst. Die Items wurden anhand einer fünfstufigen Rating-Skalen bewertet. Die Subskala *Gefühlsbezogenes Interesse an Forschung* konnte von „macht mir überhaupt keinen Spaß“ (1) bis „macht mir sehr viel Spaß“ (5) beantwortet werden. Im Rahmen der zwei weiteren Subskalen konnten die Studierenden die Optionen „stimme nicht zu“ (1) bis „stimme zu“ (5) nutzen. Weitere Informationen zu den Subskalen finden sich in Tabelle 1.

Zur Erfassung des forschungsmethodischen Wissens wurde ein selbst erstellter Fragebogen mit Single- und Multiple-Choice-Fragen (13 Items) eingesetzt, der in vorangegangenen Semestern evaluiert wurde. Ein Punkt wurde für die richtige Beantwortung der Frage und null Punkte für eine falsche oder fehlende Beantwortung vergeben.

Die intrinsische Motivation bei der Bearbeitung der Lernmodule wurde mit der Kurzskala intrinsischer Motivation über eine fünfstufige Rating-Skala von „stimmt gar nicht“ (1) bis „stimmt völlig“ (5) gemessen (Wilde et al., 2009; Tab. 1).

Tabelle 1

Itemanzahl, Beispielitems und Cronbachs α -Werte der Subskalen der Forschungskompetenz und der intrinsischen Motivation

| | Itemanzahl | Beispielitems | α Prä-test | α Post-test |
|--|------------|--|-------------------|--------------------|
| Kognitive Komponente | | | | |
| Fachliches Wissen | 8 | „Ich habe eine solide Kenntnis der wichtigsten Forschungsmethoden in der Biologiedidaktik.“ | .89 | .88 |
| Recherchekompetenz | 4 | „Ich bin in der Lage, den Forschungsstand zu einem bestimmten biologiedidaktischen Thema zu sichten und systematisch aufzubereiten.“ | .87 | .82 |
| Methodenkompetenz | 7 | „Ich kann unterschiedliche Forschungsmethoden entsprechend meiner Fragestellung anwenden.“ | .83 | .80 |
| Reflexionskompetenz | 5 | „Ich bin in der Lage, meine eigenen biologiedidaktischen Forschungsergebnisse hinsichtlich der methodischen Grenzen kritisch zu reflektieren.“ | .88 | .83 |
| Kommunikationskompetenz | 4 | „Ich bin in der Lage, Forschungsbefunde für eine Präsentation in einem wissenschaftlichen Kolloquium aufzubereiten.“ | .73 | .58 |
| Affektiv-motivationale Komponente | | | | |
| Gefühlsbezogenes Interesse an Forschung | 11 | „Im Folgenden schildern wir Ihnen verschiedene Forschungstätigkeiten im Bereich der Biologiedidaktik, die einem mehr oder auch weniger Spaß machen können.“ - „Literatur zu einem Forschungsfeld recherchieren“ | .68 | .85 |
| Wertbezogenes Interesse an Forschung | 6 | „Die Forschung in der Biologiedidaktik ist nützlich, weil man durch Forschung interessante Erkenntnisse erzielen kann.“ | .77 | .79 |
| Wahrgenommener Nutzen der Forschung für die Praxis | 6 | „Theoretische Vorstellungen über die Erziehung erleichtern den Umgang mit unerwarteten Situationen im Unterricht.“ | .76 | .78 |
| Intrinsische Motivation | | | | |
| Interesse/Vergnügen | 3 | „Die Bearbeitung der Lernmodule hat mir Spaß gemacht.“ | | .89 |
| Wahrgenommene Kompetenz | 3 | „Ich glaube, ich war bei der Bearbeitung der Lernmodule ziemlich gut.“ | | .93 |
| Wahrgenommene Wahlfreiheit | 3 | „Bei der Bearbeitung der Lernmodule konnte ich so vorgehen, wie ich es wollte.“ | | .84 |
| Druck/Anspannung | 3 | „Bei der Bearbeitung der Lernmodule fühlte ich mich unter Druck.“ | | .90 |

4. Ergebnisse

4.1 Forschungskompetenz (H1)

Zunächst wurde überprüft, ob sich der Zuwachs in der kognitiven Komponente der Forschungskompetenz zwischen den Gruppen unterschied. Eine mixed MANOVA zeigte, dass dies nicht der Fall und der Interaktionseffekt, $F(1, 45) = 0.19, p = .976, \eta_p^2 = .02$, nicht signifikant war (vgl. auch Tab. 2). Jedoch zeigte der Haupteffekt der Zeit signifikante Unterschiede zwischen Prä- zu Posttest auf, $F(1, 45) = 26.93, p < .001, \eta_p^2 = .77$. Dieser Haupteffekt geht zurück auf die Zunahme des *Fachlichen Wissens*, $F(1, 45) = 88.45, p < .001, \eta_p^2 = .66$, der *Recherche-*, $F(1, 45) = 48.81, p < .001, \eta_p^2 = .52$, und *Methodenkompetenz*, $F(1, 45) = 19.87, p < .001, \eta_p^2 = .31$. Die *Reflexions-*, $F(1, 45) = 3.14, p = .083, \eta_p^2 = .65$, und *Kommunikationskompetenz*, $F(1, 45) = 0.71, p = .403, \eta_p^2 = .02$, veränderten sich jedoch nicht.

Auch wurde überprüft, ob sich der Zuwachs in der affektiv-motivationalen Komponente der Forschungskompetenz zwischen den Gruppen unterschied. Mittels einer mixed MANOVA wurde gezeigt, dass dies nicht der Fall und der Interaktionseffekt, $F(1, 45) = 0.57, p = .565, \eta_p^2 = .01$, nicht signifikant war. Auch der Haupteffekt der Zeit zeigte keine Zunahme, $F(1, 45) = 0.01, p < .947, \eta_p^2 = .00$.

4.2 Forschungsmethodisches Wissen (H2)

Es wurde überprüft, ob sich der Zuwachs des forschungsmethodischen Wissens zwischen den Gruppen unterschied. Eine mixed ANOVA zeigte, dass der Interaktionseffekt, $F(1, 36) = 0.01, p = .918, \eta_p^2 = .00$ nicht signifikant ausfiel. Der Haupteffekt der Zeit zeigte jedoch auch hier eine Zunahme, $F(1, 36) = 29.58, p < .001, \eta_p^2 = .45$.

4.3 Intrinsische Motivation (H3)

Es wurde eine MANOVA genutzt, um Unterschiede zwischen den Gruppen in den vier Subskalen der intrinsischen Motivation zu ermitteln, wobei kein statistisch signifikanter Unterschied zwischen den Gruppen vorlag, $F(4, 42) = 1.68, p = .173, \eta_p^2 = .14$, Wilk's $\Lambda = .862$.

Tabelle 2

Mittelwerte und Standardabweichungen der Subskalen der Forschungskompetenz, des forschungsmethodischen Wissens und der intrinsischen Motivation

| | Prätest | | | Posttest | | |
|--|----------------|----------------|----------------|----------------|----------------|----------------|
| | EG | KG | Gesamt | EG | KG | Gesamt |
| Kognitive Komponente | | | | | | |
| Fachliches Wissen | 1.94 (0.80) | 1.95 (0.61) | 1.95 (0.69) | 2.89 (0.80) | 2.89 (0.71) | 2.89 (0.74) |
| Recherchekompetenz | 2.93 (1.07) | 2.81 (0.79) | 2.86 (0.90) | 3.76 (0.80) | 3.71 (0.62) | 3.73 (0.69) |
| Methodenkompetenz | 3.11 (0.67) | 3.06 (0.65) | 3.08 (0.65) | 3.57 (0.49) | 3.54 (0.55) | 3.56 (0.52) |
| Reflexionskompetenz | 3.41 (0.73) | 3.45 (0.81) | 3.43 (0.77) | 3.63 (0.53) | 3.60 (0.50) | 3.61 (0.61) |
| Kommunikationskompetenz | 3.83 (0.65) | 3.48 (0.56) | 3.62 (0.62) | 3.84 (0.49) | 3.61 (0.47) | 3.70 (0.49) |
| Affektiv-motivationale Komponente | | | | | | |
| Gefühlsbezogenes Interesse an Forschung | 3.14 (0.44) | 3.06 (0.54) | 3.09 (0.50) | 2.99 (0.66) | 2.80 (0.74) | 2.88 (0.71) |
| Wertbezogenes Interesse an Forschung | 3.49 (0.71) | 3.40 (0.91) | 3.44 (0.59) | 3.47 (0.71) | 3.29 (0.66) | 3.36 (0.68) |
| Wahrgenommener Nutzen der Forschung für die Praxis | 3.67 (0.41) | 3.43 (0.55) | 3.53 (0.50) | 3.61 (0.53) | 3.47 (0.62) | 3.53 (0.58) |
| Forschungsmethodisches Wissen | 7.78 (1.67) | 7.21 (1.72) | 7.44 (1.70) | 9.66 (2.08) | 9.16 (1.58) | 9.35 (1.78) |
| Intrinsische Motivation | | | | | | |
| Interesse/Vergnügen | | | | 2.19 (0.80) | 2.27 (1.02) | 2.23 (0.91) |
| Wahrgenommene Kompetenz | | | | 3.42 (0.75) | 3.20 (0.92) | 3.31 (0.84) |
| Wahrgenommene Wahlfreiheit | | | | 3.33 (0.88) | 3.44 (1.04) | 3.39 (0.96) |
| Druck/Anspannung | | | | 2.07 (0.85) | 2.77 (1.22) | 2.42 (1.03) |

Anmerkung. EG=Experimentalgruppe, KG=Kontrollgruppe.

5. Diskussion

In der vorliegenden Studie wurde zunächst der Einfluss verschiedener Rückmeldungen (korrektiv/unterstützend) in videobasierten Lernmodulen auf die selbsteingeschätzte affektiv-motivationale und die kognitive Komponente der Forschungskompetenz (**H1**) untersucht. Der Befund zur kognitiven Komponente steht in Einklang mit früheren Studien, in denen Rückmeldungen als wirkungsvolles Instrument zur Förderung des Kompetenzerlebens identifiziert werden konnten (z. B. Garcia et al., 2019). Der vorliegende gruppunenabhängige Effekt könnte darauf zurückzuführen sein, dass eine korrektive Rückmeldung, die den

Studierenden lediglich aufzeigt, welche Inhalte noch nicht erlernt wurden, oder das selbstständige Arbeiten mit Lernvideos ausreichen, um die wahrgenommene Kompetenz der Studierenden zu unterstützen.

Hinsichtlich der affektiv-motivationalen Komponente der Forschungskompetenz kann angeführt werden, dass frühere Studien bereits aufzeigten, dass Rückmeldungen effektiver für kognitive als für motivationale Variablen sein können (Wisniewski et al., 2020). Zu berücksichtigen ist, dass die affektiv-motivationale Komponente das Interesse an Forschung und den wahrgenommenen Nutzen von Forschung umfasst (Wessels et al., 2018). Die videobasierten Lernmodule in unserer Studie waren lediglich auf die Vermittlung von sowie die Rückmeldungen zu den Inhalten ausgelegt und konnten daher das Interesse sowie den Nutzen von Forschung möglicherweise nicht hinreichend ansprechen (Gussen et al., 2023b).

Bezüglich **H2** konnte analog zur Entwicklung der kognitiven Komponente der Forschungskompetenz ein gruppenunabhängiger Zuwachs im forschungsmethodischen Wissen festgestellt werden. Dieser Befund steht in Einklang mit der Annahme, dass eine Entwicklung der kognitiven Komponente im Zuge des Erwerbs forschungsmethodischen Wissens stattfindet (Fichten, 2010). Da Rückmeldungen den Lernenden ihren Lernfortschritt transparent machen (Hattie & Timperley, 2007), wird angenommen, dass diese im Lernprozess förderlich für den Wissenszuwachs von Studierenden sind (Mertens et al., 2022; van der Kleij et al., 2015). Die Befunde der vorliegenden Studie sprechen dafür, dass die zwei untersuchten Arten von Rückmeldungen den Wissenszuwachs gleichermaßen unterstützen. Allerdings ist zu berücksichtigen, dass sich die Studierenden beider Gruppen zum Zeitpunkt des Posttests aufgrund der anstehenden Prüfung bereits vertieft mit den Inhalten der Lehrveranstaltungen befasst haben (Gussen et al., 2023a), sodass der Effekt des Treatments möglicherweise maskiert wurde. Zugleich ist anzumerken, dass keine Informationen darüber vorliegen, ob die Studierenden die unterstützenden Rückmeldungen lediglich als Information oder tatsächlich als Anreiz zur vertieften Auseinandersetzung mit den Inhalten wahrgenommen haben.

Für die intrinsische Motivation (**H3**) wurde zuletzt ebenfalls angenommen, dass sich Rückmeldungen, die sich in ihrem Informationsgehalt unterscheiden, unterschiedlich auswirken (Wisniewski et al., 2020). Rückmeldungen, die weitere Hinweise für den Lernprozess beinhalten, sollten die Wahrnehmung von Kompetenz und in Folge die intrinsische Motivation im Vergleich zu korrekativen Rückmeldungen fördern (Großmann et al., 2020; Reeve, 2015). Da durch diese Hinweise ein Gleichgewicht von Fähigkeiten und Anforderung auf Seiten der Lernenden besser hergestellt werden kann als ohne diese Hinweise (Großmann et al.,

2020; Reeve, 2015), könnte auch der wahrgenommene Druck, sich mit den Inhalten weiter auseinanderzusetzen, vermindert werden. Diese auf Theorie und Empirie basierten Annahmen werden durch die vorliegenden Befunde nicht gestützt. Möglicherweise könnte es sein, dass bereits zu Semesterbeginn Unterschiede in der intrinsischen Motivation der Studierenden, sich mit Forschung zu beschäftigen, vorlagen. Diese Variable sollte in zukünftigen Studien einbezogen werden. Zugleich könnte die intrinsische Motivation im Verlauf des Semesters betrachtet werden, um zu überprüfen, ob sich die Effekte der Maßnahme über die Zeit verändern.

Vor dem Hintergrund unserer Befunde bleibt die Frage offen, wie Rückmeldungen gestaltet werden können, um Studierende hinsichtlich der affektiv-motivationalen Komponente der Forschungskompetenz und der intrinsischen Motivation zur Auseinandersetzung mit Forschung zu unterstützen. Auch die kognitive Komponente sowie der Wissenszuwachs sollten dabei weitere Beachtung finden, da eine positive Veränderung, jedoch keine Unterschiede zwischen den Arten der Rückmeldung festgestellt werden konnten. Möglicherweise könnten individuelle Rückmeldungen eine zusätzliche Unterstützung im Lernprozess darstellen. Hinsichtlich der Förderung der affektiv-motivationalen Komponente und der intrinsischen Motivation könnte insbesondere die Relevanz und der Nutzen von Forschung für den späteren Lehrberuf und damit einhergehend das Interesse an Forschung angesprochen werden (Gussen et al., 2023b; Wessels et al., 2021).

6. Literatur

- Besa, K., Lüking, S., Biel, A. & Wilde, M. (2023). Forschungskompetenz von Lehramtsstudierenden und Studierenden anderer Fachrichtungen. *Zeitschrift für Bildungsforschung*, 13(55), 55–74.
- Böttcher, F. & Thiel, F. (2017). Ergebnisse der Evaluation der Forschungsorientierten Lehre (FoL) an der Freien Universität Berlin. *Freie Universität Berlin*, 1–85. https://www.fu-berlin.de/sites/fof/_media/FoL-Evaluationsbericht_Dez_2017.pdf
- Böttcher-Oschmann, F., Groß-Ophoff, J. & Thiel, F. (2019). Validierung eines Fragenbogens zur Erfassung studentischer Forschungskompetenzen über Selbsteinschätzungen – Ein Instrument zur Evaluation forschungsorientierter Lehr-Lernarrangements. *Unterrichtswissenschaft*, 47, 495–521.
- Fichten, W. (2010). Forschendes Lernen in der Lehrerbildung. In U. Eberhardt (Hrsg.), *Neue Impulse in der Hochschuldidaktik* (127–182). VS Verlag.

- García, J.A., Carcedo, R.J. & Castaño, J.L. (2019). The influence of feedback on competence, motivation, vitality, and performance in a throwing task. *Research quarterly for exercise and sport*, 90(2), 172–179.
- Großmann, N., Eckes, A. & Wilde, M. (2020). Prädiktoren der Kompetenzwahrnehmung von Schülerinnen und Schülern im Biologieunterricht. *Zeitschrift für Pädagogische Psychologie*, 36(3), 153–165.
- Gussen, L., Schumacher, F., Ferreira-González, L., Schlüter, K. & Großschedl, J. (2023a). Entwicklung biologiedidaktischer Forschungskompetenz bei Lehramtsstudierenden. *die hochschullehre*, 8, 101–116.
- Gussen, L., Schumacher, F., Ferreira-González, L., Schlüter, K. & Großschedl, J. (2023b). Supporting pre-service teachers in developing research competence. *Frontiers in Education*, 8, 1–9.
- H5P (2023). Interactive Video. <https://h5p.org/interactive-video>
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112.
- Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S.I. & Reiss, K.M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computer & Education*, 153, 1–25.
- KMK [Kultusministerkonferenz] (Hrsg.) (2024). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung* (Beschluss der Kultusministerkonferenz vom 16.10.2008 i. d. F. vom 08.02.2024).
- Kuklick, L. & Lindner, M.A. (2021). Computer-based knowledge of results feedback in different delivery modes: Effects on performance, motivation, and achievement emotions. *Contemporary Educational Psychology*, 67.
- Merkt, M., Weigand, S., Heier, A. & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction*, 21, 687–704.
- Mertens, U. Finn, B., & Lindner, M.A. (2022). Effects of Computer-Based Feedback on Lower- and Higher-Order Learning Outcomes: A Network Meta-Analysis. *Journal of Educational Psychology*, 114(8), 1743–1772.
- Noetel, M., Griffith, S., Delaney, O., Sanders, T., Parker, P., del Pozo Cruz, B. & Lonsdale, C. (2021). Video Improves Learning in Higher Education: A Systematic Review. *Review of Educational Research*, 91(2), 204–236.
- Reeve, J. (2015). *Understanding motivation and emotion* (6. Aufl.). Wiley.
- Ryan, R.M. & Deci, E.L. (2017). *Self-determination theory – Basic psychological needs in motivation, development, and wellness*. Guilford Press.
- Taxipulati, S. & Lu, H. (2021). The Influence of Feedback Content and Feedback Time on Multimedia Learning Achievement of College Students and Its Mechanism. *Frontiers in Psychology*, 7, 1–14.

- Thiel, F. & Böttcher, F. (2014). Modellierung fächerübergreifender Forschungskompetenzen. *Neues Handbuch Hochschullehre*, 2(10), 109–124.
- Van der Kleij, F.M., Feskens, R.C.W. & Eggen, T.J.H.M. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes: A Meta-Analysis. *Review of Educational Research*, 85(4), 474–511.
- Weidlich, J. & Spannagel, C. (2014). Die Vorbereitungsphase im Flipped Classroom. Vorlesungsvideos versus Aufgaben. In K. Rummler (Hrsg.), *Lernräume gestalten - Bildungskontexte vielfältig denken* (237–248). Waxmann.
- Wessels, I., Rueß, J., Jenßen, L., Gess, C. & Deicke, W. (2018). Beyond cognition: experts' views on affective-motivational research dispositions in the social sciences. *Frontiers in Psychology*, 9.
- Wessels, I., Rueß, J., Gess, C., Deicke, W. & Ziegler, M. (2021). Is research-based learning effective? Evidence from a pre–post analysis in the social sciences. *Studies in Higher Education*, 46(12), 2595–2609.
- Wilde, M., Bätz, K., Kovaleva, A. & Urhahne, D. (2009). Überprüfung einer Kurzsкала intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 31–45.
- Wisniewski, B., Zierer, K. & Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, 10, 1–14.
- Wulf, C., Thiem, J. & Gess, C. (2020). Motivationale Faktoren im Wirkungskontext von Forschendem Lernen. In C. Wulf, S. Haberstroh & M. Petersen (Hrsg.), *Forschendes Lernen* (129–144). Springer VS.

4.2 Fostering Learning with Incremental Scaffolds During Chemical Experimentation: A Study on Junior High School Students Working in Peer-Groups

4.2.1 Lernstrategie Scaffolds: Theoretischer Zugang, Hypothesen, Studiendesign

Theoretischer Zugang

Die Bearbeitung komplexer (Problemlöse-) Aufgaben im Rahmen naturwissenschaftlichen Unterrichts kann zu Überforderung seitens der Lernenden führen (Schmidt-Weigand et al., 2008). Um dem zu begegnen werden Scaffolds eingesetzt (deutsch: Gerüste) als Lernstrategie (Arnold et al., 2016). Diese Scaffolds können von Lernenden optional in Anspruch genommen werden, wenn sie Schwierigkeiten haben bei der Aufgabebearbeitung (Schmidt-Weigand et al., 2008). Der Einsatz von Scaffolds im Unterricht bedeutet für Lehrende, dass sie eine stetige Einschätzung, Anpassung und sukzessive Entfernung des „Gerüsts“ beziehungsweise der Unterstützung vornehmen müssen (Arnold et al., 2016). Unterschieden wird bei Scaffolds zwischen *soft scaffolds* und *hard scaffolds* (Brush & Saye, 2008; Saye & Brush, 2002).

Während *soft scaffolds* eine situationsbezogene Unterstützung darstellen, die von den Lehrenden in der Regel bereitgestellt wird auf Anfrage der Lernenden, sind *hard scaffolds* eine im Vorhinein vorbereitete und statische Unterstützung, die von den Lernenden im Bedarfsfall genutzt werden kann (Arnold et al., 2016). Eine Form von *hard scaffolds* stellen gestufte Lernhilfen (englisch: *incremental scaffolds*) dar (Schmidt-Weigand et al., 2008, 2009). Diese bilden Unterstützung aus aufeinander aufbauenden Schritt-für-Schritt Hinweisen zur Lösung der Aufgabe (Hänze et al., 2010; Stäudel et al., 2007). Wenn Lernende eine Aufgabe erhalten, die sie nicht ohne Unterstützung meistern können, können sie einen ersten Lösungshinweis in Anspruch nehmen. Wenn sie weitere Unterstützung benötigen, können weitere Lösungshinweise schrittweise genutzt werden, bis Lernende eine vollständige Lösung erhalten (Schmidt-Weigand et al., 2008; Stäudel et al., 2007). Das Ausmaß an Unterstützung zur Bearbeitung der Aufgabe können Lernende selbst regulieren (Schmidt-Weigand et al., 2008). Durch die Nutzung gestufter Lernhilfen erleben Lernende eine höhere Kompetenzwahrnehmung und eine niedrigere kognitive Belastung (Schmidt-Weigand et al., 2009). Die Reduktion der kognitiven Belastung kann auf die Quelle extrinsischer kognitiver Belastung attribuiert werden, wobei angenommen wird, dass gestufte Lernhilfen die

Aufgabenbearbeitung stärker strukturieren und die Arbeitsschritte dadurch vereinfachen (Arnold et al., 2016).

Gestufte Lernhilfen weisen einen binnendifferenzierenden Charakter auf (Kleinert et al., 2022), da weniger kompetente Lernende bei Schwierigkeiten die Lösungshinweise in Anspruch nehmen können, während kompetentere Lernende nach der Aufgabenbearbeitung ihre eigene Lösung mit der Beispiellösung der gestuften Lernhilfen abgleichen können (Hänze et al., 2010). Besonders für das naturwissenschaftliche Lernen ist dies von Vorteil, da innerhalb der gegebenen zunehmend heterogenen Lerngruppen (siehe Kapitel 1), naturwissenschaftlich-kompetenten und weniger kompetenten Lernenden ermöglicht wird, die Aufgaben zu bearbeiten und sie ein adäquates Maß an Unterstützung erhalten (Hänze et al., 2010; Stäudel et al., 2007).

Hinsichtlich der Wirksamkeit gestufter Lernhilfen zeigen empirische Studien inkonsistente Ergebnisse bei verschiedenen Arten von Wissen (Hülsmann et al., 2024). Beispielsweise wurde in der Studie von Stiller und Wilde (2021) kein Effekt auf das konzeptuelle Wissen indiziert (Stiller & Wilde, 2021). Da konzeptuelles Wissen beim Lernen in den Naturwissenschaften von zentraler Bedeutung ist (siehe Kapitel 2.1.2) und auch im Rahmen des Experimentierens eine wichtige Rolle spielt, überrascht es, dass bisher nur wenige Studien vorliegen zum Einsatz von gestuften Lernhilfen beim Experimentieren (Kleinert et al., 2022) und deren Auswirkungen auf das konzeptuelle Wissen (Stiller & Wilde, 2021).

Hypothesen

In nachfolgender Studie wird die Wirksamkeit gestufter Lernhilfen (als Form von *hard scaffolds*) überprüft, indem diese der mündlichen instruktionalen Unterstützung durch die Lehrkraft (*soft scaffolds*) gegenübergestellt werden. Dabei soll ermittelt werden, welche Form des Scaffolds, Schüler:innen beim Erwerb konzeptuellen Wissens im Rahmen forschungsbasierten Lernens mit Experimenten effektiver unterstützt und zum Abbau fachlich inadäquater Vorstellungen führt. Folgende Überlegungen liegen der Studie zugrunde:

Das Experimentieren mit chemischen Reaktionen hilft, biologische Prozesse (z. B. Photosynthese, Enzymreaktion, Zellatmung und Energieumwandlung) besser zu verstehen. Chemische Reaktionen werden daher häufig im Rahmen des Biologieunterrichts eingesetzt (Bach, 2019). Trotz inkonsistenter Ergebnisse hinsichtlich verschiedener Arten von Wissen, kann angenommen werden, dass gestufte Lernhilfen für den Erwerb konzeptuellen Wissens

eine geeignete Lernstrategie darstellen. Im Rahmen forschungsbasierten Lernens mit Experimenten ermöglicht der Einsatz gestufter Lernhilfen, den Lernenden Schlussfolgerungen aus den Experimenten zu ziehen und in Folge dessen allgemeine Prinzipien ableiten und reflektieren zu können (Stiller & Wilde, 2021). Folglich kann folgende Hypothese formuliert werden:

Hypothese 2: Schüler:innen, die gestufte Lernhilfen nutzen, zeigen einen signifikant höheren Zuwachs an konzeptuellem Wissen, im Vergleich zu Schüler:innen, die instruktionale Unterstützung durch die Lehrkraft erhalten.

Studiendesign

Die nachfolgende Studie überprüfte in einer quasi-experimentellen Untersuchung, die Wirksamkeit gestufter Lernhilfen hinsichtlich des konzeptuellen Wissens mit $N = 105$ Schüler:innen der siebten Klasse. Das Studiendesign gliedert sich in vier Phasen. (1. Phase:) Concept Map-Trainingsphase - Als Erhebungsinstrument für das konzeptuelle Wissen wurden Concept Maps, als graphische Darstellung der Beziehung zwischen einzelnen Konzepten (Ruiz-Primo & Shavelson, 1996, siehe auch Kapitel 4.3) eingesetzt. Für den Einsatz dieser wird ein Training empfohlen (Becker, Welter, Aschermann & Großschedl, 2021), welches Schüler:innen in dieser Phase erhielten. In dieser Concept Map-Trainingsphase erarbeiteten sich die Schüler:innen in Partnerarbeit, akkurate Charakteristika einer Concept Map und erstellten selbst eine Concept Map. Zu dieser Concept Map erhielten die Schüler:innen vor Beginn der nächsten Phase Rückmeldung über die Richtigkeit. (2. Phase:) In dem Prä-Test wurden demographische Daten (z. B. Alter, Geschlecht) und das konzeptuelle Vorwissen über Concept Maps zum Thema *chemische Reaktion* erhoben. Hierzu erstellten Schüler:innen eine Concept Map in Einzelarbeit. (3. Phase:) In der Interventionsphase wurde zunächst deklaratives Wissen zu endothermen und exothermen Reaktionen über einen mündlichen Vortrag der Lehrkraft vermittelt. Anschließend führten Schüler:innen in Partnerarbeit zwei Experimente (mit Kupfersulfat) durch und arbeiteten dann in Einzelarbeit an drei experiment-bezogenen Aufgaben. Während Schüler:innen der Experimentalbedingung ($n = 59$) gestufte Lernhilfen nutzen konnten, konnten Schüler:innen der Kontrollbedingung ($n = 46$) instruktionale Unterstützung durch die Lehrkraft einholen. (4. Phase:) Im Post-Test wurden, analog zum Prä-Test, Concept Maps zum Thema chemische Reaktionen eingesetzt, mittels derer der Zuwachs an konzeptuellem Wissen ermittelt werden sollte.

4.2.2 Eigenanteil

Die **Studie II** wurde von Timo Dabrowski konzipiert und durchgeführt, sowie maßgeblich ausgewertet. Bei der Erstellung des Manuskripts waren Anestis Ioannidis und Dr. Laura Ferreira Gonzalez maßgeblich beteiligt. Die Autorin war wesentlich an den Visualisierungen, anteilig an der Datenauswertung und -interpretation sowie Erstellung des Manuskripts und wesentlich an der Überarbeitung des Manuskripts beteiligt. Die Überarbeitung und Betreuung übernahm Prof. Dr. Großschedl.

Fostering Learning with Incremental Scaffolds During Chemical Experimentation: A Study on Junior High School Students Working in Peer-Groups

Mirlinda Mustafa^{a*}, Anestis Ioannidis^{a*}, Laura Ferreira González^{a*}, Timo Dabrowski^b, Jörg Grossschedl^{a*}

Corresponding author: j.grossschedl@uni-koeln.de

^a Institute for Biology Education, University of Cologne, Köln 50931, Germany

^b Dreikönigsgymnasium, Cologne, Köln 50931, Germany

*These authors contributed equally to this article.

Keywords: conceptual knowledge, incremental scaffolds, inquiry-based learning, junior high school students

¹ Publikationsmanuskript wurde lediglich hinsichtlich formaler Aspekte angepasst

Abstract

Scaffolds are considered to be a promising method of supporting learning. In this study, we investigated the learning efficacy of scaffolds in an inquiry-based learning scenario. Three tasks posed a question / problem to facilitate inquiry-based learning, and scaffolds offered the answer / solution in multiple steps (so-called incremental scaffolds). The use of the scaffolds was voluntary and students' learning efficacy was compared with a traditional teaching approach. A total of $N = 105$ seventh graders participated in the quasi-experimental study. Incremental scaffolds were available to the students in the treatment group. Students in the control group received the same question / problem but could only ask the teacher about the answer / solution. Concept maps were used at pre- and posttest to assess conceptual knowledge acquisition. In-line with our hypothesis, results show that students in the treatment group outperformed controls concerning conceptual knowledge acquisition. Regarding the number of misconceptions students used, there were no differences between the groups. Our study indicates that incremental scaffolds are an appropriate method to provide students with the exact help they really need. Based on our findings, we offer practical implications and recommendations for future research.

Introduction

Inquiry-based learning

Inquiry-based learning is a pedagogical approach for student-centred learning, which starts by posing questions, scenarios, or problems. Learners address these issues through inquiry and intellectual engagement in order to find answers / solutions, to develop a deeper understanding of the underlying concepts, and to acquire conceptual knowledge (Hmelo-Silver, Duncan, & Chinn, 2007; Tan, Koppi, & Field, 2016; Wangdi, Precharattana, & Kanthang, 2020). Inquiry-based learning improves the students' independence, as they undergo the research process as a whole (Mieg, 2019), from developing questions and hypotheses, selecting the methods, and presenting the results (Pedaste, Mäeots, Leijen, & Sarapuu, 2012). Inquiry-based learning is cognitively demanding, requires advanced metacognitive skills and a high degree of motivation (Thomas, Bennett, & Lockyer, 2016; Zimmermann, 1998). Students facing these challenges often need instructional support (Schmidt-Weigand, Hänze, & Wodzinski, 2009). Tight instructional support (i.e., guidance) decreases students' motivation by restricting their scope of action, but too little instructional support also negatively affects motivation because of a high risk of failing the task (Bjonness & Kolsto 2015; van de Pol &

Elbers, 2013). Therefore, scholars recommend a middle ground of instructional support, which gives the students the responsibility for the learning process (Haltunen, 2003) and adapts the teacher's degree of control in such a way that guidance is tailored to students' needs. Providing an adequate level of instructional support is particularly challenging in an increasingly heterogeneous classroom (Forghani-Arani, Cerna, & Bannon, 2019). Hence, adequate instructional support which facilitates step-by-step learning, is of great importance (Hmelo-Silver et al., 2007).

Incremental scaffolds

Scaffolds provide a solution or answer to a problem or question that arise from a particular task. Sometimes, scaffolds also give additional information or include prompts (i.e., hints to find the answer / solution) to support the students. A particular type of scaffolds, so-called incremental scaffolds, presents the answer / solution step-by-step to the students (Schmidt-Wiegand, Franke-Braun, & Hänze, 2008) or gradually provides additional information or prompts. The step-by-step presentation reduces the complexity of the task with the aim to decrease the necessary amount of working memory resources (so-called cognitive load; cf. Sweller, van Marrienoer, & Paas, 1998). Scaffolding enables a learner "to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts" (Wood, Bruner, & Ross, 1976, p. 90). Franke-Braun, Schmidt-Weigand, Stäudel, and Wodzinski (2008) emphasise that incremental scaffolds are particularly useful for students requiring special assistance. Incremental scaffolds provide the support these students actually need without exposing their lack of knowledge in the classroom.

Schmidt-Weigand et al. (2009) compared three ways of supporting students' problem solving in a collaborative context: The solution was given (1) at once (worked-out examples), (2) in multiple steps without (incremental scaffolds) or (3) with particular prompts like questions, graphics, or hints to promote active thinking (incremental strategic scaffolds). They found that an incremental scaffolding increased the motivation of the students, improved their feeling of competence, and led to more problem- and regulation-focused communication between students working in pairs. This is supported by recent research, highlighting the positive effects of scaffolding on performance and motivation (Krause, Stark, & Mandl, 2004; Lou, Abrami, & D'Apollonia, 2001).

A source of scaffolding can be 'soft' or 'hard' (Saye & Brush, 2002). Soft scaffolding refers to the support provided by a teacher or a peer when required. It demands constant monitoring

of students' performance to provide the right amount of help at the right time. Due to practical difficulties with this concept in a class of 20-30 students, hard scaffolding plays an important role. Hard scaffolds are paper-and-pencil or electronic tools which anticipate the particular needs of the students when learning. They often consist of a question, a hint, or a prompt that stimulates students to think in more depths about a question or problem (Belland, Glazewski, & Richardson, 2008). However, we are not aware of any research that compares learning with incremental scaffolds (hard scaffolds) with instructional support from the teacher (soft scaffolds; Arnold, Kremer, & Mayer, 2017). As the latter is the most common way of classroom learning, this comparison has high ecological validity.

Conceptual knowledge

To be knowledgeable in science implies the understanding how scientific concepts are inter-related. These interrelated concepts are stored in long-term memory (Veríssimo Catarreira, Godinho Lopes, Casas García, & Luengo González, 2017) and represent a person's conceptual knowledge (Krathwohl, 2002). Conceptual knowledge comprises the knowledge of facts, laws, and principles, all of which are necessary for dealing with different tasks in the academic environment (Furtak, Seidel, Iverson, & Briggs, 2012). Conceptual knowledge makes it possible to edit and understand multiple stimuli, such as words, sounds, or pictures and allows the expression of knowledge in both verbal and non-verbal ways. It also mediates the generalisation or transfer of knowledge from one domain to another (Lambon-Ralph, Pobric, & Jefferies, 2009).

Research question

Schmidt-Weigand et al. (2008, p. 373) show that students supported by incremental scaffolds when learning the scientific concept of density, outperformed the control group in a paper and pencil test measuring students' understanding of density (medium effect of Cohens $d = 0.49$; $N = 63$ ninth graders). Contrarily, the research of Franke-Braun et al. (2008) shows only a minor benefit of incremental scaffolds; they explain these results with the complexity of the incremental scaffolds, the lack of students' experience with learning in pairs as well as the lack of distinct metacognitive abilities ($N = 62$ ninth graders). However, it is difficult comparing those findings as the studies differ in regard to the research foci, the control groups used, and the learning approaches. Consequently, our research question is whether

an incremental scaffolding (step-by-step; treatment group) leads to a significant increase of conceptual knowledge in comparison to a traditional approach, where students seek instructional support from the teacher if they have questions (control group).

Methods

Sample

We conducted an a priori power analysis with G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007) to determine the required sample size using the following specifications: We considered a medium effect to be desirable for our treatment ($f = .25$; cf. Hattie [2012, p. 15]), defined an α error level of .05, and a power of $1 - \beta = .90$ (according to Whitley and Ball [2002] the power represents the chance of correctly identifying a significant difference between two groups when the difference really exists in the population). Based on these specifications, the power analysis revealed a required sample size of 128 students. Our actual sample consisted of 132 students ($n = 65$ male, $n = 67$ female) from an integrative comprehensive secondary school in Germany (five seventh-grade classes) aged between 12 and 14 years old ($M = 12.43$ years, $SD = 0.63$). Due to a number of dropouts, only the data of $N = 105$ students could be included in our analysis. Despite this relatively high drop-out rate of 20%, a post-hoc power analysis (Faul et al., 2007) was conducted which still showed an adequate power of $1 - \beta = .81$ for the reduced sample size. A total of $n = 59$ students (50.8% male; $M = 12.42$ years, $SD = 0.59$) attended the treatment group, $n = 46$ students (47.8% male; $M = 12.39$ years, $SD = 0.68$) attended the control group.

Research design and procedure

To address our research question, we conducted a quasi-experimental intervention study with a pre- and posttest design. The pretest, the intervention, and the posttest were all carried out on consecutive days, and both tests involved the same concept mapping-task to assess conceptual knowledge. Concept maps are diagrams with a network structure consisting of nodes (concepts) which are connected via labelled arrows in a meaningful way (Novak & Cañas, 2008). As students are often unfamiliar with the concept mapping-technique, a training session preceded the pretest to prevent inadequate concept mapping-skills by the students. In the intervention phase, which followed the pretest, chemistry classes were randomly assigned to the treatment and control groups. A random assignment of the students to

both groups was not possible because of organisational reasons (e.g., restricted availability of parallel laboratory rooms). Both groups conducted the same two experiments on chemical reactions in pairs and had to work on the same three tasks. The tasks were presented on worksheets, related to the experiments, and posed a question / problem to facilitate inquiry-based learning. Working on the tasks was supported by incremental scaffolds in the treatment group and by the teacher in the control group (see section intervention for more detail). As “learning science means learning to do science” (Brewer & Smith, 2009, p. 14), inquiry-based learning provides “the opportunity to generate scientific knowledge through research” (The President's Council of Advisors on Science and Technology, 2012, p. 25).

Our study strictly adhered to the ethical guidelines of the Declaration of Helsinki and was permitted by our Ministry of School and Further Education (10-45 No. 2). Admission was subject to the condition that participants were informed about (1) aims and process of the investigation, (2) entire voluntariness of participation, (3) possibility of dropping out of participation at any time, (4) guaranteed protection of data privacy (collection of only anonymised data), (5) possibility of requesting data cancelation, (6) no-risk character of study participation, and (6) contact information in case of any questions or problems. Additionally, we obtained the written and informed consent of all participants, as well as of their parents prior to the study.

Concept map-training

The concept mapping-training lasted 60 minutes and the students worked in pairs. Two posters, each showing a concept map on the same topic (‘Life in the wilderness’) were presented to the students on the blackboard. One concept map was correct, the other one contained common mistakes (e.g., arrows pointing in the wrong direction). Students compared both maps and identified the main characteristics of an accurate concept map. After this, they were given a sheet of paper with a pre-constructed concept map on a different topic (‘Teaching natural sciences in school’), with empty nodes and arrows to be filled in with the right words from a given list. The training ended with the students constructing a concept map by themselves on the topic of ‘zoology’ by using a given list of concepts and linking words. This task was similar to the pre- and posttest. To ensure that students could master this task, concept maps were evaluated by the instructor after the training session; a 15-minute feedback was then provided prior implementing the pretest. None of the example-topics were linked to the content of the test- and learning phases.

Pretest

Students' conceptual knowledge was assessed by a concept mapping-task. Students got 35 minutes time to construct a concept map on a focus question on a piece of paper which referred to the target-topic 'chemical reactions'. A list of 14 concepts (e.g., activation energy, reactant) and 13 linking words (e.g., heat) were provided to the students, no collaboration among students was allowed. Although, a variety of other methods, for example, multiple choice tests (DiBattista & Kurzawa, 2011), interpretive essays (Bolte, 1999), or similarity judgements tests (Großschedl & Harms, 2013) promise a valid assessment of conceptual knowledge, concept maps seemed promising for several reasons. Concept maps start with an open-ended question (the focus question) and provide richer insights into students' thinking processes and understanding than responses to closed-ended questions like multiple choice questions or the questions of a similarity judgements test. Like interpretive essays, concept maps reveal students' misconceptions, but their scoring / evaluation requires significantly less effort than it is the case with essays (Brandstädter, Harms, & Großschedl, 2012; Großschedl, Mahler, & Harms, 2018).

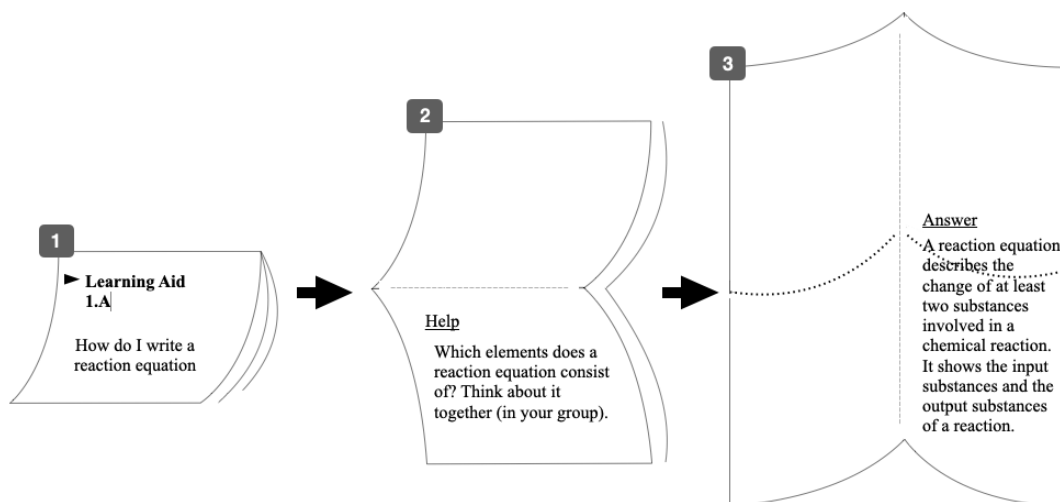
Intervention phase

The intervention phase lasted 60 minutes and took place in a chemistry laboratory of the school. All students attended a lecture given by the instructor on the topic 'endothermic chemical reaction.' In the following 15 minutes, paired students conducted two experiments with copper sulphate. They were then given 20 minutes to work individually on three tasks referring to the conducted experiments. These tasks were identical for the treatment and control groups. In the treatment group, four sets of incremental scaffolds were provided (their use was explained to the treatment group before the intervention phase), addressing each of the three tasks: two scaffold-sets on chemical equations and reactants / products for the first task; one scaffold-set on exothermic / endothermic reactions for the second task, and one scaffold-set on activation energy for the last task. Each incremental scaffold-set consisted of four (A, B, C, D) parts (see Figure 1 for an example; here part A of scaffold-set 1) and each part was printed on a sheet of paper and folded three times. By unfolding the sheet of paper, students were given the particular parts step-by-step. These parts included additional prompts or hints, an extra piece of information, or an example from everyday life. The use of incremental scaffolds was optional. The control group received a hint equivalent to those on the incremental scaffolds in the form of a simple solution given by the teacher on request. In

case of content-related questions, the students in the treatment group were encouraged to use the scaffolds, whereas the control group received an answer from the teacher when asked. No further content-related questions were answered in any group, whereas general questions (e.g., questions about the time still available) were answered in both groups. The learning phase finished with a presentation of the solution to the entire class, so that all students could correct their answers.

Figure 1

Example of an incremental scaffold from the student's point of view. Step 1: Number of the incremental scaffold (upper corner) and question. Step 2: After the first unfold, the help is revealed. Step 3: After the second unfold, the answer appears



Posttest

The posttest lasted 30 minutes and was identical to the pretest.

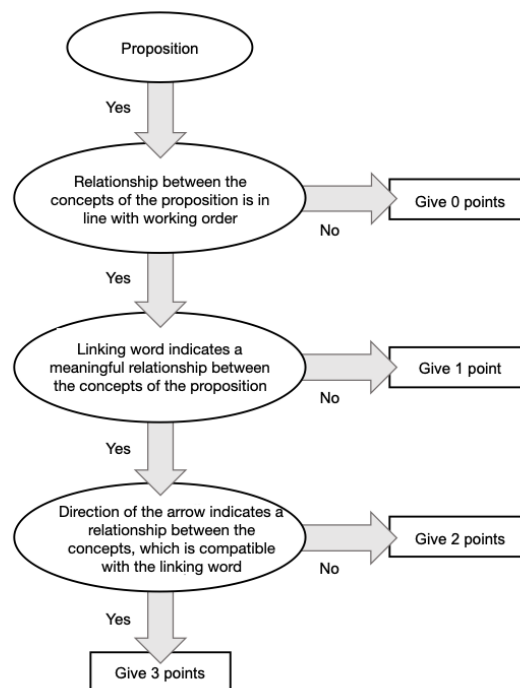
Analysis of pretest and posttest-concept maps

Students' conceptual knowledge was assessed by a quantitative and qualitative approach. The quantitative approach refers to the relational scoring method developed by McClure, Sonak, and Suen (1999). This method involves the scoring of the individual propositions of a concept map. Propositions consist of two nodes (concepts) connected to each other with a labelled arrow. A completely mistaken proposition, which depicts a relationship between two concepts that is not reasonable from the scientific point of view, was scored with a zero. If two related concepts were connected with each other, but with the wrong kind of relationship (mistaken linking word), one point was granted. In case the type of relationship was

correct, but the direction of the arrow was wrong, two points were given. An error-free proposition was given three points (for a detailed description of the coding scheme see Figure 2). Students' conceptual knowledge was then expressed as a sum score (so-called proposition accuracy score), which is derived from the summation of the individual sub-scores for each proposition. In order to check the quality of scoring, ten percent of the concept maps from pre- and posttest were randomly chosen and independently scored by a second rater (Döring & Bortz, 2016). Spearman correlation was calculated to determine interrater reliability and indicates reliable scoring ($r = .92, p < .001$).

Figure 2

Coding scheme according to McClure et al. (1999; adapted by Brandstädter et al., 2012)



The qualitative approach refers to the assessment of misconceptions. Misconceptions can be defined as ‘any conceptual idea that differs from the commonly accepted scientific consensus’ (Garnett & Treagust, 1990, p. 147). Students already have ingrained ideas and concepts that are inconsistent with, or even in strong contrast to, scientific views (Duit & Treagust, 2003). Therefore, misconceptions pose a challenge for education in science because they can be widespread amongst students and can be resistant to change (Smith, III, diSessa, &

Roschelle, 1993). Previous studies show that throughout from middle school to university students' knowledge of chemical reactions is limited and characterised by numerous misconceptions (Ahtee & Varjola, 1998). Concept maps can be used to uncover students' misconceptions (Djanette & Fouad, 2014). Following the qualitative approach of Djanette and Fouad (2014), the analysis of misconceptions involves three steps: (1) identification of all technically wrong propositions, (2) inductive identification of common misconceptions which become apparent in these propositions (see Appendix A), and (3) calculation of sum scores representing the number of misconceptions used (so-called misconceptions score). In order to determine interrater reliability, ten percent of the concept maps from pre- and post-test were also randomly chosen and independently scored by a second rater (Döring & Bortz, 2016). Spearman correlation was calculated which indicates acceptable interrater reliability ($r = .74, p < .001$).

Data analysis

We used *SPSS 23* and specified an α level of .05 for statistical analyses.

Results

According to our research question, we were interested whether an incremental scaffolding (treatment group) improves conceptual knowledge acquisition in comparison to a traditional approach (control group). We used two types of scores as indicators of students' conceptual knowledge: (1) The 'proposition accuracy score' refers to the relational scoring method developed by McClure et al. (1999). (2) The 'misconceptions score' represents the number of misconceptions and was suggested by Djanette and Fouad (2014). Conceptual knowledge acquisition should be associated with an increase of 'proposition accuracy scores' and a decrease of 'misconceptions scores' from pre- to posttest. Before testing these hypotheses, a Shapiro-Wilk normality test was performed which indicates non-normality of all variables. However, histograms showed approximately normal distribution of the 'proposition accuracy scores'. Considering the sample size (cf. central limit theorem described by Field [2013]) and the histograms, we applied a t-test and repeated measures analysis of variance to the 'proposition accuracy scores'. In contrast to the 'proposition accuracy scores', the 'misconceptions scores' showed a strong right-skewed distributions in the histograms. According to Bortz and Lienert (2008), we implemented a Solomon four-group design and

performed non-parametric Mann-Whitney U tests. No values deviating more than 3 *SD* from the mean were detected in either group. As students were not randomly assigned to the treatment and control groups, we checked whether groups differed in their conceptual knowledge at the pretest which would limit the comparability of the two groups. A t-test showed that the treatment group ($M = 8.95$, $SD = 4.83$) and the control group ($M = 8.93$, $SD = 5.26$; see Table 1) had comparable ‘proposition accuracy score’ in the pretest, $t(103) = 0.02$, $p = .988$. The same was the case for the ‘misconceptions score’ with comparable mean ranks in the treatment ($M_{\text{Rank}} = 52.17$, $n = 59$) and control groups ($M_{\text{Rank}} = 54.07$, $n = 46$), $U = 1308.00$, $p = .731$ (Mann-Whitney U test).

Table 1
Pre- and posttest mean scores (proposition accuracy scores) and standard deviations as a function of instruction condition (group)

| Group | Pretest | | Posttest | | <i>n</i> |
|-----------|----------|-----------|----------|-----------|----------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | |
| Treatment | 8.95 | 4.83 | 15.58 | 8.88 | 59 |
| Control | 8.93 | 5.29 | 12.43 | 5.97 | 46 |
| Total | 8.94 | 5.01 | 14.20 | 7.86 | 105 |

A repeated measures analysis of variance was executed to investigate whether ‘proposition accuracy scores’ increased from pre- to posttest (within-subject factor ‘time’) and whether this increase differs between the two groups (interaction effect between ‘time’ and ‘group’). Students’ ‘proposition accuracy scores’ significantly improved in both groups from pre- to posttest, $F(1, 103) = 56.87$, $p < .001$, partial $\eta^2 = 0.356$ (large effect; Richardson, 2011, p. 142). The treatment group achieved a higher increase than the control group, $F(1, 103) = 5.42$, $p < .05$, partial $\eta^2 = 0.05$ (medium effect; Richardson, 2011, p. 142). These results support our hypothesis that conceptual knowledge acquisition benefits from a step-by-step support through incremental scaffolds in comparison to a traditional approach, in which students had the opportunity to ask questions to the teacher.

Because of non-normality of the ‘misconceptions scores’, the analysis followed a Solomon four-group design (Bortz & Lienert, 2008). Difference scores between the ‘misconceptions scores’ from pre- and posttest were calculated to describe the rate of change in the treatment and control groups. These scores were used as the dependent variable in a Mann-Whitney U test with group as independent variable. The Mann-Whitney U test showed no significant

difference between the treatment group ($M_{\text{Rank}} = 54.23$, $n = 59$) and the control group ($M_{\text{Rank}} = 51.42$, $n = 46$; $U = 1284.50$, $p = .623$), indicating that incremental scaffolds did not reduce misconceptions more effectively than the traditional approach.

An explorative analysis showed that students had a variety of misconceptions regarding the target-topic ‘chemical reactions’. These misconceptions arose in both groups and included the idea that...

- a “chemical reaction” cannot take place without a “gas burner” (e.g., “chemical reaction needs a gas burner”).
- a “chemical reaction” is “energy” (e.g., “chemical reactions consist of energy”).
- “water” generates “energy” (e.g., “water provides energy”).

Beyond that, students incorrectly combined the concepts “energy”, “heat”, and “activation energy” (examples: “heat generates energy”, “activation energy generates heat”) as well as “exothermic”, “endothermic”, and “heat” (examples: “endothermic is exothermic”, “endothermic emits heat”, “exothermic absorbs heat”)

Discussion

This study investigated whether an incremental scaffolding improves conceptual knowledge acquisition in comparison to a traditional approach, where students require instructional support from the teacher if they have questions. Both approaches were implemented in a chemistry classroom following the concept of inquiry-based learning. Concept maps from the pre- and posttest provided insight into the students’ conceptual knowledge and were evaluated according to the relational scoring method developed by McClure et al. (1999). Scores that emerged from this evaluation express the average accuracy of propositions and their number (so-called proposition accuracy score). Beyond that, the number of misconceptions was determined as suggested by Djanette and Fouad (2014; so-called misconceptions score) to investigate whether incremental scaffolds help to overcome scientifically wrong conceptions (i.e., misconceptions). Both scores allow different statements about the learning efficacy of incremental scaffolds. Whereas the ‘proposition accuracy score’ supports the learning efficacy of incremental scaffolds, the ‘misconceptions score’ does not. At a first glance, this result appears contradictory, but it can be explained by the nature of learning opportunities offered to the students. Whereas these learning opportunities convey the scientific concepts, they neglect students’ misconceptions. Thus, the learning opportunities of this study enable

the acquisition of conceptual knowledge, but the misconceptions of the students remain unchanged due to a lack of learning opportunities that encourage cognitive conflict (cf. Posner, Strike, Hewson, & Gertzog, 1982). It is important to note that the standard deviation of the ‘proposition accuracy scores’ (cf. posttest results of the treatment group in Table 1) noticeably exceeds the standard deviation of the control group. This could indicate that incremental scaffolds are only suitable for certain students. Since the scores in the treatment group are distributed upwards more strongly than downwards compared to the control group, we assume that not all students used the scaffolds correctly or benefited from its use.

Only a few studies examined the effect of incremental scaffolds on learning (van de Pol, Volman, & Beishuizen, 2010). Since they use different strategies in the control group (e.g., worked-out examples and incremental strategic scaffolds in the study of Schmidt-Weigand et al. [2009]), it is very challenging to assess the learning efficacy of incremental scaffolds in the science classroom. As the comparison of incremental scaffolds to a traditional approach of teaching and learning has high ecological validity, our study provides further evidence for the learning efficacy of incremental scaffolds.

Beyond the learning efficacy of incremental scaffolds, an explorative evaluation of the concept maps revealed a series of misconceptions about chemical reactions. One misconception concerns the origin of energy and can be described as energy generation (e.g., “heat generated energy”). This misconception contradicts the concept of energy conversions and is being used in various contexts (cf. Barke, 2006; Opitz, Blankenstein, & Harms, 2016). An additional misconception refers to the relationship between the concepts of activation energy, energy, and heat. Although the students noticed a relation between these concepts, they were not able to describe this relation correctly. As a final point, students had problems to distinguish between endothermic and exothermic reactions (e.g., “exothermic is endothermic”), which can be found in other studies too (e.g., de Vos & Verdonk, 1986; Kind, 2004). Although, the comparison of incremental scaffolds with a traditional approach of teaching and learning has high ecological validity, our study has some limitations:

- The quasi-experimental design of the study restricts its internal validity. However, both groups entered the study with similar conceptual knowledge, suggesting that the conclusions drawn from our study are robust.
- The students were informed about the aim of our study in advance. This may have led to an increased effort of the students (due to extrinsic motivation to impress) and thus might have influenced the results (Rosenthal effect; Rosenthal & Fode, 1963).

- Incremental scaffolds were new to the students. Thus, a novelty effect could have positively influenced the learning efficacy of incremental scaffolds (cf. Kormi-Nouri, Nilsson, & Ohta, 2005).
- Due to a drop-out rate of 20%, we did not achieve the desired power of $1 - \beta = .90$. However, the power is considered sufficient in the range of 80% to 95% (Whitley & Ball, 2002), as is the case for this study.
- Since our study provided a voluntary use of the incremental scaffold for the students, we collected no data about how many students in the treatment group did make use of it. Furthermore, we cannot be sure whether the students used the incremental scaffold in the intended form, or if they just sought for the solution given on the last card too early.

In order to gain more valid statements about the learning efficacy of incremental scaffolds further studies are necessary. These studies should record the actual use of the scaffolds by the students (e.g., by camera observations). Moreover, they should apply incremental scaffolds in various subjects (e.g., in chemistry, physics, or mathematics), groups (e.g., students with and without special need), or social settings (e.g., individual work, partner work, group work) to explore the conditions which are suitable for successful scaffolding. Van de Pol, Volman, Oort, and Beishuizen (2015) found that the learning efficacy of incremental scaffolds depends on how much time students have to complete a particular task. The more time they have, the more scaffolds they use. Furthermore, the long-term embedding of scaffolds in everyday school life should be investigated. This way a potential novelty bias could be examined, and it could be also prevented that student are not able to deal adequately with incremental scaffolds.

It appears noteworthy that incremental scaffolds are not typically utilised in German schools and only a small number of studies have explored the influence of this strategy in inquiry-based learning scenarios. There are various possible reasons why scaffolds are rarely used in the classroom. Some teachers may not be familiar with incremental scaffolds and others may draw back from the effort creating scaffolds. This is suboptimal as incremental scaffolding seems to be a powerful tool in classroom teaching. Giving students the right amount of instructional support is an important challenge in school. Incremental scaffolding gives the students the opportunity to decide for themselves when they need help. This way they can work independently and in a self-regulated manner. Incremental scaffolds can be seen as an efficient way to give the exact help they really need. They also meet the preference of

the students to get anonymous help rather than asking the teacher (Franke-Braun et al., 2008). Therefore, our study could serve as a useful example of how incremental scaffolds could be used to support teaching science.

Acknowledgements

We would like to thank Naemi Welter and Britta Mörs for their comments on an earlier version of the manuscript.

The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

- Ahtee, M., & Varjola, I. (1998). Students' understanding of chemical reaction. *International Journal of Science Education*, 20(3), 305–316.
- Arnold, J., Kremer, K., & Mayer, J. (2017). Scaffolding beim Forschenden Lernen [Scaffolding in inquiry learning]. *Zeitschrift für Didaktik der Naturwissenschaften*, 23(1), 21–37.
- Barke, H. (2006). *Chemiedidaktik: Diagnose und Korrektur von Schülervorstellungen [Chemistry education: Diagnosis and correction of student ideas]*. Heidelberg: Springer.
- Belland, B. R., Glazewski, K. D., & Richardson, J. C. (2008). A scaffolding framework to support the construction of evidence-based arguments among middle school students. *Educational Technology Research and Development*, 56(4), 401–422.
- Bjonness, B., & Kolsto, S. D. (2015). Scaffolding open inquiry: How a teacher provides students with structure and space. *Nordina*, 11(3), 223–237.
- Bolte, L. A. (1999). Using concept maps and interpretive essays for assessment in mathematics. *School Science and Mathematics*, 99(1), 19–30.
- Bortz, J., & Lienert, G. A. (2008). *Kurzgefasste Statistik für die klinische Forschung [Brief statistics for clinical research]* (2nd ed.). Heidelberg: Springer.
- Brandstädter, K., Harms, U., & Großschedl, J. (2012). Assessing system thinking through different concept-mapping practices. *International Journal of Science Education*, 34(14), 2147–2170.
- Brewer, C. A., & Smith, D. (2009). *Vision and change in undergraduate biology education*. American Association for the Advancement of Science & National Science Foundation. Washington. Retrieved July 18, 2021, from <https://www.jsg.utexas.edu/events/files/biology-vision-and-change.pdf>
- de Vos, W., & Verdonk, A. H. (1986). A new road to reactions: Part III. Teaching the heat effect of reactions. *Journal of Chemical Education*, 63(11), 972.

- DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *Canadian Journal for the Scholarship of Teaching and Learning*, 2(2).
- Djanette, B., & Fouad, C. (2014). Determination of university students' misconceptions about light using concept maps. *Procedia Social and Behavioral Sciences*, 152, 582–589.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften [Research methods and evaluation in social and human sciences]*. Heidelberg: Springer.
- Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25(6), 671–688.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage.
- Forghani-Arani, N., Cerna, L., & Bannon, M. (2019). *The lives of teachers in diverse classrooms, OECD Education Working Papers* (Report No. 198), OECD Publishing. Retrieved July 18, 2021, from <https://doi.org/10.1787/8c26fee5-en>
- Franke–Braun, G., Schmidt-Weigand, F., Stäudel, L., & Wodzinski, R. (2008). Aufgaben mit gestuften Lernhilfen [Tasks with incremental scaffolds]. In R. Messner & W. Blum (Eds.), *Lernumgebungen auf dem Prüfstand* (pp. 27–42). Kassel: Kassel University Press.
- Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, 82(3), 300–329.
- Garnett, P. J., & Treagust, D. F. (1990). Implications of research on students' understanding of electrochemistry for improving science curricula and classroom practice. *International Journal of Science Education*, 12(2), 147–156.
- Großschedl, J., & Harms, U. (2013). Assessing conceptual knowledge using similarity judgments. *Studies in Educational Evaluation*, 39, 71–81.
- Großschedl, J., Mahler, D., & Harms, U. (2018). Construction and evaluation of an instrument to measure content knowledge in biology: The CK-IBI. *Education Sciences*, 8(3), 145.
- Haltunen, K. (2003). Scaffolding performance in IR instruction: Exploring learning experiences and performance in two learning environments. *Journal of Information Science*, 29(5), 375–390.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. London: Routledge.
- Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99–107.

- Kind, V. (2004). Beyond appearances: Students' misconceptions about basic chemical ideas (2nd ed.). Durham: Durham University. Retrieved July 18, 2021, from https://www.researchgate.net/publication/228799159_Beyond_Apearances_Students%27_Misconceptions_About_Basic_Chemical_Ideas
- Kormi-Nouri, R., Nilsson, L.-G., & Ohta, N. (2005). The novelty effect: Support for the novelty-encoding hypothesis. *Scandinavian Journal of Psychology*, *46*, 133–43.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212–218.
- Krause, U. M., Stark, R., & Mandl, H. (2004). Förderung des computerbasierten Wissenserwerbs durch kooperatives Lernen und eine Feedbackmaßnahme. [Promotion of computer-based knowledge acquisition through cooperative learning and a feedback activity]. *Zeitschrift für Pädagogische Psychologie*, *18*(2), 125–136.
- Lambon-Ralph, M. A., Pobric, G., & Jefferies, E. (2009). Conceptual knowledge is underpinned by the temporal pole bilaterally: Convergent evidence from rTMS. *Cerebral Cortex*, *19*(4), 832–838.
- Lou, Y., Abrami, P.C., & D'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, *71*(3), 449–521.
- 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.
- Mieg, H. A. (2019). Introduction: Inquiry-based learning – Initial assessment. In A. H. Mieg (Eds.), *Inquiry-based learning – Undergraduate research: The German multidisciplinary experience* (pp. 1–16). Cham: Springer Open.
- Novak, J. D., & Cañas A. J. (2008). The theory underlying concept maps and how to construct them. Pensacola: IHMC. Retrieved July 18, 2021, from <http://cmap.ihmc.us/publications/researchpapers/TheoryUnderlyingConceptMaps.pdf>
- Opitz, S. T., Blankenstein, A., & Harms, U. (2016). Student conceptions about energy in biological contexts. *Journal of Biological Education*, *51*(4), 427–440.
- Pedaste, M., Mäeots, M., Leijen, Ä., & Sarapuu, S. (2012). Improving students' inquiry skills through reflection and self-regulation scaffolds. *Technology, Instruction, Cognition and Learning*, *9*, 81–95.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211–227.
- Richardson, J. T. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147.
- Rosenthal, R., & Fode, K. L. (1963). The effect of experimenter bias on the performance of the albino rat. *Behavioral Science*, *8*, 183–189.

- Saye, J. W., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia supported learning environments. *Educational Technology Research and Development*, 50(3), 77–96.
- Schmidt-Weigand, F., Franke-Braun, G., & Hänze, M. (2008). Erhöhen gestufte Lernhilfen die Effektivität von Lösungsbeispielen? Eine Studie zur kooperativen Bearbeitung von Aufgaben in den Naturwissenschaften [The influence of different presentation modes of worked examples on learning]. *Unterrichtswissenschaft*, 36(4), 365–384.
- Schmidt-Weigand, F., Hänze, M., & Wodzinski, R. (2009). Complex problem solving and worked examples. *Zeitschrift für Pädagogische Psychologie*, 23(2), 129–138.
- Smith, III, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, 3(2), 115–163.
- Sweller, J., van Marrienoer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296.
- Tan, D. K. Y., Koppi, A., & Field, D. J. (2016). First year agricultural science student perspectives in graduate attribute development through problem-based learning. *International Journal of Innovation in Science and Mathematics Education*, 24(1), 54–66.
- The President's Council of Advisors on Science and Technology (2012). *Report to the President, engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Retrieved July 18, 2021, from <https://files.eric.ed.gov/fulltext/ED541511.pdf>
- Thomas, L., Bennett, S., & Lockyer, L. (2016). Using concept maps and goal-setting to support the development of self-regulated learning in a problem-based learning curriculum. *Medical Teacher*, 38(9), 930–935.
- van de Pol, J., & Elbers, E. (2013). Scaffolding student learning: A micro-analysis of teacher–student interaction. *Learning, Culture and Social Interaction*, 2(1), 32–41.
- 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.
- van de Pol, J., Volman, M., Oort, F., & Beishuizen, J. (2015). The effects of scaffolding in the classroom: Support contingency and student independent working time in relation to student achievement, task effort and appreciation of support. *Instructional Science*, 43(5), 615–641.
- Veríssimo Catarreira, S. M., Godinho Lopes, V., Casas García, L. M., & Luengo González, R. (2017). Evaluation of changes in cognitive structures after the learning process in mathematics. *International Journal of Innovation in Science and Mathematics Education*, 25(2), 17–33.
- Wangdi, D., Precharattana, M., & Kanthang, P. (2020). A guided inquiry laboratory to enhance students' understanding of the law of mechanical energy conservation. *International Journal of Innovation in Science and Mathematics Education*, 28(1), 29–43.

- Whitley, E., & Ball, J. (2002). Statistics review 4: Sample size calculations. *Critical Care*, 6, 335–341.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 17(2), 89–100.
- Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman, (Eds.). *Self-regulated learning: From teaching to self-reflective practice* (pp. 1–19). New York: The Guilford Press.

Appendix A

Misconceptions in pre- and posttest concept maps as a function of instruction condition (group)

| Misconceptions | Treatment group (n = 59) | | Control group (n = 46) | |
|---|-----------------------------|-----------------|---------------------------|-----------------|
| | Pretest (%) | Posttest (%) | Pretest (%) | Posttest (%) |
| Chemical reaction consists of (is) energy | 8.5 | 5.1 | 10.9 | 2.2 |
| Chemical reaction needs gas burner | 10.2 | 11.9 | 17.4 | 6.5 |
| Gas burner needs energy | 6.8 | 1.7 | 10.9 | 8.7 |
| Heat generates energy | 22.0 | 6.8 | 10.9 | 6.5 |
| Heat needs energy | 0 | 3.4 | 6.5 | 0 |
| Energy generates heat | 11.9 | 1.7 | 13.0 | 2.2 |
| Energy consists of heat | 5.1 | 0 | 0 | 0 |
| Energy generates activation energy | 3.4 | 0 | 6.5 | 0 |
| Activation energy generates energy | 1.7 | 0 | 2.2 | 2.2 |
| Activation energy generates heat | 3.4 | 1.7 | 0 | 0 |
| Endothermic is exothermic | 8.5 | 0 | 0 | 0 |
| Endothermic emits heat | 0 | 8.5 | 0 | 6.5 |
| Exothermic absorbs heat | 0 | 6.8 | 0 | 6.5 |
| Water generates energy | 13.6 | 8.5 | 23.9 | 8.7 |

Note. The values (in %) represent the proportion of students who use a corresponding misconception in their concept map.

4.3 Concept Mapping – Increased Potential as a Retrieval-based Task

4.3.1 Lernstrategie Concept Maps: Theoretischer Zugang, Hypothesen, Studiendesign

Theoretischer Zugang

Concept Maps, welche ursprünglich als Diagnoseinstrument entwickelt wurden (Novak & Musonda, 1991, siehe auch Studie II), sollen Aufschluss über die kognitive Struktur der Lernenden und damit ihre Wissensorganisation zu einem bestimmten Lerngegenstand geben (Kinchin et al., 2000; Novak & Cañas, 2008). Innerhalb einer Concept Map werden einzelne Konzepte und ihre Zusammenhänge zueinander graphisch dargestellt (Blunt & Karpicke, 2014; Schroeder et al., 2017). Die Konzepte werden als Knoten in der Concept Map (räumlich) angeordnet. Diese Knoten werden über beschriftete Pfeile verbunden, die die Beziehung zwischen den Konzepten kennzeichnen. Zwei, durch einen Pfeil verbundene, Knoten stellen die kleinste Einheit einer Concept Map dar und werden Proposition bezeichnet. Die Pfeilrichtung gibt hierbei die Leserichtung der Proposition vor (Ruiz-Primo & Shavelson, 1996).

Das Konstruieren von Concept Maps (bezeichnet als: Concept Mapping) stellt einen komplexen Prozess dar und erfordert von Lernenden, dass sie sich mit dem Lernmaterial vertieft auseinandersetzen, um die wichtigsten Informationen zu identifizieren, ihre Beziehungen zueinander zu ermitteln und anschließend in einer Concept Map anzuordnen (Schroeder et al., 2017). Concept Mapping stellt für Lernende, welche vormals das Auswendiglernen (*rote learning*, siehe Kapitel 2.1.2) praktizieren (Novak & Cañas, 2008) eine große Herausforderung und damit eine Beanspruchung kognitiver Ressourcen dar (Kriegelstein et al., 2021). Für einen wirksamen Einsatz von Concept Maps wird daher das Üben der Strategie und somit der Erwerb prozeduralen Wissens zu Concept Maps empfohlen (Becker, Welter, Aschermann & Großschedl, 2021).

Concept Maps fungieren sowohl als Organisations- als auch als Elaborationsstrategie (Becker, 2022). So wird für das Concept Mapping nicht nur die (räumliche) Organisation und Strukturierung der Konzepte in einer übersichtlichen Form über eine Concept Map erforderlich (Kinchin et al., 2000), sondern darüber hinaus auch die Implementierung des Vorwissens und die anschließende Verknüpfung von bereits erlernten und neu erworbenen Konzepten (Karpicke & Blunt, 2011). Folglich stellen Concept Maps eine Lernstrategie dar, die zum bedeutungsvollen Lernen beitragen soll (González et al., 2008; Novak, 2002; Novak &

Cañas, 2008) und durch die Entwicklung stabiler Wissensstrukturen den Transfer des Wissens auf neue Kontexte und das Behalten des Wissens über einen längeren Zeitraum ermöglichen soll (Novak & Cañas, 2008).

Concept Maps werden üblicherweise in Anwesenheit des Lernmaterials als Elaborationsstrategie eingesetzt (Blunt & Karpicke, 2014). Allerdings verspricht die Konstruktion von Concept Maps in Abwesenheit des Lernmaterials als Abrufstrategie ein hohes Potential (Blunt & Karpicke, 2014; O’Day & Karpicke, 2020). Es wird angenommen, dass abrufbasiertes Concept Mapping zum Lernen beiträgt, da das Lernmaterial in zweifacher Weise erschlossen wird: Während bei Concept Maps hauptsächlich die Organisation und Verknüpfung beziehungsweise Elaboration von Konzepten erfolgt, könnte das abrufbasierte Concept Mapping zusätzlich den Abruf der Informationen erwirken und so die Lerninhalte tiefer verarbeiten (O’Day & Karpicke, 2020). Da das Praktizieren des Abrufens beziehungsweise das Erinnern von Informationen aus dem Langzeitgedächtnis von Lernenden als schwierig eingeschätzt wird und Lernende aufgrund der Abrufstrategie eine erhöhte kognitive Belastung erfahren können (Blunt & Karpicke, 2014), ist der Vergleich von Concept Maps, die in Anwesenheit des Lernmaterials (*study-based concept mapping*; SCM; deutsch: materialbasiertes Concept Mapping) und von Concept Maps, die in Abwesenheit des Lernmaterials (*retrieval-based concept mapping*, RCM, deutsch: abrufbasiertes Concept Mapping) konstruiert werden, erforderlich. Zwar liegen aus empirischer Sicht zum Concept Mapping (in Form von materialbasiertem Concept Mapping) eine Vielzahl von Befunden vor (Nesbit & Adescope, 2006; Schroeder et al., 2017), zu dem Vergleich von Concept Maps mit Abrufstrategien (Schroeder et al., 2017) oder gar Concept Maps als Abrufstrategie liegen nur wenige empirische Befunde vor (O’Day & Karpicke, 2020).

Hypothesen

Das Lernen in der Biologie erfordert das Erlernen von Zusammenhängen zwischen einzelnen Konzepten (siehe Kapitel 2.1.2). Die Lernstrategie der Concept Maps scheint hierfür eine geeignete Möglichkeit darzustellen (Becker, Welter, Aschermann & Großschedl, 2021). Fraglich bleibt jedoch, ob das abrufbasierte Concept Mapping oder die bewährte Lernstrategie des (materialbasierten) Concept Mappings, Lernende beim Erwerb biologischen Wissens effektiver unterstützen. So zielt nachfolgende Studie darauf ab, die Effektivität abrufbasierter Concept Mappings durch den Vergleich mit materialbasiertem Concept Mapping zu

untersuchen. Die der Untersuchung zugrunde liegenden Hypothesen werden im Folgenden dargestellt.

Das Praktizieren des Abrufs und damit die Abrufstrategie stellt eine aktive Form des Lernens dar (Jonsson et al., 2020). Es wird angenommen, dass das Testen der Lerninhalte während des Lernens (hier: infolge der Entfernung des Lernmaterials), Informationen aus dem Langzeitgedächtnis im Arbeitsgedächtnis verfügbar macht und dort verknüpft (Wiklund-Hörnqvist et al., 2022). Diese aktive Form des Lernens ermöglicht eine tiefere Verarbeitung und Speicherung von Informationen im Gedächtnis (Antony et al., 2017; O’Day & Karpicke, 2020) und führt, empirischen Studien zu Folge, zu einer höheren Lernperformanz im Vergleich zu Elaborationsstrategien (Blunt & Karpicke, 2014; Goossens et al., 2014; Karpicke & Blunt, 2011). Folglich kann angenommen werden, dass abrufbasiertes Concept Mapping zu einer höheren Lernperformanz in einem (deklarativen) Wissenstest führt, im Vergleich zum materialbasierten Concept Mapping. Folgende Hypothese lässt sich somit formulieren:

Hypothese 3.1: Schüler:innen, welche über abrufbasiertes Concept Mapping lernen, zeigen eine höhere Lernperformanz als Schüler:innen, die über materialbasiertes Concept Mapping lernen.

Da der Einsatz von Lernstrategien auch Auswirkungen auf die wahrgenommene kognitive Belastung hat (siehe Kapitel 2.2.1) und somit auch Abrufstrategien die kognitive Belastung beeinflussen können (Blunt & Karpicke, 2014), ist die Untersuchung der verschiedenen Quellen der kognitiven Belastung in den beiden Bedingungen erforderlich. Es kann angenommen werden, dass abrufbasiertes Concept Mapping, im Vergleich zu materialbasiertem Concept Mapping, zu einer Verringerung der extrinsischen kognitiven Belastung führt, da Schüler:innen durch die Abwesenheit des Lernmaterials entlastet werden und somit nicht verschiedene visuelle Quellen verarbeiten müssen (siehe *split attention*, Ayres & Sweller, 2014; Sweller, 2010, siehe auch Kapitel 2.2.1). Zudem kann angenommen werden, dass das abrufbasierte Concept Mapping zu einer höheren intrinsischen kognitiven Belastung führt, da Schüler:innen durch die Abwesenheit des Lernmaterials, mehr Konzepte gleichzeitig im Arbeitsgedächtnis behalten müssen, im Vergleich zum materialbasierten Concept Mapping (Blunt & Karpicke, 2014). Da angenommen wird, dass Abrufstrategien zu einer Steigerung der lernbezogenen kognitiven Belastung führen (Hultberg et al., 2018), sollten Schüler:innen der Bedingung mit dem abrufbasierten Concept Mapping eine höhere lernbezogene kognitive Belastung aufzeigen als Schüler:innen der materialbasierten Concept Mapping-Bedingung. Folgende Hypothesen können somit formuliert werden:

Hypothese 3.2.1: Schüler:innen, die über abrufbasiertes Concept Mapping lernen, schätzen ihre extrinsische kognitive Belastung geringer ein, als Schüler:innen, die über materialbasiertes Concept Mapping lernen.

Hypothese 3.2.2: Schüler:innen, die über abrufbasiertes Concept Mapping lernen, schätzen ihre intrinsische kognitive Belastung höher ein, als Schüler:innen, die über materialbasiertes Concept Mapping lernen.

Hypothese 3.2.3: Schüler:innen, die über abrufbasiertes Concept Mapping lernen, schätzen ihre lernbezogene kognitive Belastung höher ein, als Schüler:innen, die über materialbasiertes Concept Mapping lernen.

Das Wissen von Schüler:innen kann über verschiedene Methoden ermittelt werden (McClure et al., 1999). Besonders gut eignen sich Wissenstest, welche meist über mehrere Fragen, vornehmlich das deklarative Wissen zu einem Lerngegenstand erfassen (Novak & Cañas, 2008; siehe Hypothese 1, Kapitel 4.3). Auch die Analyse der Qualität von Lernprodukten (hier: Concept Maps) kann eingesetzt werden als valide Methode zur Erfassung des (konzeptuellen) Wissens (McClure et al., 1999, Ruiz-Primo & Shavelson, 1996). Die Einschätzung der Concept Map-Qualität erfolgt auf Basis von Propositionen (McClure et al., 1999). Bereits vorangegangene Studien indizieren, dass Schüler:innen, welche Abrufstrategien nutzen, weniger Konzepte nennen (Blunt & Karpicke, 2014). Dies könnte im Rahmen des abrufbasierten Concept Mappings in einer geringeren Anzahl an Propositionen und folglich in einer geringeren Concept Map-Qualität münden (siehe Clausen & Christian, 2012). Folglich kann folgende Hypothese formuliert werden:

Hypothese 3.3: Concept Maps, die unter abrufbasierten Bedingungen erstellt werden, weisen eine geringere Concept Map-Qualität auf als solche, die unter materialbasierten Bedingungen erstellt werden.

Da durch Abrufstrategien Elaborationsprozesse initiiert werden (Eshuis et al., 2022), kann angenommen werden, dass Propositionen in Concept Maps aus der abrufbasierten Bedingung vermehrt Vorwissen inkludieren (Becker, Welter & Großschedl, 2021). Folglich würden in Concept Maps, welche in der abrufbasierten Bedingung erstellt werden, mehr Propositionen generiert, die auf Vorwissen basieren im Vergleich zu Concept Maps aus der materialbasierten Bedingung. Da in nachfolgender Studie mehrere aufeinanderfolgende Lernmaterialien eingesetzt werden, werden Rückschlüsse darüber gezogen, ob die gebildeten Propositionen auf dem Vorwissen der Schüler:innen oder den vorhergehenden Lernmaterialien beruhen. Folgende Hypothesen können formuliert werden:

Hypothese 3.4.1: Schüler:innen, welche mittels abrufbasierten Concept Mapping lernen, generieren mehr Propositionen, welche auf Vorwissen basieren, als Schüler:innen, die mittels materialbasiertem Concept Mapping lernen.

Hypothese 3.4.2: Schüler:innen, welche mittels abrufbasierten Concept Mapping lernen, generieren mehr Propositionen, welche auf vorangegangene Lernmaterialien beruhen, als Schüler:innen, die mittels materialbasiertem Concept Mapping lernen.

Studiendesign

Zur Überprüfung der Wirksamkeit des abrufbasierten Concept Mappings wurden $n = 129$ Schüler:innen der neunten Klasse entweder der abrufbasierten oder der materialbasierten Concept Mapping-Bedingung zugeordnet. Das Interventionsstudiendesign beinhaltete insgesamt vier aufeinanderfolgende Phasen: (1. Phase:) Im Prä-Test wurden demographische Informationen (Alter, Geschlecht) der Schüler:innen, sowie Lesefähigkeiten und das Vorwissen zu Ökosystemen (über einen Wissenstest) ermittelt. In der (2. Phase:) Trainingsphase wurde die Lernstrategie des Concept Mappings sowohl theoretisch eingeführt, als auch praktisch geübt, sodass Schüler:innen prozedurales Wissen zum Concept Mapping erwarben. In der (3. Phase:) Lernphase erhielten die Lernenden zunächst Lernmaterial zum Ökosystem See, welches (für fünf Minuten) gelesen wurde. Nach Ablauf der Zeit durften Schüler:innen der materialbasierten Bedingung ihr Lernmaterial behalten, während Schüler:innen der abrufbasierten Bedingung ihr Lernmaterial (Lerntext) abgaben. In beiden Bedingungen wurden anschließend Concept Maps (für 15 Minuten) erstellt. Nach Ablauf dieser Zeit wurden die Concept Maps eingesammelt und dieser Prozess wurde mit zwei weiteren Lerntexten wiederholt. Insgesamt erhielten die Lernenden somit drei Lerntexte, die sie entweder über materialbasiertes oder abrufbasiertes Concept Mapping bearbeiteten. Anschließend schätzten die Schüler:innen ihre kognitive Belastung ein und führten im Post-Test (4. Phase:), einen Wissenstest zum Ökosystem See durch.

4.3.2 Eigenanteil

Die Studie III wurde von Dr. Sina Lenski konzipiert und durchgeführt. Die Autorin war wesentlich an der Datenauswertung, wesentlich an den Visualisierungen, der Erstellung, Einreichung und Überarbeitung des Manuskripts beteiligt. Prof. Dr. Großschedl überarbeitet das Manuskript und übernahm die Betreuung.

Concept Mapping – Increased Potential as a Retrieval-based Task

Sina Lenski^a^b†, Mirlinda Mustafa^a† and Jörg Großschedl^{a*}

^a*Institute for Biology Education, Faculty of Mathematics and Natural Sciences, University of Cologne, Cologne, Germany;* ^b*German Institute for Adult Education, Bonn, North Rhine-Westphalia, Germany*

† These authors contributed equally and share the first authorship

*Corresponding author

Jörg Großschedl, Institute for Biology Education, Faculty of Mathematics and Natural Sciences, University of Cologne, Cologne, Germany

Email: j.grossschedl@uni-koeln.de

² Publikationsmanuskript wurde lediglich hinsichtlich formaler Aspekte angepasst

Concept Mapping – Increased Potential as a Retrieval-based Task

Concept mapping is a practical task for enhancing learning performance. Learners usually construct concept maps while studying the learning material or after studying. In the first case, the learning material is available during construction, and learners are less involved in retrieval practice from memory (study-based concept mapping; SCM). In the second case, the learning material is absent during construction, and the learners rely on retrieving information from memory (retrieval-based concept mapping, RCM). RCM is assumed to be associated with lower concept map quality and higher cognitive load but better elaboration and learning performance than SCM. This study investigated how the availability of the learning material influenced these variables in biology classrooms. Unlike other studies, this study provided learners with an authentic learning environment and prior concept mapping training. After the concept mapping training, $n = 129$ secondary school students were assigned to an SCM or RCM condition in a quasi-experimental design. As expected, students in the RCM condition constructed concept maps of lower quality but outperformed SCM students concerning elaboration activities and learning performance. The perceived intrinsic cognitive load was higher in the RCM condition. The results indicate that using concept mapping as a retrieval practice could support students' learning in biology.

Keywords: concept mapping, cognitive load, elaboration, retrieval practice, testing effect

Introduction

Concept maps organise information in node-and-link diagrams. Nodes represent concepts, while links represent semantic relationships between concepts (Novak & Gowin, 1984). The smallest meaningful unit of a concept map is a *proposition*. A proposition connects at least two concepts and describes their relationship by a labelled arrow (Novak & Cañas, 2008). Constructing a concept map (concept mapping) is a practical task to promote meaningful learning (Horton et al., 1993; Schroeder et al., 2017). It fosters elaboration processes that help anchor new information in learners' prior knowledge (e.g., Eshuis et al., 2022; Schroeder et al., 2017).

Learners construct concept maps mostly in an elaborative study setting, where the learning material is present during construction (*study-based concept mapping* = SCM; Becker et al., 2021b). If learners construct concept maps in a retrieval practice setting (*retrieval-based concept mapping* = RCM), the learning material is available before but not during the construction process (e.g., O'Day & Karpicke, 2021). Research has shown that

retrieval practice facilitates a more active form of learning than elaborative study because it enables retrieval activities from long-term memory (Chan et al., 2018; Jonsson et al., 2020; Wiklund-Hörnqvist et al., 2022) and promotes memory consolidation (Antony et al., 2017) and meaningful learning (Blunt & Karpicke, 2014; Karpicke & Blunt, 2011; Karpicke & Grimaldi, 2012; O'Day & Karpicke, 2021). Moreover, retrieval practice could offer learners the opportunity to evaluate their own memory performance in a context that resembles the later testing conditions (Jonssen et al., 2021). This evaluative process enables learners to gauge their knowledge more accurately and identify areas of uncertainty (Roediger & Karpicke, 2006a). Thus, retrieval practice could be a key aspect in developing more effective learning strategies (Chan et al., 2018). These strategies, grounded in the principle of *transfer appropriate learning* (see Jonssen et al., 2021), would aid learners in transferring skills, knowledge, or strategies to new or different contexts.

Although research showed that retrieval practice is a more efficient method for learning than elaborative study, learners still perceive elaborative study as more effective (Karpicke & Blunt, 2011; Karpicke & Grimaldi, 2012; Lechuga et al., 2015). Because the learning material is not available during the construction process in retrieval practice, retrieval practice requires continuous search and constant refreshment of information within the working memory (Carpenter, 2009), which increases cognitive load (Blunt & Karpicke, 2014). High cognitive load takes up a significant portion of a learner's cognitive resources. Fewer resources are available for metacognitive processes, such as self-assessment and self-monitoring, so learners may misperceive learning in retrieval practice settings (Karpicke & Grimaldi, 2012). Three types of cognitive load can impact learning (Sweller, 2010), and little research focuses on the type of cognitive load in retrieval practice and elaborative study.

Previous results of Becker et al. (2021b) showed no differences in learning performance between SCM and RCM. However, due to the small sample size in their study, the results should be considered with caution. Other studies have been criticised for the "lack of authentic forms of retrieval-based learning strategies" (Bae et al., 2019, p. 207), as they were conducted under laboratory conditions with university students (e.g., Blunt & Karpicke, 2014). Our study aims to compare the effectiveness of SCM and RCM in an authentic learning environment focussing learning performance, concept map quality, cognitive load, and elaboration activities. Since sufficient concept mapping training is crucial for effective concept mapping (but is rarely implemented in investigations; Becker et al., 2021a; Mintzes et al., 2011), our study provided concept mapping training before the intervention. The results could guide teachers on how to use concept mapping in biology classrooms.

Theoretical Background

Retrieval Practice and Learning

Previous studies have shown that learners practising retrieval perform better than learners practising elaborative study (e.g., Blunt & Karpicke, 2014). While learning tasks that involve elaboration processes during the encoding of information (*elaborative study*) have long been considered particularly conducive to learning, the importance of retrieval processes has been overlooked by learning theories for a long time (Blunt & Karpicke, 2014; McDermott, 2021). Karpicke and O’Day (in press) hypothesised that retrieval practice is particularly effective for learning because it improves memory search processes by forming and reinforcing memory traces that facilitate future knowledge access. According to the *elaborative retrieval hypothesis* (Carpenter, 2009, 2011; Rawson et al., 2015), it can be assumed that the retrieval of target information activates further semantically related information. This process is thought to “make target knowledge more distinctive and retrievable in the future” (Karpicke & O’Day, in press, p. 37). Despite scientific evidence of the effectiveness of retrieval practice, learners often do not perceive retrieval practice as a learning opportunity. This is supported by findings showing that learners rarely incorporate retrieval practice in their learning routines (Hui et al., 2022; Karpicke et al., 2009) as processing during retrieval practice may not be perceived as “fluent and easy” (Blunt & Karpicke, 2014, p. 857). These results suggest that learners experience greater cognitive effort due to retrieval practice.

Cognitive load and retrieval-based concept mapping

Understanding human cognition is crucial for organising learning processes and environments effectively. *Cognitive load theory* focuses on the limited capacity of working memory (Leahy & Sweller, 2019; Sweller et al., 1998) and identifies three sources of cognitive load which lead to three types of load: extraneous, intrinsic, and germane cognitive load (Sweller, 2005).

Unfavourable features of the learning environment that are not directly related to the learning goals (e.g., distractions, confusing instructions, or irrelevant information) cause *extraneous cognitive load* (Sweller et al., 1998). When learners try to comprehend information from multiple sources simultaneously, such as texts and diagrams, these sources compete for attention, which is called the *split-attention effect* (Sweller, 2010). This means that in the context of RCM, the absence of learning material during concept mapping could help reduce unnecessary extraneous cognitive load compared to SCM.

Intrinsic cognitive load refers to the mental effort needed to understand a particular

learning topic, which is determined by its complexity and the level of expertise needed to comprehend it (Sweller, 2010). The complexity of the topic is referred to as "element interactivity", which is the number of interacting components that must be processed simultaneously in working memory (Hanham et al., 2017; Sweller et al., 1998). The more elements that interact with each other, the higher the intrinsic cognitive load. Regarding RCM, the absence of learning material during concept mapping could increase intrinsic cognitive load because learners must simultaneously keep more information in their minds when the learning material is unavailable.

Germane cognitive load is associated with the meaningful processing and integration of information into long-term memory. It relates to the effort required to organise and make sense of information, create connections between new and existing knowledge, and engage in meaningful learning. In essence, germane cognitive load is the mental effort contributing to learning and understanding. Retrieval practice serves as a strategy that is expected to effectively increase germane cognitive load (e.g., Hultberg et al., 2018). Becker et al. (2021b) have shown that learners link new concepts more often to existing knowledge in RCM than in SCM, which could indicate a higher germane cognitive load in RCM.

Elaboration processes in RCM and SCM and concept map quality

Learning in biology means dealing with abstract concepts that are interconnected in multiple ways. Thus, it means understanding connections/interrelationships and requires the organisation and elaboration of knowledge (Becker et al., 2021b). Elaboration activities help learners connect prior knowledge with new information (Kalyga, 2009) and thus have long been considered the core learning strategy for understanding interrelationships (Weinstein & Mayer, 1986). Elaborative study should lead to more propositions in learners' concept maps. In contrast, retrieval practice has been shown to result in a reduced number of concepts and propositions (see Blunt & Karpicke, 2014). This can be explained as follows: When the source of new information (e.g., the text or learning material) is removed, learners must first retrieve that information from long-term memory to make it available in working memory. These memory processes establish an organisational structure in which information is more easily retrieved, and corresponding concepts are more easily discriminated (Karpicke & Blunt, 2011). Although SCM is associated with more propositions in learners' concept maps, a recent study has shown that the number of propositions linking new concepts (concepts from the learning material) to prior knowledge concepts (*elaboration-suggesting propositions*) is significantly higher in RCM than in SCM concept maps (Becker et al., 2021b). The

authors concluded that SCM strongly binds learners' attention to the learning material, while RCM encourages them to activate prior knowledge and thus activate elaboration processes.

Hypotheses

Our study examines the effect of RCM and SCM on secondary school students' learning performance, concept map quality, cognitive load, and elaboration activities. We hypothesise that students who practice RCM will outperform students who practice SCM regarding *learning performance* (**H1**). In addition, we expect that students who practice RCM will experience lower *extraneous cognitive load* (**H2.1**), higher *intrinsic cognitive load* (**H2.2**), and higher *germane cognitive load* (**H2.3**) than students who practice SCM. We also hypothesise that the *concept map quality*, which considers the number of propositions, will be lower for RCM than SCM concept maps (**H3**). Finally, we expect that RCM leads to more elaboration activities than SCM: We assume that RCM results in more propositions that connect new concepts (concepts from the learning material [in our case: learning texts]) to prior knowledge concepts (*propositions based on prior knowledge*). By conducting multiple runs with different learning texts, we can examine those propositions that connect new concepts from a current learning text with concepts from previous learning texts (*propositions based on previous learning opportunities*). We expect to find more *propositions based on prior knowledge* (**H4.1**) and *more propositions based on previous learning opportunities* (**H4.2**) in the RCM condition than in the SCM condition.

Materials and methods

This study was conducted in line with current school law agreements (North Rhine-Westphalian Ministry of Education Science and Research, 2005) and the ethical principles and guidelines for protecting human subjects in research (Department of Health, 2014). Among others, this means that we informed participants about the study procedure and the possibility of stopping the study at any time. We obtained written informed consent from parents and school principals prior to the study and ensured data protection.

Design and procedure

Our study was structured as a quasi-experimental intervention, divided into four main phases (see Figure 1): pre-testing, training, learning, and post-testing. Our instructors carried out these four phases during three visits to the classroom. A different pair of instructors appeared on each visit. This approach not only minimised scheduling conflicts but also guaranteed

that the study was executed uniformly and consistently in every session. Additionally, this core team of six instructors was supported by a reserve duo, ready to step in as needed, ensuring uninterrupted and smooth progress throughout the study. The study was conducted over a period of three weeks for each participating group. Due to organisational constraints within the school, it was not feasible to maintain a consistent time interval between all phases, but an interval of precisely two days between the second and third visit was ensured. At the first classroom visit, the pre-testing phase was implemented to gather demographical information about the students, their reading fluency, and prior knowledge about ecosystems (35 min). The remaining time of the lesson was used to introduce the concept map training phase (10 min). At the second classroom visit, the training phase continued for a further 90 min. The training phase provided declarative knowledge about concept mapping and offered intensive practice opportunities for this task (procedural knowledge). The training itself consisted of a slightly modified version of the training by Lenski et al. (2022) for junior high school students and incorporated training design recommendations from a broad spectrum of educational research (Großschedl & Tröbst, 2018; Holley & Dansereau, 1984; Lechuga et al., 2015; Mintzes et al., 2011). At the third classroom visit, the learning phase was administered and consisted of three sections in which students studied the topic “Lake Ecosystem” using learning texts on “Living Organisms in a Lake” (section I), “Zones of a Lake” (section II), and “Limnetic Zones in a Lake” (section III) in consecutive order (60 minutes in total). The learning texts ranged from 177 to 206 words each and were taken from a study by Lenski et al. (2022) with slight modifications in sentence order and grammar (see supplementary material 1). According to the LIX (readability index; Lenhard & Lenhard, 2022), learning text I shows a moderate complexity (LIX = 49.8), learning texts II (LIX = 42.19) and III (LIX = 39.32) show a low complexity. In each section (I-III), students had five minutes to read the learning text. Since 9th-grade students can read about 240 words per minute (Christodoulou et al., 2014), five minutes were considered sufficient to read a text with a maximum of 206 words twice. Afterwards, students had 15 min to graphically represent the content of the learning text in a concept map. The learning text contained highlighted concepts that students could use and expand with further concepts. Additionally, possible linking phrases were underlined, a technique that was introduced during the concept map training. While students in the SCM condition constructed concept maps with the learning text present throughout the construction process, in the RCM condition, the learning texts were taken away from students at the end of the reading time. In both conditions, students were encouraged to use the entire time to create their concept maps and to review and refine

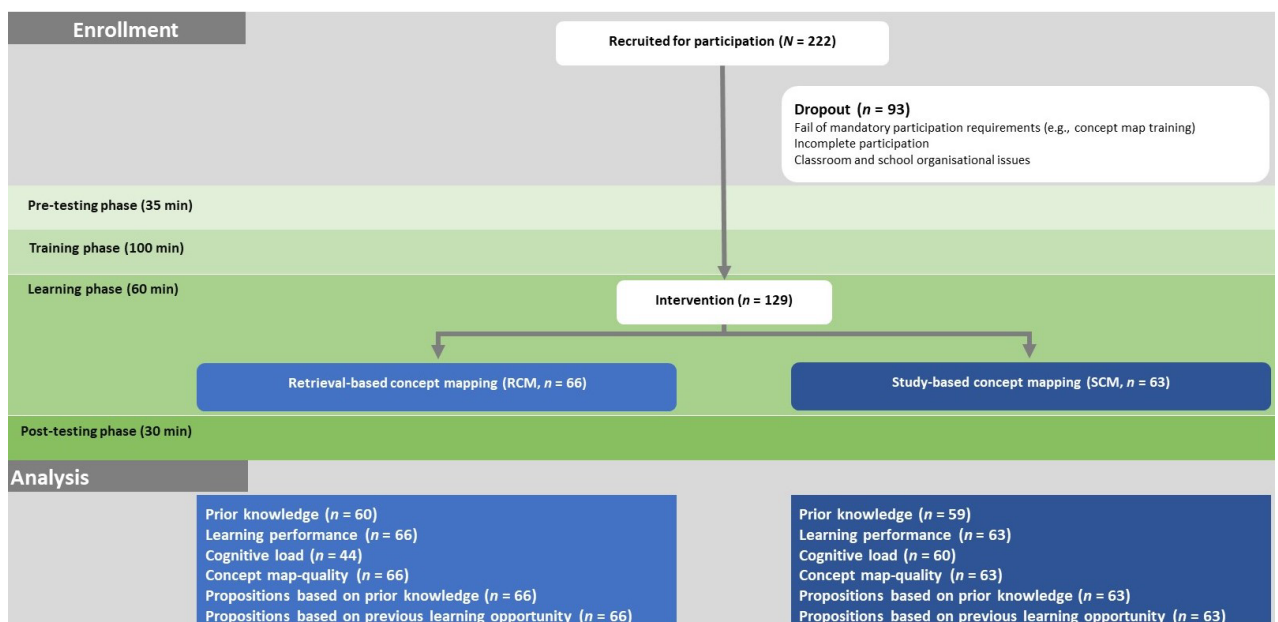
them as needed. At the end of the learning phase, the concept maps were collected and kept for later assessment. However, the students did not receive any feedback on the quality of their concept maps. The topic of the learning phase was not the subject of an upcoming examination or part of the curriculum in any of the classes. Since a random allocation of individual students to the SCM and RCM conditions was impossible because of school organisational reasons, we randomly assigned classes. After the learning phase, students rated their perceived cognitive load. Finally, the post-testing phase (30 min) was initiated to measure learning performance.

Participants

Nine classes with a total of $N = 222$ secondary school (German school type: *Gymnasium*) students from five different schools in North Rhine-Westphalia (Germany) were recruited for our study. Ninety-three students were unable to complete the study due to various reasons, including failure to meet mandatory participation requirements (such as the concept map training), incomplete participation, and organisational issues within classrooms and schools. **Figure 1** overviews student allocation, exclusion criteria, and the variables analysed. A total of $n = 129$ 9th-grade students completed the entire study (48.8 % female, age: $M = 14.18$, $SD = 0.55$; maximum classroom size: 33 students; RCM in 3 classes; SCM in 4 classes).

Figure 1

Flowchart showing participant allocation and procedure. This figure was designed based on the criteria recommended by the CONSORT guidelines (Moher et al., 2012)



Instruments

Reading Fluency

We captured students' reading fluency through the Salzburg Reading Screening *SLS 2-9* (Wimmer & Mayringer, 2014) to check whether differences in reading speed existed between groups. The *SLS 2-9* is a classroom-administered standardised reading speed task (test-retest reliability for normative sample in grade 8 [$r = .87$]). It contains sentences with a simple semantic and syntactic structure that the students should read in silence (e.g., "Apple pie is made from strawberries."). After a short introduction and an exercise phase, students had to read 100 sentences in three minutes and mark them as correct or incorrect. Based on the number of correctly categorised sentences, a raw score was determined, which was transferred to a "reading quotient" based on age and gender norms differences using the *SLS 2-9* manual.

Prior knowledge

To assess students' prior knowledge of the topic "Ecosystem", we utilised a 20-item instrument that included single-correct, multiple-correct, and matching items. The instrument was a modified version of Lenski et al. (2022) and consisted of factual and conceptual knowledge items (see supplementary material 2). Correct responses were scored 1, while incorrect or missing responses were scored 0. We excluded two items with low item discrimination (below .20; Durrheim & Tredoux, 2004) and calculated sum scores ranging from 0 to 18. The scale showed sufficient reliability with Cronbach's $\alpha = .69$ (Rost, 2013).

Learning performance

We measured learning performance using a modified test created by Lenski et al. (2022; see supplementary material 1). This test covered similar topics and types of knowledge as the prior knowledge test but included more challenging items to reduce practice effects, cueing and recall bias (Brennan et al., 2006). The test consisted of questions on the knowledge conveyed in the learning texts (I. - III.) The items were coded dichotomously, and two items had to be removed due to low item discrimination (below .20; Durrheim & Tredoux, 2004). The reliability of the test scores was good with a Cronbach's α score of .79 (Rost, 2013).

Cognitive load

Cognitive load was assessed using an eight-item self-report instrument developed by Klep-sch et al. (2017). As recommended by the authors, we adapted the items by replacing the

placeholder “the task” with “the concept map” (e.g., “When constructing concept maps, it was exhausting to find the important information.” [*extraneous cognitive load*]; “When constructing concept maps, many things must be kept in mind simultaneously.” [*intrinsic cognitive load*]; “When constructing concept maps, I had to engage myself highly.” [*germane cognitive load*]). Items were rated on a 7-point Likert scale from “I fully disagree” to “I fully agree”, with higher points representing higher *cognitive load*. Sum scores were calculated and showed acceptable reliability according to Rost (2013) for the subscales *extraneous cognitive load* (three items, Cronbach’s $\alpha = .66$), *intrinsic cognitive load* (two items, Cronbach’s $\alpha = .65$), and *germane cognitive load* (three items, Cronbach’s $\alpha = .69$).

Concept map quality

When learners consciously link new information to concepts stored in long-term memory, meaningful learning occurs (Karpicke & Grimaldi, 2012). Concept maps represent the learner’s hypothetical memory structure and are a suitable outcome measure for meaningful learning (Ekinci & Sen, 2020). To assess meaningful learning, concept map quality is usually determined by the number and correctness of the concept map propositions (Austin & Shore, 1995; McClure et al., 1999). Clausen and Christian’s (2012) scoring system adds the type of proposition to these evaluation criteria. It is used in our study as this type of scoring is recommended for younger students.

Clausen and Christian (2012) refined the scoring scheme initially developed by McClure et al. (1999). They adapted it to the characteristics of biological content knowledge, in which cause-and-effect relationships play a particular role. The adapted scoring scheme emphasises the semantic coherence between paired elements, such as terms or concepts (for a detailed description of the scoring method, see supplementary material 3). The cumulative sum of the scores for all propositions within a concept map yields an overall score, which reflects the map’s coherence and accuracy in representing conceptual relationships (for a scoring of an exemplary concept map, see supplementary material 3).

To determine interrater reliability, two independent raters scored a proportion of 13 % of the concept maps (see Döring & Bortz, 2016). Cohen’s κ was calculated with the Cohen’s κ Calculator (Hemmerich, 2019) using sum scores. Interrater reliability was substantial to almost perfect (Landis & Koch, 1977), with Cohen’s κ values of .77 for section I maps, .76 for section II maps, and .82 for section III maps of the learning phase. Each learner’s overall concept map quality was calculated as the mean of the three sum scores.

Elaboration in the learning process

Elaboration was determined from the propositions of the concept maps. Propositions were identified that link the new information from the learning texts to prior knowledge (*propositions based on prior knowledge*). Since the students created a total of three concept maps (in learning phases I, II, and III), it was also necessary to identify propositions that link information from the previous learning texts to information from the current learning text (*propositions based on previous learning opportunities*). For the evaluation, all possible propositions from the learning texts were first extracted. Then students' concept maps were analysed by identifying propositions that do not stem from the present learning text. In the following all propositions were evaluated by a single rater, who used a dichotomous categorisation system, classifying the propositions as (1) *propositions based on prior knowledge* or (2) *propositions based on previous learning opportunities*. To determine interrater reliability, 20% of the propositions were analysed by a second rater (see Döring & Bortz, 2016). Cohen's κ was calculated using a contingency table using Cohen's κ Calculator (Hemmerich, 2019). Our results showed a Cohen's κ value of 0.74, indicating substantial interrater reliability according to Landis and Koch (1977). Finally, sum scores were calculated for propositions based on prior knowledge and propositions based on previous learning opportunities for each student across all the concept maps.

Results

Data Analysis

Data analysis was performed using SPSS (version 24.0). Depending on the statistical quality of the data (e.g., normal distribution, variance homogeneity), parametric or nonparametric tests were performed using an alpha level of .05 unless stated otherwise. In addition, the Bayes factors were calculated using JASP (version 0.18.3).

Preliminary tests and statistical analyses

We checked for baseline differences between the SCM and RCM conditions regarding age, reading fluency, prior knowledge, and gender distribution. We set the α -level for these analyses to .10 to detect even minor potential differences between conditions (see Döring & Bortz, 2016). A Mann-Whitney-U-test showed no difference for age, $U= 1575.00$, $Z= -1.17$, $p = .243$. Independent sample t-tests showed no statistically significant differences between conditions in reading fluency, $t(113) = 1.35$, $p = .179$, and prior knowledge, $t(117) = 0.41$, $p = .684$. Also, gender distribution did not reveal a statistically significant discrepancy between

conditions, $\chi^2(1) = 0.00, p = .990$, examined by a chi-squared test (see Table 1 for descriptive data). For all dependent variables, we ran further tests for standard distribution and outlier identification ($> 3x$ interquartile range; Ghasemi & Zahediasl, 2012).

Table 1

Baseline tests and descriptive statistics of control variables by conditions

| Control variables | SCM | | | RCM | | | <i>p</i> |
|--|----------|----------|-----------|----------|----------|-----------|----------|
| | <i>n</i> | <i>M</i> | <i>SD</i> | <i>n</i> | <i>M</i> | <i>SD</i> | |
| Age (years) | 58 | 14.10 | 0.48 | 60 | 14.25 | 0.60 | .243 |
| Reading fluency | 56 | 96.04 | 16.17 | 59 | 99.90 | 14.43 | .179 |
| Prior knowledge (Score _{min-max} = 1-18) | 59 | 12.81 | 2.77 | 60 | 12.60 | 2.93 | .684 |
| Gender | | | | | | | |
| Male | 31 | | | 32 | | | |
| Female | 27 | | | 28 | | | .990 |

Note. SCM = study-based concept mapping; RCM = retrieval-based concept mapping. Score min-max of prior knowledge in the observed values.

Table 2

Descriptive statistics of dependent variables by condition.

| Dependent variables | SCM | | | RCM | | | <i>p</i> |
|---|----------|----------|-----------|----------|----------|-----------|----------|
| | <i>n</i> | <i>M</i> | <i>SD</i> | <i>n</i> | <i>M</i> | <i>SD</i> | |
| Learning performance (Score _{min-max} = 1-16) | 59 | 10.83 | 3.22 | 60 | 12.87 | 2.95 | <.001 |
| Cognitive load (Score _{min-max} = 1-7) | | | | | | | |
| ECL | 44 | 3.13 | 1.44 | 60 | 2.82 | 1.43 | .201 |
| ICL | 44 | 3.76 | 1.45 | 61 | 4.35 | 1.40 | .038 |
| GCL | 44 | 4.96 | 1.35 | 61 | 5.33 | 1.27 | .182 |
| Concept map quality | 63 | 36.76 | 19.96 | 66 | 27.86 | 15.66 | .005 |
| Number of propositions | 63 | 14.53 | 6.65 | 66 | 11.64 | 4.39 | |
| PPK | 63 | 0.10 | 0.30 | 66 | 0.50 | 0.97 | .006 |
| PPL | 63 | 0.03 | 0.25 | 66 | 0.36 | 0.78 | <.001 |

Note. SCM = study-based concept mapping; RCM = retrieval-based concept mapping; ECL = extraneous cognitive load; ICL = intrinsic cognitive load; GCL = germane cognitive load; PPK = propositions based on prior knowledge; PPL = propositions based on previous learning opportunities.

Hypotheses testing

Learning performance (H1)

In line with our first hypothesis, a one-factorial analysis of covariance (ANCOVA) showed, after adjusting for prior knowledge (covariate), that the RCM condition ($M = 12.87, SD = 2.95$) outperformed the SCM condition ($M = 10.83, SD = 3.22$) concerning learning

performance, $F(1, 118) = 28.43, p < .001, \eta^2_p = .197$ (support of **H1**). Supporting this, the Bayes Factor ($BF_{10} = 55.58$) indicates very strong evidence in favour of the alternative hypothesis over the null hypothesis, suggesting that the observed data are over 55 times more likely under the alternative hypothesis. This assessment aligns with the guidelines for interpreting Bayes Factors, suggesting substantial support for our findings (Hemmerich, 2020; Jeffreys, 1961).

Cognitive load (H2.1 – H2.3)

We expected that the presence or absence of the learning text during the construction of concept maps would influence perceived cognitive load. To test this, we conducted Mann-Whitney-U-tests on the three sub-scales of cognitive load. Our analysis revealed no significant differences between conditions in terms of extraneous cognitive load, $U = 1126.50, Z = -1.28, p = .201, r = -.125$ (lack of support for **H2.1**), and germane load, $U = 1137.50, Z = -1.33, p = .182, r = -.130$ (lack of support for **H2.3**). However, intrinsic cognitive load was higher in the RCM condition ($M = 4.35, SD = 1.40$) compared to the SCM condition ($M = 3.76, SD = 1.45$), $U = 1024.00, Z = -2.08, p = .038, r = -.203$ (support for **H2.2**). The effect size for intrinsic cognitive load suggests a small to medium practical impact. Regarding the Bayes Factor analysis, BF_{10} values for extraneous (0.97) and germane cognitive load (0.59) suggest anecdotal evidence in favour of the null hypothesis, indicating that the presence of the learning text does not significantly affect these aspects of cognitive load. This level of evidence, often termed “anecdotal”, indicates that while the data lean towards supporting the alternative hypothesis, further research or data might be necessary to firmly establish these findings (Hemmerich, 2020). Conversely, the BF_{10} for intrinsic cognitive load (1.98) provides anecdotal evidence in support of the alternative hypothesis, implying a more nuanced effect. This nuanced evidence suggests that while the absence of the text may not impact all forms of cognitive load equally, it does have a discernible effect on the intrinsic cognitive load associated with task comprehension and processing (Hemmerich, 2020).

Concept map quality (H3)

In alignment with our third hypothesis, which posited that concept map quality would differ between conditions, our findings revealed a superior concept map quality in the SCM condition ($M = 36.76, SD = 19.96$) compared to the RCM condition ($M = 27.86, SD = 15.66$). Conducting an independent samples t-test, we observed that this difference was statistically significant, $t(127) = -2.83, p = .005$, indicating a moderate effect size of $d = .498$ (support of

H3). Importantly, the Bayes Factor ($BF_{10} = 13.22$) suggests strong evidence in favour of the alternative hypothesis, reinforcing the conclusion that the quality of concept maps is indeed higher in the SCM condition as compared to the RCM condition (Hemmerich, 2020).

Elaboration in the learning process (H4.1 – H4.2)

According to our hypotheses, specifically H4.1 and H4.2, which anticipated differences between conditions in the number of propositions indicating elaboration processes, our descriptive analysis revealed that the RCM condition generated a greater number of *propositions based on prior knowledge* ($M = 0.50, SD = 0.97$) and *based on previous learning opportunities* ($M = 0.36, SD = 0.78$) compared to the SCM condition ($M = 0.10, SD = 0.30$ and $M = 0.03, SD = 0.25$, respectively). Subsequent Mann-Whitney-U-tests confirmed these observations as statistically significant, with $U = 1680.00, Z = 2.77, p = .006, r = -.244$ for propositions based on prior knowledge, and $U = 1612.50, Z = -3.74, p < .001, r = -.330$ for propositions based on previous learning opportunities, thereby providing robust support for both **H4.1** and **H4.2**. Notably, the Bayes Factor analysis, yielding $BF_{10} = 2.25$ for propositions based on prior knowledge and $BF_{10} = 2.51$ for those based on previous learning opportunities, suggests modest evidence in favour of the alternative hypothesis (Hemmerich, 2020).

Discussion

In our study we investigated the impact of retrieval-based concept mapping (RCM) and study-based concept mapping (SCM) on cognitive processes within an authentic educational setting. Following Schmuckler's (2001) guidelines to preserve ecological validity, we initiated our research with a comprehensive training phase. This preparatory step ensured participants were well-versed in concept mapping techniques prior to engaging in concept mapping.

Learning performance

Previous studies have consistently highlighted the learning effectiveness of retrieval practice, illustrating its benefits across various domains (Karpicke & Blunt, 2011; Karpicke & Roediger, 2008; O'Day & Karpicke, 2021; Roediger & Karpicke, 2006a). In line with these findings, our study shows that learning biological content is more effective in RCM than in SCM (see **H1**). According to O'Day and Karpicke (2020), concept maps serve as visual and spatial tools organising concepts. When paired with retrieval practice, concept maps not only

allow for the active reconstruction of the learning text but also facilitate its reorganisation, enhancing learning outcomes.

Supporting the effectiveness of retrieval practice, Carpenter's *elaborative retrieval hypothesis* (2009) proposes that retrieval practice stimulates the activation of semantically related knowledge in long-term memory, promoting the linkage of new information through the activation of semantic associates. This coactivation is believed to contribute to enhanced memory consolidation, thereby improving retention (Antony et al., 2017; Carpenter, 2009; Roelle & Nücker, 2019). Such an elaborative retrieval process seems particularly pertinent to concept mapping, where the integration of new information into existing knowledge frameworks is crucial (Blunt & Karpicke, 2014). Evidence from both, traditional assessments (Rowland, 2014) and neuroimaging studies (Jonsson et al., 2021), underscores the role of activated semantic networks in learning. While recent research into cortical network activity provides intriguing insights (Hong et al., 2023), a more direct discussion of the mechanisms behind the observed learning benefits of RCM over SCM may offer greater clarity. Specifically, the interaction between concept mapping and retrieval practice likely facilitates an elaborative process, enabling easier schema construction and integration of new information into long-term memory. This assumption is aligned with findings that concept mapping benefits learning by supporting the organisation and representation of knowledge (O'Day & Karpicke, 2020), and that the addition of a retrieval component could further enhance these benefits by activating information in long-term memory.

Another possible explanation of the advantage of RCM over SCM in learning performance could be offered by the *expertise-reversal effect* (Rey & Buchwald, 2011). This effect describes the phenomenon where learning strategies that are effective for novices, become less effective (or even counterproductive) for learners with higher knowledge and expertise. Since our participants show a good baseline of prior knowledge, it could be assumed that they may not benefit from less-demanding learning strategies (in this case: SCM). This is in line with the *desired difficulty effect*, where learning strategies of moderate difficulty are assumed to promote greater learning gains compared to less demanding tasks (Roediger & Karpicke, 2006b; Schmidt & Bjork, 1992). Thus, the RCM condition may have provided a more demanding learning environment that might have foster deeper stages of learning.

Future investigations could further explore how retrieval practice through concept mapping (i.e., without direct access to learning materials) not only influences learning outcomes but may also amplify learners' perceptions of their test preparedness. Moreover, it

could be investigated how learners' perceptions of effort and task difficulty influence their engagement and the effectiveness of retrieval-based learning strategies. This exploration into students' self-assessment capabilities and identification of transfer-appropriate processes could enrich the understanding of the mechanism of retrieval practice and its potential benefits in educational settings.

Cognitive load

As expected (**H2.2**), our findings revealed a higher intrinsic cognitive load in the RCM condition. This could be explained by the high element interactivity which is characteristic for retrieval-based learning (Buchin, 2021). Thus, for the understanding of the used learning texts, several concepts need to be held in working memory at the same time (e.g., understanding photosynthesis: photosynthesis requires water, carbon dioxide, and sunlight; photosynthesis is produced by autotrophic organisms; autotrophic organisms are producers, etc.). This aligns with the principles of the cognitive load theory, which posits that tasks with higher element interactivity demand more cognitive resources, thereby explaining the observed increase in intrinsic cognitive load within the RCM condition (see Sweller, 2010). It has been observed that learners with minimal or no prior knowledge tend to struggle more when faced with tasks characterised by high element interactivity, leading to poorer performance (Chen et al., 2017). Therefore, future research could explore the effectiveness of RCM across learners possessing varying degrees of prior knowledge.

It is important to acknowledge that the evidence supporting the effect of RCM on intrinsic cognitive load was relatively modest in our study, urging a cautious interpretation of these results. Additionally, there is a possibility that our measurement tool lacked the finesse to accurately differentiate between extraneous and intrinsic cognitive load, a limitation highlighted by Klepsch and Seufert (2020). In situations where extraneous and intrinsic cognitive load are intricately connected, Klepsch and Seufert (2020) recommend the use of more advanced measurement techniques. Furthermore, it's worth noting that certain studies suggest intrinsic and germane cognitive loads may have a common theoretical basis (Jiang & Kalyuga, 2020). Consequently, these studies advocate for a focus primarily on distinguishing between two loads and suggest to focus on extraneous and intrinsic cognitive loads (Kalyuga, 2011).

Contrary to our hypothesis (**H2.1**), RCM did not lead to a lower extraneous cognitive load compared to SCM. We assumed that RCM would reduce extraneous cognitive load because learners would not have to divide their visual attention between the learning text

and the concept map in this learning setting (Ayres & Sweller, 2014; Mayer, 2002, 2014; Sweller, 2010). The underlying assumption was that extraneous cognitive load is increased when learners have to use several sources of information (Chen et al., 2017) to complete the task (in this case, to create the concept map). As students in the SCM condition have a learning text present to create a concept map, one could argue that the learning text already provides all necessary propositions that only need to be transferred into a concept map. In consequence, this could reduce (intrinsic) cognitive load. Nonetheless, students in the SCM condition must divide their attention between the text and the map, potentially resulting in extraneous load. The competing cognitive effects in the SCM condition together with methodological shortcomings outlined above could explain why we observed no discernible difference in extraneous cognitive load between the two conditions. Consequently, there is a need for further research using more sensitive measures of cognitive load in retrieval-based tasks.

According to hypothesis **H2.3**, RCM should increase germane cognitive load compared to the SCM. Although we observed more elaboration activities in RCM (see **H4**), which suggests a higher germane cognitive load (Klepsch & Seufert, 2020), we had to reject hypothesis **H2.3**. Students' effort is essential for germane cognitive load since they do not automatically use all their available cognitive capacity in a learning task (Schnotz & Kürschner, 2007). Learners perceive learning tasks with retrieval practice as difficult (Blunt & Karpicke, 2014), so it is reasonable to assume that they are not willing to invest their free mental capacities in RCM. This could mask the advantages of RCM over SCM in perceived germane cognitive load. In addition, the familiarity and comfort with the study-based task (SCM) might lead students to engage more efficiently, without the perceived strain associated with retrieval efforts in RCM (Poppenk et al., 2010). This familiarity effect could bias learners towards allocating their efforts more readily to SCM, despite the potential cognitive benefits of RCM. However, since learning performance was observed to be higher in RCM, this hypothesis seems less likely, suggesting that germane cognitive load may not directly correlate with learning outcomes in this study.

The observed unexpected patterns across different types of cognitive load hint at a nuanced interaction that may not be directly predictable solely by cognitive load theory. Furthermore, it seems to challenge the capabilities of measurement instruments derived from this theory. It also underscores the need for future research into understanding cognitive processes involved in retrieval-based learning strategies.

Concept map quality

As hypothesised in **H3**, the concept maps from the SCM condition were of higher quality than those from the RCM condition. We assumed that the presence of the learning text during SCM would allow students to read concepts in the learning text and integrate them directly into the concept map. This integration would lead to a higher number of correct propositions and to a higher concept map quality (see McClure et al., 1999) in the SCM compared to the RCM condition.

There could be multiple reasons explaining that learners in the RCM created concept maps of low quality while outperforming SCM learners in the post-test. Initially, this outcome could be linked to the evaluation method used for concept map quality, where the quantity of propositions significantly influences the assessed quality of the concept map (Clausen & Christian, 2012). While students of the RCM condition may retain only the most important concepts and integrate them into their concept map, students in the SCM condition may transfer the concepts from the learning text into the concept map and thus achieve a higher number of propositions, which results in higher concept map quality. Underpinning this phenomenon could be the inherent requirement of RCM for learners to prioritise and retain only the most salient concepts, a strategy that, while reducing proposition quantity, may enhance conceptual relevance and depth. This aligns with findings suggesting that retrieval practice promotes a more selective, meaningful integration of knowledge (Blunt & Karpicke, 2014), necessitating students to restructure their understanding and distinguish between important and less important concepts (Congleton & Rajaram, 2012). Moreover, the execution disparity between the conditions underscores a distinct cognitive demand in the RCM condition, where integrating new information with existing knowledge necessitated additional cognitive resources, potentially reducing the representational quality of concept maps (McClure et al., 1999). The manual process of creating maps with paper and pencil might have added complexity to the task, especially for RCM, where frequent revisions and reliance on memory play a crucial role. Contrarily, SCM participants, supported by immediate text reference, likely faced fewer obstacles from this manual mapping process. This suggests a potential advantage for digital concept mapping tools for retrieval tasks, as they could equalise the operational challenges between conditions. Future investigations would benefit from comparing concept map quality of RCM and SCM conditions within digitalized mapping environments (such as *CmapTools*) to ascertain if the observed disparities persist or if digital tools can mitigate the manual constraints, thereby providing clearer insights into the cognitive processes at play (see Watson et al., 2015).

Elaboration in the learning process

As discussed in **H1**, prior knowledge affects activity in brain networks and further moderates the processing of new information, according to fMRI analyses (Raykov et al., 2021). Elaboration processes and, thus, the amount of prior knowledge incorporated into the concept maps were examined through the analysis of propositions. The analysis revealed that students of the RCM condition used more *propositions based on prior knowledge* and more *propositions based on previous learning opportunities* than the students of the SCM condition (see **H4.1 – 4.2**). This is in line with studies that have shown the effectiveness of RCM in terms of elaboration (Becker et al., 2021b). It is assumed that in RCM, the absence of the learning text encourages the students to actively recall and use their prior knowledge to represent the connections in a concept map graphically (Becker et al., 2021b). As already explained, the activation of prior knowledge can lead to an increase in intrinsic cognitive load on the one hand (see **H2.2**) and a higher learning performance on the other (see **H1**). Thus, it can be assumed that RCM provides adequate support and challenge (*desired difficulty*; Roediger & Karpicke, 2006b; Schmidt & Bjork, 1992). According to our results, the extent of elaboration may play a more crucial role in learning performance rather than the number of propositions and, thus, the concept map quality.

Limitations

Our study's exploration of retention is limited because we assessed knowledge immediately after the learning phase. While this approach aimed to capture the efficacy of retrieval practice, the literature suggests that extending the delay to several days or weeks might yield a more accurate assessment of learning outcomes (Rowland, 2014). Delayed tests are noted for their heightened sensitivity in detecting genuine learning effects compared to immediate assessments (Dunlosky et al., 2013). Future research should contemplate longer intervals between learning activities and subsequent testing to more definitively ascertain the durability of knowledge acquired through RCM versus SCM.

In regard to the cognitive load measurement, future efforts might explore more objective and sensitive instruments. Additionally, we did not assess cognitive load immediately following each learning text, a choice that has constrained our ability to fully document fluctuations in cognitive load throughout the study. Our decision to measure cognitive load at the study's end was driven by a desire to minimise potential participant fatigue and response bias.

The concept map quality assessment focused primarily on procedural aspects, such

as structural and relational accuracy, rather than content alone. Recognising the diversity of potential metrics, future investigations should aim to develop and validate measures tailored for RCM scenarios. These measures should be designed to capture the intricate dynamics of retrieval practice and its impact on learning effectiveness, encompassing various dimensions (e.g., the total number of concepts or terms, the accuracy of propositions, the relevance of concepts and propositions to the learning objectives).

Conclusions

The present study shows that creating concept maps without the learning text available (i.e., including a retrieval practice component) improves learning relative to creating concept maps with the text available. Due to retrieval practice, students must recall information from memory, which is challenging and seems to result in concept maps comprising a limited number of propositions. However, these retrieved concepts are often deeply connected to prior knowledge, leading to better understanding and improved learning outcomes. Considering these results, our study supports that retrieval practice deserves more focus in learning settings.

Acknowledgements

We would like to express our gratitude to the reviewers for their valuable feedback, which has significantly improved our manuscript. We also thank Britta Mörs for her insightful comments on an earlier version of the manuscript. Additionally, we extend our appreciation to the students of the module “Research Projects in Biology Education” for their assistance in developing survey materials and supporting data collection.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the German Research Foundation (DFG) [grant number GR 4763/2-1].

Data Availability Statement

The data that support the findings of this study are openly available in OSF at <https://doi.org/10.17605/OSF.IO/RW67C>.

ORCID

Sina Lenski <http://orcid.org/0000-0001-6379-0941>

Mirlinda Mustafa <http://orcid.org/0009-0008-4495-4953>

Jörg Großschedl <http://orcid.org/0000-0002-7943-4818>

References

- Antony, J.W., Ferreira, C.S., Norman, K.A., & Wimber, M. (2017). Retrieval as a fast route to memory consolidation. *Trends in Cognitive Sciences*, 21(8), 573–576. <https://doi.org/10.1016/j.tics.2017.05.001>
- Austin, L.B., & Shore, B.M. (1995). Using concept mapping for assessment in physics. *Physics Education*, 30(1), 41–45. <https://doi.org/10.1088/0031-9120/30/1/009>
- Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R.E. Mayer (Eds.), *The Cambridge handbook of multimedia learning* (pp. 206–226). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.011>
- Bae, C.L., Therriault, D.J., & Redifer, J.L. (2019). Investigating the testing effect: Retrieval as a characteristic of effective study strategies. *Learning and Instruction*, 60, 206–214. <https://doi.org/10.1016/j.learninstruc.2017.12.008>
- Becker, L.B., Welter, V.D.E., Aschermann, E., & Großschedl, J. (2021a). Comprehension-oriented learning of cell biology: Do different training conditions affect students' learning success differentially? *Education sciences*, 11(438), 1–31. <https://doi.org/10.3390/educsci11080438>
- Becker, L.B., Welter, V.D.E., & Großschedl, J. (2021b). Effects of strategy training and elaboration vs. retrieval settings on learning of cell biology using concept mapping. *Education Sciences*, 11(9), 530–525. <https://doi.org/10.3390/educsci11090530>
- Blunt, J.R., & Karpicke, J.D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology*, 106(3), 849–858. <https://doi.org/10.1037/a0035934>
- Brennan, R.L. (2006). *Educational Measurement* (4th ed.). National Council on Measurement in Education.
- Buchin, Z.L. (2021). *Retrieval-based learning and element interactivity: The role of prior knowledge* (Doctoral dissertation). The University of North Carolina at Chapel Hill).

- 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. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(6), 1547–1552. <https://doi.org/10.1037/a0024140>
- Chan, J., Meissner, C., & Davis, S. (2018). Retrieval potentiates new learning: A theoretical and meta-analytic review. *Psychological Bulletin*, 144(11), 1–133. <https://doi.org/10.1037/bul0000166>.
- Chen, O., Woolcott, G., & Sweller, J. (2017). Using cognitive load theory to structure computer-based learning including MOOCs. *Journal of Computer Assisted Learning*, 33(4), 293–305. <https://doi.org/10.1111/jcal.12188>
- Christodoulou, J.A., Del Tufo, S.N., Lymberis, J., Saxler, P.K., Ghosh, S.S., Triantafyllou, C., Whitfield-Gabrieli, S., & Gabrieli, J.D. (2014). Brain bases of reading fluency in typical reading and impaired fluency in dyslexia. *PLoS One*, 9(7), 1–14. <https://doi.org/10.1371/journal.pone.0100552>
- Clausen, S., & Christian, A. (2012). Concept Mapping als Messverfahren für den außerschulischen Bereich [Concept Mapping for measurement in a non-scholar context]. *Journal für Didaktik der Biowissenschaft*, 18–31.
- Congleton, A., & Rajaram, S. (2012). The origin of the interaction between learning method and delay in the testing effect: The roles of processing and conceptual retrieval organization. *Memory & Cognition*, 40(4), 528–539. <https://doi.org/10.3758/s13421-011-0168-y>
- Department of Health, E. (2014). The Belmont Report. Ethical principles and guidelines for the protection of human subjects of research. *The Journal of the American College of Dentists*, 81(3), 4.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* [Research methods and evaluation in the social and human sciences] (5th ed.). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-41089-5>
- Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., & Willingham, D.T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4–58. <https://doi.org/10.1177/1529100612453266>
- Durrheim, K., & Tredoux, C. (2004). *Numbers, hypotheses & conclusions: A course in statistics for the social sciences*. Juta and Company Ltd.
- Ekinci, S., & Şen, A.I. (2020). Investigating grade-12 students' cognitive structures about the atomic structure: a content analysis of student concept maps, *International*

- Eshuis, E.H., ter Vrugte, J., & de Jong, T. (2022). Supporting reflection to improve learning from self-generated concept maps. *Metacognition and Learning*, 17(3), 691–713. <https://doi.org/10.1007/s11409-022-09299-7>
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>
- Großschedl, J., & Tröbst, S. (2018). Biologie lernen durch Concept Mapping: Bedeutung eines Lernstrategietrainings für kognitive Belastung, kognitive Prozesse und Lernleistung – Kurzdarstellung des DFG-Projekts [Learning biology by concept mapping: the importance of 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*, 22(1), 20–30. <https://doi.org/10.4119/zdb-1630>
- Hanham, J., Leahy, W., & Sweller, J. (2017). Cognitive load theory, element interactivity, and the testing and reverse testing effects. *Applied Cognitive Psychology*, 31(3), 265–280. <https://doi.org/10.1002/acp.3324>
- Hemmerich, W. (2019). *StatistikGuru: Cohen's Kappa für zwei Rater berechnen* [StatisticsGuru: Calculate cohen's kappa for two raters]. <https://statistikguru.de/rechner/cohens-kappa-zwei-rater-berechnen.html>
- Hemmerich, W. (2020). *StatistikGuru: Bayes Faktor*. <https://statistikguru.de/lexikon/bayes-faktor.html>
- Holley, C.D., & Dansereau, D.F. (1984). Networking: The technique and the empirical evidence. In D.F. Dansereau & C.D. Holley (Eds.), *Spatial learning strategies* (pp. 81–108). Elsevier.
- Hong, Y., Moore, I.L., Smith, D.E., & Long, N.M. (2023). Spatiotemporal dynamics of memory encoding and memory retrieval states. *Journal of Cognitive Neuroscience*, 35(9), 1463–1477. https://doi.org/10.1162/jocn_a_02022
- 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>
- Hui, L., de Bruin, A.B.H., Donkers, J., & van Merriënboer, J.J.G. (2022). Why students do (or do not) choose retrieval practice: Their perceptions of mental effort during task performance matter. *Applied Cognitive Psychology*, 1–12. <https://doi.org/10.1002/acp.3933>
- Hultberg, P., Calonge, D.S., & Lee, A.E.S. (2018). Promoting long-lasting learning through instructional design. *Journal of the Scholarship of Teaching and Learning*, 18(3), 26–43. <https://doi.org/10.14434/josotl.v18i3.23179>
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Oxford University Press.

- Jiang, D., & Kalyuga, S. (2020). Confirmatory factor analysis of cognitive load ratings supports a two-factor model. *The Quantitative Methods for Psychology*, *16*(3), 216–225. <https://doi.org/10.20982/tqmp.16.3.p216>
- Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2021). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology*, 1–14. <https://doi.org/10.1037/edu0000627>
- Kalyga, 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. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review*, *23*(1), 1–19. <https://doi.org/10.1007/s10648-010-9150-7>
- Karpicke, J.D., & Blunt, J.R. (2011). 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 III, H.L. (2009). Metacognitive strategies in student learning: do students practise retrieval when they study on their own? *Memory (Hove, England)*, *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., & O’Day, G.M. (in press). Elements of effective learning. In M.J. Kahana & A.D. Wagner (Eds.), *Oxford handbook of human memory, volume II: Applications* (pp. 3–58) Oxford University Press.
- 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>
- Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, *8*. <https://doi.org/10.3389/fpsyg.2017.01997>
- Klepsch, M., & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science*, *48*(1), 45–77. <https://doi.org/10.1007/s11251-020-09502-9>
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159–174. <https://doi.org/10.2307/2529310>
- Leahy, W., & Sweller, J. (2019). Cognitive load theory, resource depletion and the delayed testing effect. *Educational Psychology Review*, *31*(2), 457–478. <https://doi.org/10.1007/s10648-019-09476-2>
- Lechuga, M.T., Ortega-Tudela, J.M., & Gómez-Ariza, C.J. (2015). Further evidence that concept mapping is not better than repeated retrieval as a tool for learning from texts.

Learning and Instruction, 40, 61–68. <https://doi.org/10.1016/j.learninstruc.2015.08.002>

- Lenhard, W., & Lenhard, A. (2022). *Berechnung des Lesbarkeitsindex LIX nach Björnson* [Calculation of the readability index LIX according to Björnson]. <http://www.psychometrica.de/lix.html>. <https://doi.org/10.13140/RG.2.1.1512.3447>.
- Lenski, S., Elsner, S., & Großschedl, J. (2022). Comparing construction and study of concept maps – an intervention study on learning outcome, self-evaluation and enjoyment through training and learning. *Frontiers in Education*. <https://doi.org/10.3389/feduc.2022.892312>
- Mayer, R.E. (2002). Cognitive theory and the design of multimedia instruction: An example of the two-way street between cognition and instruction. *New Directions for Teaching & Learning*, 89. <https://doi.org/10.1002/tl.47>
- Mayer, R.E. (2014). Cognitive theory of multimedia learning. In R. Mayer (Eds.), *The Cambridge Handbook of multimedia learning* (pp. 43–71). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- 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)
- McDermott, K. (2021). Practicing retrieval facilitates learning. *Annual Review of Psychology*, 72(1), 609–633. <https://doi.org/10.1146/annurev-psych-010419-051019>
- Mintzes, J.J., Canas, 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*, 334(6055), 453. <https://doi.org/10.1126/science.1203698>
- Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P., Elbourne, D., Egger, M., & Altman, D.G. (2012). CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *International Journal of Surgery*, 10(1), 28–55. <https://doi.org/10.1016/j.ijsu.2011.10.001>
- North Rhine-Westphalian Ministry of Education Science and Research (2005). *Schulgesetz für das Land Nordrhein-Westfalen* [School law for the state of North Rhine-Westphalia]. Retrieved January 2023, from https://recht.nrw.de/lmi/owa/br_vbl_detail_text?anw_nr=6&vd_id=3928&vd_back=N102&sg=&menu=1.
- Novak, J.D., & Cañas, A.J. (2008). *The theory underlying concept maps and how to construct and use them* (Technical Report IHMC CmapTools 2006-01 Rev 01-2008, Issue. P. Florida Institute for Human and Machine Cognition, Florida, USA. <http://cmap.ihmc.us/publications/researchpapers/theoryunderlyingconceptmaps.pdf>
- Novak, J.D., & Gowin, D.B. (1984). *Learning how to learn*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139173469>

- O'Day, G.M., & Karpicke, J.D. (2021). Comparing and combining retrieval practice and concept mapping. *Journal of Educational Psychology, 113*(5), 986–997. <https://doi.org/10.1037/edu0000486>
- Poppenk, J., Köhler, S., & Moscovitch, M. (2010). Revisiting the novelty effect: When familiarity, not novelty, enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1321–1330. <https://doi.org/10.1037/a0019900>
- Rawson, K.A., Vaughn, K.E., & Carpenter, S.K. (2015). Does the benefit of testing depend on lag, and if so, why? Evaluating the elaborative retrieval hypothesis. *Memory & Cognition, 43*(4), 619–633. <https://doi.org/10.3758/s13421-014-0477-z>
- Raykov, P.P., Keidel, J.L., Oakhill, J., & Bird, C.M. (2021). Activation of person knowledge in medial prefrontal cortex during the encoding of new lifelike events. *Cerebral Cortex, 31*(7), 3494–3505. <https://doi.org/10.1093/cercor/bhab027>
- Rey, G.D., & Buchwald, F. (2011). The expertise reversal effect: Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied, 17*(1), 33–48. <https://doi.org/10.1037/a0022243>
- Roediger, H.L., & Karpicke, J.D. (2006a). Test enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Roediger, H.L., & Karpicke, J.D. (2006b). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*(3), 181–210. <https://doi.org/10.1111/j.1745-6916.2006.00012.x>
- Roelle, J., & Nückles, M. (2019). Generative learning vs. Retrieval practice in learning from text: The cohesion and elaboration of the text matters. *Journal of Educational Psychology, 1–61*. <https://doi.org/10.1037/edu0000345>
- Rost, D. (2013). *Interpretation und Bewertung pädagogisch-psychologischer Studien* [Interpretation and evaluation of educational-psychological studies] (3rd ed.). Klinkhardt utb.
- Rowland C.A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin, 140*(6), 1432–1463. <https://doi.org/10.1037/a0037559>
- Schmidt, R.A., & Bjork, R.A. (1992). New conceptualizations of practice: Common principle in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207–218. <https://doi.org/10.1111/j.1467-9280.1992.tb00029.x>
- Schmuckler, M.A. (2001). What is ecological validity? A dimensional analysis. *Infancy, 2*(4), 419–436. https://doi.org/10.1207/S15327078IN0204_02
- 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>

- Schroeder, N.L., Nesbit, J.C., Anguiano, C.J., & Adesope, O.O. (2017). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review*, 30(2), 431–455. <https://doi.org/10.1007/s10648-007-9053-4>
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511816819.003>
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123–138. <https://doi.org/10.1007/s10648-010-9128-5>
- Sweller, J., van Merriënboer, J.J. G., & Paas, F.G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296. <https://doi.org/10.1023/A:1022193728205>
- Watson, M.K., Pelkey, J., Noyes, C., & Rodgers, M. (2015). Assessing conceptual knowledge using three concept map scoring methods. *Journal of Engineering Education*, 105. <https://doi.org/10.1002/jee.20111>
- Weinstein, C.E., & Mayer, R. (1986). The teaching of learning strategies. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315-327). Macmillan.
- Wiklund-Hörnqvist C., Stillesjö S., Andersson M., Jonsson, B., & Nyberg, L. (2022). Retrieval practice is effective regardless of self-reported need for cognition - behavioral and brain imaging evidence. *Frontiers in Psychology*, 12, 1–10. <https://doi.org/10.3389/fpsyg.2021.797395>
- Wimmer, H., & Mayringer, H. (2014). *SLS 2-9: Salzburger Lese-Screening für die Schulstufen 2-9* [SLS 2-9: Salzburg Reading screening for grades 2-9]. Hogrefe.

5. Diskussion

Das Lernen in der Biologie stellt für Lernende im schulischen und universitären Kontext Herausforderungen dar. Diese sind aufgrund der Komplexität der Lerninhalte gegeben und resultieren in meist unzureichenden Bemühungen, sich mit biologischen Inhalten zu beschäftigen. Als Reaktion darauf werden verschiedene Lernstrategien in der schulischen und universitären Lehre eingesetzt. Diese Lernstrategien zielen darauf ab, einerseits kognitive Aspekte zu verbessern und somit den Wissenserwerb und damit die Lernperformanz zu steigern, sowie die (extrinsische) kognitive Belastung zu verringern (siehe Kapitel 2.2.1). Andererseits wird durch den Einsatz von Lernstrategien intendiert, affektive Aspekte wie Interesse und intrinsische Motivation positiv zu beeinflussen (siehe Kapitel 2.2.2). Die Betrachtung dieser beiden Aspekte ist, in Anlehnung an die hier genutzte Definition von Lernstrategien (nach Friedrich & Mandl, 2006; siehe auch Kapitel 2.2), welche ebenfalls kognitive und affektive Aspekte einbezieht, sinnvoll, um die Wirksamkeit der Lernstrategien beurteilen zu können.

Nachfolgend wird inhaltlich mit Rückbezug auf die Ergebnisse der Studien, der Einfluss der Lernstrategien auf kognitive und affektive Aspekte diskutiert. Anschließend folgen die methodische Diskussion und die Limitationen der Studien sowie abschließend der Ausblick und die Implikationen für Forschung sowie schulische und universitäre Praxis im Fach Biologie.

5.1 Inhaltliche Diskussion

Angesichts der sich wandelnden Lernbedingungen, die gekennzeichnet sind durch zunehmend vielfältigere Lerngruppen und die fortschreitende Digitalisierung (siehe Kapitel 1), ist es erforderlich bewährte Lernstrategien kontinuierlich zu evaluieren und zu ergänzen und damit zu modifizieren durch weitere Elemente. Um dies zu bewerkstelligen, wurden im Rahmen dieser Arbeit bewährte Lernstrategien und somit jene die in vergangenen empirischen Studien sowohl hinsichtlich kognitiver als auch affektiver Aspekte, Wirksamkeit gezeigt haben, untersucht. In (quasi-) experimentellen Feldstudien wurden zwei Formen einer Lernstrategie gegenübergestellt. In den jeweiligen Studien wurde in einer Bedingung eine bewährte Lernstrategien eingesetzt und in der anderen eine modifizierte Lernstrategie, welche „neue“ Elemente enthielt. Diese Arbeit verfolgte das Ziel, zu untersuchen, inwiefern sich Lernende der Biologie hinsichtlich kognitiver und affektiver Aspekte zwischen den

angewendeten Formen der jeweiligen Lernstrategien unterscheiden. Die Ergebnisse der einzelnen Studien werden in Tabelle 2 dargestellt.

Tabelle 2

Zusammenfassung der Ergebnisse der drei Studien (I-III) mit Hypothesenkennzeichnungen

| Studie | Hypothese | Ergebnisse |
|------------|------------|--|
| I | 1.1 | <ul style="list-style-type: none"> - Sowohl hinsichtlich der kognitiven als auch der affektiven Komponente der Forschungskompetenz konnten keine signifikanten Unterschiede zwischen den Bedingungen gezeigt werden. - Bei der kognitiven Komponente der Forschungskompetenz lag jedoch ein Zuwachs von Semesterbeginn zu Semesterende vor, welcher bei der affektiven Komponente der Forschungskompetenz nicht gegeben war. |
| | 1.2 | <ul style="list-style-type: none"> - Hinsichtlich des forschungsmethodischen Wissens konnte ein Zuwachs von Semesterbeginn zu Semesterende gezeigt werden, welcher jedoch unabhängig von der Bedingung war. |
| | 1.3 | <ul style="list-style-type: none"> - Die Bedingungen unterschieden sich nicht signifikant hinsichtlich der intrinsischen Motivation bei der Bearbeitung der Lernmodule. |
| II | 2 | <ul style="list-style-type: none"> - In der Bedingung mit gestuften Lernhilfen zeigten die Lernenden einen signifikant höheren Zuwachs an konzeptuellem Wissen im Vergleich zu der Bedingung, die instruktionale Unterstützung durch die Lehrkraft enthielt. - Hinsichtlich fachlich inadäquater Vorstellungen konnte jedoch kein signifikanter Unterschied zwischen den beiden Bedingungen gezeigt werden. |
| | | |
| III | 3.1 | <ul style="list-style-type: none"> - Lernende der Bedingung mit abrufbasierten Concept Mapping zeigten eine signifikant höhere Lernperformanz als Lernende des materialbasierten Concept Mappings. |
| | 3.2 | <ul style="list-style-type: none"> - Hinsichtlich der extrinsischen und lernbezogenen kognitiven Belastung lagen keine signifikanten Unterschiede zwischen den Bedingungen vor. Jedoch zeigten Lernende der Bedingung mit dem abrufbasierten Concept Mapping eine signifikant höhere intrinsische kognitive Belastung als Lernende des materialbasierten Concept Mappings. |
| | 3.3 | <ul style="list-style-type: none"> - Concept Maps, die in der Bedingung des abrufbasierten Concept Mappings erstellt wurden, zeigten eine signifikant niedrigere Concept Map-Qualität als Concept Maps aus der Bedingung mit materialbasiertem Concept Mapping. |
| | 3.4 | <ul style="list-style-type: none"> - Concept Maps der abrufbasierten Bedingung zeigten mehr Propositionen, die auf Vorwissen basierten und mehr Propositionen, die auf vorangegangenen Lernmaterialien beruhen als Concept Maps der materialbasierten Bedingung. |

5.1.1 Inhaltliche Diskussion zu kognitiven Aspekten

Mit Blick auf die Ergebnisse der Studien (I-III) konnte gezeigt werden, dass die hier eingesetzten Lernstrategien kognitive Aspekte des Lernens beeinflussen. Die Ergebnisse indizieren, wie bereits in Ausubels (1963, 1968) Assimilationstheorie (siehe Kapitel 2.1.2) ausgeführt, dass das Vorwissen für das Lernen von großer Bedeutung ist (Binder et al., 2021).

In Studie I wurde das Vorwissen als wesentliche Komponente für die Wirksamkeit der Rückmeldungen erkannt, auch wenn dieses im Rahmen der Studie nicht explizit untersucht wurde. Rückmeldungen verknüpfen das, was die Lernenden wissen, mit dem, was sie wissen sollen und damit dem Lernziel (Hattie & Timperley, 2007). Rückmeldungen können Lernende über falsche Antworten informieren (korrektive Rückmeldungen) oder sie darüber hinaus bei der Korrektur ihrer Antworten anleiten (unterstützende Rückmeldungen). Zwar indizieren bisherige Studien, dass Rückmeldungen förderlich sind für den Wissenszuwachs (Mertens et al., 2022; van der Kleij et al., 2015) und den Kompetenzzuwachs (Garcia et al., 2019) und dass dies abhängig ist von der Art der eingesetzten Rückmeldung (Mertens et al., 2022), doch konnten entgegen der Hypothesen 1.1 und 1.2 keine Unterschiede im forschungsmethodischen Wissen und in der kognitiven Komponente der Forschungskompetenz zwischen den beiden Bedingungen festgestellt werden. Es lag jedoch eine Zunahme betreffender Variablen von Anfang bis zum Ende des Semesters vor. Es ist möglich, dass die Unterschiede zwischen den Bedingungen, durch die Vorbereitung der Lernenden auf die zu Semesterende anstehende Modulabschlussprüfung, überdeckt wurden (siehe Gussen et al., 2023). Die Ergebnisse indizieren, dass Rückmeldungen, unabhängig davon, ob sie korrektiver oder unterstützender Natur sind, zu einem Wissens- und Kompetenzzuwachs führen. In zukünftigen Studien sollten auch Bedingungen ohne jegliche Form von Rückmeldungen eingesetzt werden, um dies eingehender zu überprüfen.

Auch durch Studie III wurde die Relevanz des Vorwissens deutlich, da in beiden Bedingungen, beim abrufbasierten Concept Mapping jedoch signifikant mehr, Propositionen in die Concept Maps integriert wurden, die auf Vorwissen oder auf vorangegangene Lernmaterialien basieren (siehe Hypothesen 3.4.1, 3.4.2). Über die Lernstrategie der Concept Maps wurden die Lernenden vermutlich dazu veranlasst, ihr Wissen zu reorganisieren, zu strukturieren und visuell-räumlich darzustellen (O'Day & Karpicke, 2020). Das Praktizieren des Abrufens beim Concept Mapping erforderte von Lernenden, welche nur eine begrenzte Anzahl von Konzepten in ihrem Arbeitsgedächtnis halten können (Miller, 1956), dass sie verstärkt Konzepte aus dem (Langzeit-) Gedächtnis abriefen, um die Konzept-Zusammenhänge in ihren

Concept Maps adäquat darstellen zu können. Folglich kann angenommen werden, dass abrufbasiertes Concept Mapping, durch die Verknüpfung bisherigen und neuen Wissens, Elaborationsprozesse begünstigte (Becker, Welter & Großschedl, 2021) und so zum bedeutungsvollen Lernen beitrug (Blunt & Karpicke, 2014; O'Day & Karpicke, 2020). Es wird angenommen, dass das Praktizieren des Abrufens, die langfristige Speicherung (Konsolidierung) der Informationen im Gedächtnis ermöglicht (Antony et al., 2017), indem semantisch verwandtes Wissen im Langzeitgedächtnis aktiviert wird und dieses mit neuen Informationen verknüpft wird (*elaborative retrieval hypothesis*: Carpenter, 2009). In Einklang mit diesen Annahmen können auch die Ergebnisse der Lernperformanztests (siehe Hypothese 3.1) verortet werden, welche indizieren, dass abrufbasiertes Concept Mapping zu einer signifikant höheren Lernperformanz führte im Vergleich zum materialbasiertem Concept Mapping. Neben der gängigen Auswertung des Wissens über Lernperformanztests, können auch die Lernprodukte, nämlich die Concept Maps, valide Einblicke geben in das Wissen der Lernenden (McClure et al., 1999; Ruiz-Primo & Shavelson, 1996). Im Vergleich zum materialbasierten Concept Mapping zeigten die Lernenden des abrufbasierten Concept Mappings bessere Ergebnisse im Lernperformanztest, jedoch eine geringere Concept Map-Qualität (siehe Hypothese 3.3). Diese vermeintliche Diskrepanz lässt sich dadurch erklären, dass die Auswertungsmethode über die Concept Map-Qualität vormals mit materialbasiertem Concept Mapping erprobt wurde, bei der das konzeptuelle Wissen der Lernenden valide gemessen werden konnte (McClure et al., 1999). Während Lernende beim materialbasierten Concept Mapping die Konzepte und ihre Verbindungen kontinuierlich im Lernmaterial nachlesen konnten, fehlte diese Unterstützung bei Lernenden des abrufbasierten Concept Mappings, wodurch diese nur auf die im Gedächtnis behaltenen Konzepte zugreifen konnten. So kann angenommen werden, dass das abrufbasierte Concept Mapping Lernende dazu auffordert, ihr Wissen neu zu strukturieren und zwischen wichtigen und weniger wichtigen Konzepten zu unterscheiden (Congleton & Rajaram, 2011). Folglich führt die begrenzte Anzahl der im Gedächtnis behaltenen Konzepte, zu einer begrenzten Anzahl an Propositionen in der erstellten Concept Map, was eine geringere Concept Map-Qualität zur Folge hatte (siehe Clausen & Christian, 2012). Zudem ist beim Vergleich der Auswertung des Wissens der Lernenden durch die Concept Map-Qualität und der Lernperformanztests anzumerken, dass die Lernperformanztests eine validere Messung darstellen. Dies ist deshalb gegeben, da Lernperformanztests mehrere Items und nicht wie die Concept Map-Qualität nur einen einzigen Wert umfassen (McClure et al., 1999).

Wie in Studie III dargestellt, können Concept Maps einerseits als Lernstrategie verwendet und andererseits auch zur Diagnose des Wissens von Lernenden genutzt werden. Ursprünglich wurden Concept Maps zur Diagnose der Wissensstruktur Lernender entwickelt (Novak & Musonda, 1991) und auch in Studie II als solche eingesetzt. Mittels Concept Maps wurde in Studie II das konzeptuelle Wissen erhoben und ermittelt, dass gestufte Lernhilfen im Vergleich zu instruktionalen Unterstützungen durch die Lehrkraft einen signifikant höheren Zuwachs im konzeptuellen Wissen ermöglichten. So wird angenommen, dass Lernende von der Möglichkeit profitierten, anonyme Unterstützung zu erhalten, ohne die Lehrkraft persönlich um Hilfe bitten zu müssen (Franke-Braun et al., 2008). Lernende konnten selbstgesteuert die Unterstützung in Anspruch nehmen, die sie für das Lösen komplexer Problemlöseaufgaben benötigen und nach der Aufgabenlösung die, durch die gestuften Lernhilfen bereitgestellte, Beispiel-Lösung abgleichen (Hänze et al., 2010; Kleinert et al., 2022). Der differenzierende Charakter gestufter Lernhilfen (Kleinert, 2023) bietet insbesondere für heterogene Lerngruppen, die im Lernalltag zunehmend gängiger werden, eine Entlastung für Lehrende und eine Unterstützung für Lernende.

Concept Maps eignen sich nicht nur, um das konzeptuelle Wissen selbst (in Studie II bezeichnet als *proposition accuracy score*) zu erfassen, sondern auch um mögliche fachlich inadäquate Vorstellungen (bezeichnet als *misconceptions score*) zu erfassen (Wang et al., 2018, siehe Hypothese 2). Ursprünglich wurde angenommen, dass gestufte Lernhilfen es ermöglichen, fachlich inadäquate Vorstellungen zu korrigieren. Diese Annahme konnte jedoch durch die Ergebnisse der Studie nicht gestützt werden. Es ist davon auszugehen, dass die Korrektur fachlich inadäquater Vorstellungen ein aktives Engagement erfordert, um einen kognitiven Konflikt zu erreichen (siehe Posner et al., 1982). Da die Lernenden jedoch nicht mit ihren fachlich inadäquaten Vorstellungen konfrontiert wurden und somit keine Korrektur der Vorstellungen in Richtung wissenschaftliche Vorstellungen vorgenommen wurde, erscheint das Nichtvorliegen von Unterschieden nicht überraschend. Auch ist zu beachten, dass Vorstellungen, aktuellen Annahmen zu Folge, nicht schlicht *korrigiert* werden können, sondern wissenschaftliche und fachlich inadäquate Vorstellungen Lernender nebeneinander koexistieren (Bryce & Blown, 2023, siehe Kapitel 2.1.2). Durch die vorliegende Analyse konnte eine weitreichende qualitative Einschätzung der Vorstellungen Lernender gewonnen werden, wobei sich zeigte, dass Lernende die Konzepte Energie, Wärme, Aktivierungsenergie sowie endotherme und exotherme Reaktionen nicht eindeutig unterscheiden konnten. Zudem zeigten sich auch fachlich inadäquaten Vorstellungen, wobei Lernende annahmen, dass eine chemische Reaktion ohne Gasbrenner nicht möglich sei oder dass Wasser

Energie erzeugen könne. Diese Ergebnisse sind insbesondere für zukünftige Studien nützlich, in denen es möglich sein könnte, Anreize für kognitive Konflikte innerhalb der gestuften Lernhilfen zu schaffen. Erneut konnte auch mittels Studie II aufgezeigt werden, dass Lernende stets Vorstellungen und Vorwissen zu den betreffenden Lerngegenständen mitbringen, welche als Ressource statt als Hindernis betrachtet werden sollten (Hattie & Clarke, 2018). Lehrende sollten somit die Vorstellungen und das Vorwissen der Lernenden erfassen und in der Planung des Unterrichts und beim Einsatz von Lernstrategien mitberücksichtigen. Rückblickend zeigen alle drei Studien auf, dass Lernstrategien zu einem Zuwachs an verschiedenen Arten von Wissen führen. Von besonderer Relevanz ist hierbei, was Lernende an Vorstellungen und Vorwissen mitbringen. Dies ist nicht nur für den Wissens- und auch Kompetenzzuwachs von Bedeutung, sondern auch maßgeblich für die kognitive Belastung. Die kognitive Belastung, als wesentlicher kognitiver Aspekt des Lernens, wurde in Studie III explizit behandelt (siehe Hypothese 3.2), war jedoch auch für die Planung von Studie I und II maßgeblich.

In Studie III wurde angenommen, dass die extrinsische kognitive Belastung erhöht wird, wenn Lernende mehrere Informationsquellen zur Bearbeitung ihrer Aufgaben (hier der Konstruktion von Concept Maps) nutzen müssen (Chen et al., 2017, siehe *split attention*: Kapitel 2.2.1). Es wurde angenommen, dass beim abrufbasierten Concept Mapping die extrinsische kognitive Belastung geringer ausfällt, da die Lernenden ihre visuelle Aufmerksamkeit nicht zwischen dem Lernmaterial und der Concept Map aufteilen müssen (Ayres & Sweller, 2014; Mayer, 2002; Sweller, 2010, siehe Hypothese 3.2.1). Für diese Annahme konnte in den Ergebnissen jedoch keine Evidenz gefunden werden, da zwischen den Bedingungen keine signifikanten Unterschiede festgestellt wurden. Dies ist möglicherweise auf das eingesetzte Messinstrument nach Klepsch et al. (2017) zurückzuführen, welches nicht ausreichend sensibel zwischen den Quellen der extrinsischen und intrinsischen kognitiven Belastung unterscheiden konnte, insbesondere wenn eine hohe Elementinteraktivität vorliegt (Klepsch & Seufert, 2020). Da die Konzepte in der vorliegenden Studie nicht unabhängig voneinander, sondern nur im Zusammenhang zueinander erschlossen werden konnten (siehe Beispiel zu Ökosystemen, Kapitel 2.1.1), kann die Elementinteraktivität als hoch eingestuft werden, so dass der Einsatz sensiblerer Instrumente zur Messung der kognitiven Belastung empfohlen wird (Klepsch & Seufert, 2020).

Grundsätzlich ist, Annahmen der *cognitive load theory* zu Folge, eine hohe Elementinteraktivität für Lernende mit wenig (oder keinem Vorwissen) herausfordernd (Chen et al., 2017;

Sweller, 2010). Da jedoch die Lernenden der vorliegenden Studie in beiden Bedingungen über ein gewisses Maß an Vorwissen über Ökosysteme verfügten (siehe deskriptive Ergebnisse zum Prä-Test in Studie III) und durch das Training prozedurales Wissen zu Concept Maps erwarben (siehe Becker, Welter, Aschermann & Großschedl, 2021), war es für sie vermutlich einfacher, die Aufgabe der Concept Map-Erstellung zu bewältigen. Da Vorwissen weniger Raum im Arbeitsgedächtnis einnimmt, sind Strategien zur Reduzierung der Belastung des Arbeitsgedächtnisses (z. B. materialbasiertes Concept Mapping) nicht vorteilhafter für das Lernen als anspruchsvollere Strategien (z. B. abrufbasiertes Concept Mapping; *expertise-reversal-effect*: Rey & Buchwald, 2011). Dies liefert eine mögliche Erklärung für die höhere intrinsische kognitive Belastung beim abrufbasierten im Vergleich zum materialbasierten Concept Mapping (siehe Hypothese 3.2.2). Zudem kann davon ausgegangen werden, dass beim abrufbasierten Concept Mapping, durch die Abwesenheit des Lernmaterials, mehrere Konzepte gleichzeitig im Arbeitsgedächtnis gehalten werden müssen, was eine höhere intrinsische kognitive Belastung zur Folge hat (siehe Sweller et al., 1998).

An dieser Stelle sollte erwähnt werden, dass innerhalb wissenschaftlicher Diskurse diskutiert wird, ob die intrinsische und lernbezogene kognitive Belastung, wie sie mit dem Fragebogen von Leppink et al. (2013) gemessen werden, auf der gleichen theoretischen Grundlage basieren (Jiang & Kalyuga, 2020). So wird vorgeschlagen, lediglich zwischen extrinsischer und intrinsischer Belastung zu unterscheiden (Kalyuga, 2011). Untersuchungen des hierverwendeten Fragebogens haben jedoch gezeigt, dass das traditionelle Modell mit drei Quellen der kognitiven Belastung in konfirmatorischen Analysen am geeignetsten ist (Klepsch et al., 2017). Auch in vorliegender Studie gibt es Hinweise darauf, dass intrinsische und lernbezogene kognitive Belastung unterschieden werden müssen (unterschiedliche Ergebnisse hinsichtlich der intrinsischen und lernbezogenen kognitiven Belastung).

Eingangs wurde angenommen, dass abrufbasiertes Concept Mapping die lernbezogene kognitive Belastung erhöht (siehe Hypothese 3.2.3). Zwar deuteten die verstärkten Elaborationsprozesse in der abrufbasierten Concept Mapping-Bedingung (siehe Hypothese: 3.4) auf diese Annahme hin, jedoch lag in den Ergebnissen zur lernbezogenen kognitiven Belastung keine Evidenz für diese Annahme vor. Dies könnte durch die Motivation der Lernenden erklärt werden, die nicht automatisch ihre verfügbare kognitive Kapazität in eine Lernaufgabe investieren (Schnotz & Kürschner, 2007). Da Lernende das Praktizieren des Abrufs als schwierig empfinden (Blunt & Karpicke, 2014), ist es naheliegend anzunehmen, dass sie nicht bereit waren, ihre freien kognitiven Kapazitäten in abrufbasiertes Concept Mapping zu

investieren. Dies könnte die Vorteile von abrufbasiertem gegenüber materialbasiertem Concept Mapping in der wahrgenommenen lernbezogenen kognitiven Belastung verschleiern haben und betreffende Ergebnisse erklären.

5.1.2 Inhaltliche Diskussion zu affektiven Aspekten

Bereits im vorangegangenen Kapitel wurde hervorgehoben, dass der Einsatz verschiedener Lernstrategien einen wesentlichen Einfluss auf die kognitiven Aspekte des Lernens hat. Lernstrategien initiieren nicht nur in unterschiedlichem Maße die Einbindung des Vorwissens, sondern erwirken darüber hinaus den Zuwachs unterschiedlicher Formen des Wissens. In diesem Zusammenhang beeinflussen Lernstrategien zwangsläufig die kognitiven Wissensstrukturen der Lernenden und ihre wahrgenommene kognitive Belastung. Obwohl ein erhöhter Fokus dieser Arbeit auf kognitiven Aspekten der Lernstrategien lag, wird an verschiedenen Stellen, im Besonderen hinsichtlich der zuletzt genannten lernbezogenen kognitiven Belastung deutlich, dass auch affektive Aspekte eine zentrale Rolle beim Lernen einnehmen.

Während innerhalb der Studien II und III affektive Aspekte nur tangiert wurden, wurden sie in Studie I explizit untersucht. Hierzu wurden in Studie I sowohl die affektiv-motivationale Komponente der Forschungskompetenz (siehe Hypothese 1.1) als auch die intrinsische Motivation bei der Bearbeitung der Lernmodule untersucht (siehe Hypothese 1.3). Hinsichtlich der affektiv-motivationalen Komponente der Forschungskompetenz zeigten die Ergebnisse, dass weder ein Anstieg von Semesterbeginn zu Semesterende, noch ein Unterschied zwischen den Bedingungen mit korrektiven und unterstützenden Rückmeldungen vorlag. Dieses Ergebnis zeigt erneut auf, dass Rückmeldungen verstärkt kognitive und weniger affektive Aspekte beeinflussen (Wisniewski et al., 2020). Da die Rückmeldungen auf forschungsmethodisches Wissen abzielten und somit kognitive Aspekte adressierten, erscheint dieses Ergebnis nicht überraschend. Um einen Anstieg in der affektiv-motivationalen Komponente der Forschungskompetenz zu erreichen, wäre möglicherweise die Verdeutlichung der Relevanz von Forschung für Lehramtsstudierende innerhalb der Lernmodule förderlich gewesen. So könnte den Lehramtsstudierenden durch die Verdeutlichung der Relevanz von Forschung, ein erhöhtes Interesse an Forschung und eine erhöhte Motivation sich mit Forschung auseinanderzusetzen, ermöglicht werden (Gussen et al, 2023; Reeve, 2015). Dies könnte in zukünftigen Studien untersucht werden.

Ein weiterer affektiver Aspekt betrifft die intrinsische Motivation bei der Bearbeitung der Lernmodule. So wurde angenommen, dass unterstützende Rückmeldungen, und damit jene, die weitere Hinweise zum Beantworten der Fragen bereitstellen, die Wahrnehmung von Kompetenz und Autonomie und somit auch die intrinsische Motivation im Vergleich zu korrektiven Rückmeldungen fördern (Großmann et al., 2020; Reeve, 2015). Jedoch konnten auch hinsichtlich der intrinsischen Motivation bei der Bearbeitung der Lernmodule keine Unterschiede zwischen den beiden Bedingungen identifiziert werden. Zurückzuführen ist dies auf mögliche Unterschiede, die bereits zu Semesterbeginn vorgelegen haben könnten. Da im Prä-Test die Motivation nicht erhoben wurde, können hierzu keine weiteren Angaben gemacht werden. Diese Variable (dispositionale Motivation) sollte in zukünftigen Studien zu Beginn der Intervention beispielsweise über den SMR-LS (Skala zur motivationalen Regulation bei Lernen im Studium, Thomas et al., 2018) miterhoben werden. Auch ist zu beachten, dass die hier angewendete Form der unterstützenden Rückmeldungen die intrinsische Motivation verringert haben könnte, da Lernende die Kontrolle über den eigenen Lernprozess abgeben mussten (siehe Noetel et al., 2021). Zwar indizieren vorliegende Ergebnisse keinen Nachteil unterstützender Rückmeldungen für die intrinsische Motivation, jedoch könnten in zukünftigen Studien auch unterstützende Rückmeldungen implementiert werden, die optional von den Lernenden genutzt werden.

5.2 Methodische Diskussion und Limitationen

Die hier untersuchten Studien (I-III) erfordern eine methodische Diskussion und die Darstellung von Faktoren, die die Studien möglicherweise eingeschränkt haben könnten.

Alle hier untersuchten Studien wurden im Feld durchgeführt, wodurch angenommen wird, dass sie im Vergleich zu Laborstudien eine hohe externe Validität aufweisen (Döring & Bortz, 2016). Da die Studien unter authentischen Lernbedingungen durchgeführt wurden und nicht innerhalb künstlich erstellter Lernbedingungen, ist zudem anzunehmen, dass auch eine hohe ökologische Validität vorliegt (siehe Schmuckler et al., 2001). Da es sich jedoch bei den Studien II und III um quasi-experimentelle Untersuchungen handelt, in denen keine randomisierte Zuteilung zu den Bedingungen vorlag, ist anzunehmen, dass die interne Validität und somit auch die Generalisierbarkeit eingeschränkt sind (Döring & Bortz, 2016). In Studie I ist die externe Validität durch einen nicht hinreichend großen Stichprobenumfang eingeschränkt. Sowohl der Stichprobenumfang als auch die nicht-randomisierte Zuteilung zu den Bedingungen ergab sich aus der Durchführung der Untersuchungen als Feldstudien,

in denen meist die organisationalen Strukturen der Institutionen (hier: Schule) nicht zugunsten des Studiendesigns überwunden werden konnten. Um möglichen Konfundierungen zu begegnen, wurden beispielsweise Kovariaten (Vorwissen, Studie III) einbezogen, sodass die Vergleichbarkeit der Bedingungen weitreichend ermöglicht werden sollte.

Angestrebt wurde die Vergleichbarkeit der Bedingungen weiterhin dadurch, dass vor Beginn der Intervention ein Lernstrategietraining (Studie II und III) durchgeführt wurde, um für alle Lernenden eine Grundlage zu schaffen und so prozedurales Wissen zu Concept Maps bei allen Lernenden (möglichst gleichermaßen) zu vermitteln. Auch wurde in den Studien vorab geprüft, ob Unterschiede zwischen demographischen Angaben (Alter, Geschlecht), zwischen Fähigkeiten (z. B. Lesefähigkeit in Studie III) oder Unterschiede im Vorwissen in den Prä-Tests zwischen den beiden Bedingungen vorlagen (Studie I-III). Für Studie I wäre es jedoch sinnvoll gewesen, die Motivation zu Beginn der Untersuchung (zu Semesteranfang) zu erheben, da nicht vorliegende Unterschiede der intrinsischen Motivation bei der Bearbeitung der Lernmodule auf bestehende Motivationsunterschiede vor Beginn der Intervention zurückführbar sein könnten. Dies limitiert die betreffende Studie.

Eine weitere Limitation betrifft die Nutzung der Lernstrategie. In den Studien (I-III) wurde nicht überprüft, ob die Lernstrategien in adäquater Weise von Lernenden genutzt wurden. So indizieren beispielsweise die hohen Standardabweichungen in der Bedingung mit gestuften Lernhilfen (Studie II), dass es Unterschiede in der Nutzung der gestuften Lernhilfen gegeben haben könnte. Auch hinsichtlich der integrierten unterstützenden Rückmeldungen in interaktiven Lernvideos (Studie I) bleibt fraglich, ob die Video-Sequenz tatsächlich erneut rezipiert wurde. In Hinblick auf die Konstruktion von Concept Maps in Studie III lag ebenfalls keine Überprüfung der Strategie vor. Allerdings kann hier angenommen werden, dass das (ausführliche) Concept Map-Training zu einer adäquaten Anwendung der Strategie geführt hat (siehe Becker, Welter, Aschermann & Großschedl, 2021). Insgesamt wäre für zukünftige Studien die Überprüfung der Nutzung von Lernstrategien sinnvoll, um die Wirksamkeit der Lernstrategien umfänglicher beurteilen zu können. Besonders hohes Potential bieten hierfür auch Tutoring-Systeme, welche die Nutzung von Lernstrategien miterfassen könnten. Besonders für Studie I, in welcher bereits mit einem Tutoring-System gearbeitet wurde, wäre dies von Vorteil gewesen. Allerdings wurde mit einer Software (*H5P*) gearbeitet, welche lediglich als Lernwerkzeug und nicht als Diagnoseinstrument angelegt ist. Dies erschien aus Gründen der praktischen Bedeutsamkeit sinnvoll (Kirk, 1996), da die Software in der üblicherweise-genutzten Lernplattform der Lernenden (*Ilias*) implementierbar war

und Lernende somit kein neues Nutzerprofil erstellen mussten. Auch Lehrende, welche die Lernstrategie in ihren Unterricht implementieren würden, hätten keine zusätzlichen Kosten. Die untersuchten Lernstrategien aus Studien II und III weisen ebenfalls praktische Bedeutsamkeit auf, da Lehrende diese ebenso ohne Mehrkosten einsetzen können.

Abgeleitet werden kann, dass angesichts der fortschreitenden digitalen Entwicklung, zunehmend auch Lernsysteme etabliert werden sollten, die einerseits als Lerninstrument als auch andererseits als Diagnoseinstrument einsetzbar sind (siehe Alonzo, 2024), um zukünftig auch die Nutzung von Lernstrategien gezielter zu ermitteln.

5.3 Implikationen und Ausblick

Unter den untersuchten Lernstrategien, darunter Rückmeldungen (Studie I), Scaffolds (Studie II) und Concept Maps (Studie III), kristallisiert sich die Gemeinsamkeit heraus, welche das Rückmelden an Lernende zu ihrem Wissen und gegebenenfalls zu unzureichendem Wissen betrifft. So ermöglichen Rückmeldungen, was in ihrer Natur selbst verankert ist, das Rückmelden zum Wissenstand der Lernenden (Hattie & Timperley, 2007). Lernende könnten, wie in Studie I untersucht, beim Erwerb forschungsmethodischen unzureichendes Wissen aufweisen, auf welches sie mit korrektiven Rückmeldungen aufmerksam gemacht werden. Unterstützende Rückmeldungen leiteten zusätzlich zur Video-Sequenz, in welcher das Wissen vermittelt wurde. Die Funktion des Rückmeldens bei Scaffolds (Studie II) wird ersichtlich, wenn Lernende bei der Bearbeitung einer mehrstufigen Aufgabe Unterstützung benötigen (Hänze et al., 2010). Lernende wurden damit indirekt auf ihr unzureichendes Wissen aufmerksam gemacht und beanspruchten Unterstützung, für welche sie entweder die Lehrkraft nach instruktionaler Unterstützung fragen mussten oder in anonymer Weise gestufte Lernhilfen in Anspruch nahmen (Franke-Braun et al., 2008). Zudem erhielten Lernende auch über die Musterlösung der (Teil-) Aufgaben, eine Rückmeldung über ihren Lernprozess (siehe Schmidt-Weigand et al., 2008). Die Funktion des Rückmeldens wird in Studie III, besonders beim abrufbasierten Concept Mapping, deutlich (siehe Donoghue & Hattie, 2021). So kann Lernenden, welche in Abwesenheit des Lernmaterials, beim Versuch Konzepte miteinander in Beziehung zu setzen und in einer Concept Map darzustellen, verdeutlicht werden, dass sie Lücken im Wissen haben. Dies initiiert möglicherweise die Aktivierung von Vorwissen, welches für das Bilden von Zusammenhängen zwischen den Konzepten hinzugezogen wird (siehe Hypothese 3.4).

Eine weitere Gemeinsamkeit zwischen den Lernstrategien könnte, mit Abstrichen, in der Entlastung der Lehrenden gesehen werden. Beim abrufbasierten Concept Mapping könnten die Lernenden als Unterstützung ihr Vorwissen einbeziehen, statt Lehrende um Unterstützung zu fragen. Auch gestufte Lernhilfen, in denen Lernende selbstregulativ Lösungshinweise in Anspruch nehmen, entlasten die Lehrkraft. Dies ist auch für die Rückmeldungen innerhalb der interaktiven Lernvideos der Fall, wo Lehrende zwar die Inhalte audiovisuell aufbereiten, die Lernenden die Inhalte aber anschließend über die Lernvideos erarbeiten. In den Lernstrategien werden die Lehrenden zwar im Lernprozess der Lernenden entlastet, müssen jedoch vor dem Einsatz der Lernstrategien diese ausreichend aufbereiten.

Insgesamt lassen sich aus der vorliegenden Arbeit Implikationen für die Forschung und für die Praxis in Schule und Universität ableiten.

Für die Praxis in Schule und Universität lassen sich folgende Implikationen ableiten:

- Unterschiedliche Lernstrategien beeinflussen verschiedene kognitive Aspekte des Lernens bei verschiedenen Gruppen von Lernenden. Die Wirksamkeit von Lernstrategien wird insbesondere durch das Vorwissen der Lernenden geprägt. Dies zeigt sich zunächst darin, dass Lernende (hier beim abrufbasierte Concept Mapping, Studie III), auf ihr Vorwissen zurückgreifen, um die Aufgaben zu bewältigen. Zu beachten bleibt hier, dass das vorhandene Wissen und die Vorstellungen der Lernenden fachlich inadäquat sein können. Daher ist es Aufgabe und Herausforderung für Lehrende, das Vorwissen der Lernenden gezielt zu aktivieren und dabei mögliche fachlich inadäquate Vorstellungen zu berücksichtigen.
- Lehrende stehen vor der Aufgabe, Lernstrategien einzusetzen, die sowohl kognitive als auch affektive Aspekte berücksichtigen und unterstützen. Um dies zu erreichen, könnte beispielsweise die Relevanz des Lerngegenstands (in Studie I am Beispiel durch die Relevanz von Forschung) verdeutlicht werden.
- Zudem wird in vorliegender Arbeit die Bedeutung der Auswahl und Anwendung geeigneter Lernstrategien akzentuiert, um die kognitiven Strukturen und die kognitive Belastung der Lernenden zu optimieren.

Für die Forschung lassen sich folgende Implikationen ableiten:

- Insbesondere die digitalen Möglichkeiten bei der Gestaltung von Lernstrategien sollten weiter untersucht und ausgeschöpft werden. Hierbei könnte darauf abgezielt werden, nicht nur das Lernen zu ermöglichen, sondern dieses auch während des

Prozesses zu überwachen. Beispielsweise könnte der Einsatz von Lernstrategien erfasst und anschließend die Wirksamkeit der Lernstrategien eingehender bewertet werden.

- Unter der Prämisse, dass kognitive und affektive Aspekte gleichermaßen zentral sind für den Lernerfolg, ist es erforderlich, dass affektive Aspekte zunehmend berücksichtigt werden. Dieses Erfordernis zeigt sich unter anderem hinsichtlich der lernbezogenen kognitiven Belastung, welche von den motivationalen Dispositionen Lernender abhängt (Studie III). Zudem bleibt zu beachten, dass die vorliegende Arbeit nur ein Bruchstück möglicher kognitiver und affektiver Aspekte hinsichtlich des Lernens in der Biologie berücksichtigt. Zukünftige Studien könnten weitere Variablen untersuchen und ebenfalls die Wechselwirkung zwischen den Variablen gezielter erfassen (z. B. affektive Aspekte des Interesses etc.).
- Auf methodischer Ebene wird deutlich, dass für die Erfassung des Wissens, verschiedene Instrumente genutzt werden können. Hierbei können herkömmliche Lernperformanztests eingesetzt werden oder auch eine Analyse der Lernprodukte (Concept Map-Qualität in Studie III) vorgenommen werden. Indem weitere qualitative Untersuchungen initiiert werden, können Lehrende beispielsweise über fachlich inadäquate Vorstellungen (Studie II) oder das Vorwissen der Lernenden, Aufschluss erhalten (Studie III).
- Zudem sollten verschiedene Variablen auf kognitiver und affektiver Ebene berücksichtigt werden und diese bestenfalls zu Beginn der Intervention und im weiteren Verlauf der Intervention erhoben werden, um auch die Entwicklung betreffender Variablen umfänglich erfassen zu können.

6. Fazit

Die vorliegende Dissertation thematisiert die Wirksamkeit und den Einfluss verschiedener Lernstrategien auf kognitive und affektive Aspekte des Lernens in der Biologie. Die Ergebnisse der durchgeführten Studien reihen sich in den bestehenden empirischen Hintergrund ein. Im Zuge der durchgeführten Studien konnten Hinweise und praktische Empfehlungen für den Einsatz von Lernstrategien beim schulischen und universitären Lernen der Biologie abgeleitet werden.

Die Dissertation leistet einen Beitrag zur Frage der Wirksamkeit verschiedener Lernstrategien und zeigt, dass nicht nur bewährte Lernstrategien relevant sind, sondern dass auch modifizierte Lernstrategien mitunter besser den aktuellen Lernbedingungen begegnen und so besonders kognitive Aspekte beeinflussen. Insgesamt ergänzt und erweitert diese Dissertation das vorhandene Wissen über die Wirksamkeit von Lernstrategien und ermöglicht ein Verständnis über die Rolle von Lernstrategien beim Lernen in der Biologie.

Literaturverzeichnis

- Abernethy, B., Maxwell, J. P., Masters, R. S. W., Kamp, J. V. D., & Jackson, R. C. (2007). Attentional Processes in Skill Learning and Expert Performance. *Handbook of Sport Psychology*, 245–263. DOI: 10.1002/9781118270011.ch11
- Adams, N.E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association*, 103(3), 152–153. DOI: 10.3163/1536-5050.103.3.010
- Alexander, P.A., & Judy, J.E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Educational Research*, 58(4). DOI: 10.2307/1170279
- Alexander, P.A., Graham, S., & Harris, K.R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review*, 10(2), 129–154. DOI: 10.1023/a:1022185502996
- Alonzo, D. (2024). *Assessment to Support Learning and Teaching*. Routledge. DOI: 10.4324/9781003396277
- Anderson, J. R. (2014). *Rules of the mind*. Psychology Press. DOI: 10.4324/9781315806938
- Antony, J.W., Ferreira, C.S., Norman, K.A., & Wimber, M. (2017). Retrieval as a fast route to memory consolidation. *Trends in Cognitive Sciences*, 21(8), 573–576. DOI: 10.1016/j.tics.2017.05.001
- Arnold, J., Kremer, K., & Mayer, J. (2016). Scaffolding beim Forschenden Lernen. *Zeitschrift für Didaktik der Naturwissenschaften*, 23(1), 21–37. DOI: 10.1007/s40573-016-0053-0
- Artelt, C. (1999). Lernstrategien und Lernerfolg - Eine handlungsnahe Studie. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 31(2), 86–96. DOI: 10.1026//0049-8637.31.2.86
- Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Schümer, G., Stanat, P., Tillmann, K. J., & Weiß, M. (Hrsg.). (2001). *PISA 2000: Zusammenfassung zentraler Befunde*. Max-Planck-Institut für Bildungsforschung.
- Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation*, 89–195. DOI: 10.1016/s0079-7421(08)60422-3
- Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. Grune and Stratton.
- Ausubel, D.P. (1968). *Educational psychology: a cognitive view*. Holt, Rinehart and Winston.
- Ayres, P., & Sweller, J. (2014). The split-attention principle in multimedia learning. In R.E. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (2. Aufl., S. 206–226). Cambridge University Press. DOI: 10.1017/cbo9781139547369.011
- Bach, A. (2019). *Chemie für Biologie*. Wbg Academic.
- Baddeley, A., & Hitch, G. (2007). Working memory: Past, present ... and future? In N. Osaka, R.H. Logie, & M. D'Esposito (Hrsg.), *The Cognitive Neuroscience of Working Memory* (Oxford Academic, S. 1–20), DOI: 10.1093/acprof:oso/9780198570394.003.0001

- Baddeley, A.D., & Hitch, G. (1974). Working Memory. *Psychology of Learning and Motivation*, 47–89. DOI: 10.1016/s0079-7421(08)60452-1
- Baloo, K. (2019). Students' difficulties during research methods training acting as potential barriers to their development of scientific thinking. In M. Murtonen, & K. Baloo (Hrsg.), *Redefining scientific thinking for higher education* (S. 107–137). Palgrave Macmillan. DOI: 10.1007/978-3-030-24215-2_5
- Bannwarth, H., Kremer, B.P., & Schulz, A. (2019). *Basiswissen Physik, Chemie und Biochemie*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-58250-3
- Baur, A. (2023). Which student problems in experimentation are related to one another? *International Journal of Science Education*, 45(10), 781–805. DOI: 10.1080/09500693.2023.2175334
- Becker, L.B. (2022). Die Kunst, Concept Mapping zu trainieren und einzusetzen: Ein Vergleich unterschiedlicher Ansätze und ihre Bedeutung für kognitive Prozesse, kognitive Belastung und Lernleistung [Dissertation, Universität zu Köln - Institut für Biologiedidaktik].
- 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(8), 438. DOI: 10.3390/educsci11080438
- Binder, T., Schmiemann, P., & Theyssen, H. (2019). Knowledge acquisition of biology and physics university students—the role of prior knowledge. *Education Sciences*, 9(4), 281. DOI: 10.3390/educsci9040281
- Binder, T., Waldeyer, J., & Schmiemann, P. (2021). Studienerfolg von Fachstudierenden im Anfangsstudium der Biologie. *ZfDN*, 27, 73–81. DOI: 10.1007/s40573-021-00123-4
- Bloom, B.S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. Longman and Green.
- Blundell, T.L. (2020). Interdisciplinary research in physics, chemistry and biology is central to understanding biological processes. *Progress in Biophysics and Molecular Biology*, 156, 1–2. DOI: 10.1016/j.pbiomolbio.2020.09.002
- Blunt, J.R., & Karpicke, J.D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology*, 106(3), 849–858. DOI: 10.1037/a0035934
- BMBF [Bundesministerium für Bildung und Forschung] (Hrsg.) (2023). *Fortschrittsbericht DigitalPakt Schule 2022–2023*. https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.digitalpaktschule.de/files/Fortschrittsbericht_DigitalPakt_Schule_2022-2023.pdf
- Braithwaite, D.W., & Sprague, L. (2021). Conceptual knowledge, procedural knowledge, and metacognition in routine and nonroutine problem solving. *Cognitive Science*, 45(10). Portico. DOI: 10.1111/cogs.13048

- Brand, M., & Markowitsch, H.J. (2006). Was weiß die Hirnforschung über Lernen? *Zeitschrift Für Erziehungswissenschaft*, 21–42. DOI: 10.1007/978-3-531-90607-2_3
- Bratz, J., & Gaida, D. (2024). Was Bildungspolitik leisten muss, damit Inklusion gelingen kann. Das Modellprojekt Klassenassistentz. In B. Jürgens, W.W. Steinert, & H. Dietlinde (Hrsg.), *Gelingensbedingungen Für Inklusive Schulen* (S. 107–113). DOI: 10.35468/6082-11
- Brush, T., & Saye, J. (2008). The effects of multimedia-Supported problem-based inquiry on student engagement, empathy, and assumptions about history. *Interdisciplinary Journal of Problem-Based Learning*, 2(1). DOI: 10.7771/1541-5015.1052
- Bryce, T.G.K., & Blown, E.J. (2023). Ausubel’s meaningful learning re-visited. *Current Psychology*, 43(5), 4579–4598. DOI: 10.1007/s12144-023-04440-4
- Burgers, C., Eden, A., van Engelenburg, M.D., & Buningh, S. (2015). How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior*, 48, 94–103. DOI: 10.1016/j.chb.2015.01.038
- 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, 1563–1569. DOI:10.1037/a0017021
- Chen, O., Woolcott, G., & Sweller, J. (2017). Using cognitive load theory to structure computer-based learning including MOOCs. *Journal of Computer Assisted Learning*, 33(4), 293–305. DOI: 10.1111/jcal.12188
- Clark, I., & Dumas, G. (2016). The regulation of task performance: A trans-disciplinary review. *Frontiers in Psychology*, 6. DOI: 10.3389/fpsyg.2015.01862
- Clausen, S., & Christian, A. (2012). Concept Mapping als Messverfahren für den außerschulischen Bereich. *Journal für Didaktik der Biowissenschaft*, 18–31.
- Collins, A.M., & Quillian, M.R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 240–247. DOI: 10.1016/s0022-5371(69)80069-1
- Congleton, A., & Rajaram, S. (2011). The origin of the interaction between learning method and delay in the testing effect: The roles of processing and conceptual retrieval organization. *Memory & Cognition*, 40(4), 528–539. DOI: 10.3758/s13421-011-0168-y
- 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. DOI: 10.1016/s0022-5371(72)80001-x
- De Boer, H.D., Donker-Bergstra, A., Kostons, D., Korpershoek, H., & Werf, M.P. (2013). Effective strategies for self-regulated learning : A meta-analysis. *GION*, 6–72.
- De Jong, T., & Ferguson-Hessler, M.G.M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31(2), 105–113. DOI: 10.1207/s15326985ep3102_2
- Donoghue, G.M., & Hattie, J.A.C. (2021). A meta-analysis of ten learning techniques. *Frontiers in Education*, 6. DOI: 10.3389/educ.2021.581216
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-41089-5

- Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., & Willingham, D.T. (2013). Improving students' learning with effective learning techniques. *Psychological Science in the Public Interest*, 14(1), 4–58. DOI: 10.1177/1529100612453266
- Earley, M. A. (2014). A synthesis of the literature on research methods education. *Teaching in Higher Education*, 19(3), 242-253.
- Ekinci, S., & Şen, A.I. (2020). Investigating grade-12 students' cognitive structures about the atomic structure: a content analysis of student concept maps. *International Journal of Science Education*, 42(6), 977–996, DOI: 10.1080/09500693.2020.1744045
- Eshuis, E.H., ter Vrugte, J., & de Jong, T. (2022). Supporting reflection to improve learning from self-generated concept maps. *Metacognition Learning*, 17, 691–713. DOI: 10.1007/s11409-022-09299-7
- Fend, H. (2013, September 2). *Individuelle Förderung: Ideen, Hintergründe und Fallstricke*. Bundeszentrale für politische Bildung. <https://www.bpb.de/themen/bildung/dossier-bildung/162108/individuelle-foerderung-ideen-hintergruende-und-fallstricke/>
- Fichten, W. (2010). Forschendes Lernen in der Lehrerbildung. In U. Eberhardt (Hrsg.), *Neue Impulse in der Hochschuldidaktik* (S. 127–182). VS Verlag.
- Franke–Braun, G., Schmidt-Weigand, F., Stäudel, L., & Wodzinski, R. (2008). Aufgaben mit gestuften Lernhilfen. In R. Messner, & W. Blum (Hrsg.), *Lernumgebungen auf dem Prüfstand* (S. 27–42). Kassel University Press.
- Friedrich, H.F., & Mandl, H. (2006). Lernstrategien: Zur Strukturierung des Forschungsfeldes. In H. Mandl, & H.F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 1–23). Hogrefe.
- García, J.A., Carcedo, R.J., & Castaño, J.L. (2019). The influence of feed-back on competence, motivation, vitality, and performance in a throwing task. *Research quarterly for exercise and sport*, 90(2), 172–179.
- Gogus, A. (2012). Bloom's taxonomy of learning objectives. In N.M. Seel (Hrsg.), *Encyclopedia of the Sciences of Learning* (S. 469–473). Springer. DOI: 10.1007/978-1-4419-1428-6_141
- González, H.L., Palencia, A.P., Umaña, L.A., Galindo, L., & Villafrade, M.L.A. (2008). Mediated learning experience and concept maps: a pedagogical tool for achieving meaningful learning in medical physiology students. *Advances in Physiology Education*, 32(4), 312–316. DOI:10.1152/advan.00021.2007
- Goossens, N.A.M.C., Camp, G., Verkoeijen, P.P.J.L., Tabbers, H.K., & Zwaan, R.A. (2014). The benefit of retrieval practice over elaborative restudy in primary school vocabulary learning. *Journal of Applied Research in Memory & Cognition*, 3, 177–182. DOI: 10.1016/j.jar-mac.2014.05.003
- Grassinger, R., Dickhäuser, O., & Dresel, M. (2019). Motivation. In D. Urhahne, M. Drese, & F. Fischer (Hrsg.), *Psychologie Für Den Lehrberuf* (S. 207–227). DOI: 10.1007/978-3-662-55754-9_11
- Gropengießer, H. (2013). Experimentieren. In H. Gropengießer, U. Harms, & U. Kattmann (Hrsg.), *Fachdidaktik Biologie* (9. Auf., S. 284–293). Aulis.

- Großmann, N., Eckes, A., & Wilde, M. (2020). Prädiktoren der Kompetenzwahrnehmung von Schülerinnen und Schülern im Biologieunterricht. *Zeitschrift für Pädagogische Psychologie*, 36(3), 153–165.
- Gruber, T. (2018). *Gedächtnis*. Springer Berlin Heidelberg. DOI: .1007/978-3-662-56362-5
- Gussen, L., Schumacher, F., Großmann, N., Ferreira González, L., Schlüter, K., & Grobschedl, J. (2023). Supporting pre-service teachers in developing research competence. *Frontiers in Education*, 8. DOI: 10.3389/educ.2023.1197938
- Hammann, M. (2019). Organisationsebenen biologischer Systeme unterscheiden und vernetzen: Empirische Befunde und Empfehlungen für die Praxis. In J. Groß, M. Hammann, P. Schmie-mann, & J. Zabel (Hrsg.), *Biologiedidaktische Forschung: Erträge für die Praxis* (S.77–91). DOI: 10.1007/978-3-662-58443-9_5
- Hammann, M., Phan, T.T.H., Ehmer, M., & Bayrhuber, H. (2006). Fehlerfrei experimentieren. *MNU*, 59(5), 292–299.
- Hänze, M., Schmidt-Weigand, F., & Stäudel, L. (2010). Gestufte Lernhilfen. In S. Boller, & R. Lau (Hrsg.), *Innere Differenzierung in der Sekundarstufe II. Ein Praxishandbuch für Lehrer/innen* (S. 63–73). Beltz Verlag.
- Harms, U., Mayer, J., Hammann, M., Bayrhuber, H., & Kattmann, U. (2004). Kerncurriculum und Standards für den Biologieunterricht in der gymnasialen Oberstufe. In H.E. Tenorth (Hrsg.), *Kerncurriculum Oberstufe II – Biologie, Chemie, Physik, Geschichte, Politik*. Weinheim. Beltz Verlag.
- Hattie, J. (2008). *Visible Learning*. Routledge. DOI: 10.4324/9780203887332
- Hattie, J. (2023). *Visible Learning: The Sequel*. DOI: 10.4324/9781003380542
- Hattie, J., & Clarke, S. (2018). *Visible Learning: Feedback*. Routledge. DOI: 10.4324/9780429485480
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. DOI: 10.3102/003465430298487
- Hattie, J.A.C., & Donoghue, G.M. (2016). Learning strategies: a synthesis and conceptual model. *Npj Science of Learning*, 1(1). DOI: 10.1038/npjscilearn.2016.13
- Herrmann, U. (2013, Oktober 23). *Zwischenruf: Lernen im digitalen Zeitalter*. Bundeszentrale für politische Bildung. <https://www.bpb.de/themen/bildung/dossier-bildung/170984/zwischen-ruf-lernen-im-digitalen-zeitalter/>
- Higgins, N., Frankland, S., & Rathner, J. (2021). Self-regulated learning in undergraduate science. *International Journal of Innovation in Science and Mathematics Education*, 29. DOI:10.30722/IJISME.29.01.005.
- Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S.I., & Reiss, K.M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, 153(103897). DOI: 10.1016/j.com-pedu.2020.103897

- Hofer, R. (2013). Forschendes Lernen in der Lehrerinnen- und Lehrerbildung: Widersprüchliche Anforderungen zwischen Forschung und Profession. *BzL - Beiträge Zur Lehrerinnen- Und Lehrerbildung*, 31(3), 310–320. DOI: 10.36950/bzl.31.3.2013.9634
- Horstmann, G., & Dreisbach, G. (2012). *Allgemeine Psychologie 2 kompakt* (2. Aufl.). Beltz.
- Hüfner, C., Isaak, R., & Wilde, M. (2020). Der Zusammenhang zwischen Elaborationsstrategien und Lernleistung im Biologieunterricht: Analysen mit handlungsfernen und handlungsnahen Erhebungsverfahren der Elaborationsstrategien. *Zeitschrift für Didaktik der Biologie (ZDB) – Biologie Lehren und Lernen*, 24, 14–28. DOI: 10.4119/zdb-3396
- Hülsmann, M., Stiller, C., & Wilde, M. (2024). The effect of incremental scaffolds in experimentation on cognitive load. *Science Journal of Education*, 12(1), 1–10. DOI: 10.11648/j.sjedu.20241201.11
- Hultberg, P., Santandreu Calonge, D., & Safiullin Lee, A.E. (2018). Promoting long-lasting learning through instructional design. *Journal of the Scholarship of Teaching and Learning*, 18(3). DOI: 10.14434/josotl.v18i3.23179
- Husna, H., Nerita, S., & Safitri, E. (2023). Analysis of student difficulties in learning biology. *Journal of Biology Education Research*, 4(1), 1–8. DOI: 10.55215/jber.v4i1.5963
- Jiang, D., & Kalyuga, S. (2020). Confirmatory factor analysis of cognitive load ratings supports a two-factor model. *Tutorial in Quantitative Methods for Psychology*, 16(3), 216–225. DOI: 10.20982/tqmp.16.3.p216
- Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2020). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology*, 1–14. DOI: 10.1037/edu0000627
- Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational Psychology Review*, 23(1), 1–19. DOI: 10.1007/s10648-010-9150-7
- Karpicke, J.D., & Blunt, J.R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772–775. DOI: 10.1126/science.1199327
- 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. DOI: 10.1080/001318800363908
- Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, 56(5), 746–759. DOI: 10.1177/0013164496056005002
- Kleinert, S.I. (2023). Förderung von Selbstregulation und Heterogenitätssensibilität im Fach Biologie zusammengedacht: Zur Wirksamkeit von gestuften Lernhilfen und Peer-Tutorien mit integriertem Lernstrategietraining als Fördermaßnahmen im Umgang mit heterogenem Lernverhalten in Schule und Universität sowie als Professionalisierungsmaßnahmen in der Lehrkräftebildung im Fach Biologie [Dissertation, Universität Bielefeld - Fakultät der Biologie].

- Kleinert, S.I., Besa, K.S., & Wilde, M. (2022). Der Einsatz gestufter Lernhilfen als Unterstützung für Lernende im Kontext des biologischen Experimentierens. Einfluss auf die Schüler*innen-Motivation im Naturwissenschafts- und Mathematikunterricht. *Zeitschrift für Didaktik der Biologie*, 26, 1-18.
- Klepsch, M., & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science*, 48(1), 45–77. DOI: 10.1007/s11251-020-09502-9.
- Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in psychology*, 8. DOI: 10.3389/fpsyg.2017.01997
- KMK [Kultusministerkonferenz] (Hrsg.) (2004). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss* (Beschluss der Kultusministerkonferenz vom 16.12.2004). KMK
- KMK [Kultusministerkonferenz] (Hrsg.) (2024). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung* (Beschluss der Kultusministerkonferenz vom 16.10.2008 i. d. F. vom 08.02.2024).
- Kranz, J., Baur, A., & Möller, A. (2022). Learners' challenges in understanding and performing experiments: a systematic review of the literature. *Studies in Science Education*, 59(2), 321–367. DOI: 10.1080/03057267.2022.2138151
- Krapp, A. (1992). Interesse, Lernen und Leistung. Neue Forschungsansätze in der Pädagogischen Psychologie. *Zeitschrift für Pädagogik*, 38(5), 747–770.
- Krapp, A. (1998). Entwicklung und Förderung von Interessen im Unterricht. *Psychologie in Erziehung und Unterricht*, 45, 186–203.
- Krathwohl, D.R. (2002). A revision of bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 212–218. DOI: 10.1207/s15430421tip4104_2
- Krieglstein, F., Schneider, S., Beege, M., & Rey, G.D. (2021). How the design and complexity of concept maps influence cognitive and motivational learning processes. *Educational technology research and development*, 70, 99–118, DOI: 10.21203/rs.3.rs-467941/v1
- Kuklick, L., & Lindner, M.A. (2021). Computer-based knowledge of results feedback in different delivery modes: Effects on performance, motivation, and achievement emotions. *Contemporary Educational Psychology*, 67, 1–17.
- Lenski, S.J. (2022). *Concept Maps im Biologieunterricht: Kognitive und emotionale Wirkung von Trainingsmaßnahmen und visuellen Gestaltungsmerkmalen* (Dissertation, Universität zu Köln - Institut für Biologiedidaktik).
- Leppink, J., Paas, F., Van der Vleuten, C.P.M., Van Gog, T., & Van Merriënboer, J.J.G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(4), 1058–1072. DOI: 10.3758/s13428-013-0334-1
- Lipnevich, A. A., & Panadero, E. (2021). A review of feedback models and theories: Descriptions, definitions, and conclusions. *Frontiers in Education*, 6. DOI: 10.3389/educ.2021.720195

- Looß, M. (2007). Lernstrategien, Lernorientierungen, Lern(er)typen. In *Theorien in Der Biologie-didaktischen Forschung* (S. 141–152). DOI: 10.1007/978-3-540-68166-3_13
- Luo, T., & Liu, C. (2023). The impact of feedback on metacognition: Enhancing in easy tasks, impeding in difficult ones. *Consciousness and Cognition*, 116(103601). DOI: 10.1016/j.con-cog.2023.103601
- Mandouit, L., & Hattie, J. (2023). Revisiting “The Power of Feedback” from the perspective of the learner. *Learning and Instruction*, 84(101718). DOI: 10.1016/j.learninstruc.2022.101718
- Marton, F., & Saljö, R. (1976). On qualitative differences in learning: I-outcome and process*. *British Journal of Educational Psychology*, 46(1), 4–11. DOI: 10.1111/j.2044-8279.1976.tb02980.x
- Marton, F., & Säljö, R. (1984). Approaches to learning. In F. Marton, D. Hounsell, & N. Entwistle (Hrsg.), *The experience of learning* (S. 39–58). Scottish Academic Press.
- Mayer, J. (2013). Erkenntnisse mit naturwissenschaftlichen Methoden gewinnen. In H. Gropengießer, U. Harms, U. Kattmann, D. Eschenhagen, S. Bögeholz, & R. Bühs (Hrsg.), *Fachdidaktik Biologie. Die Biologiedidaktik* (S. 56–61). Aulis.
- Mayer, R.E. (1998). Cognitive theory for education: What teachers need to know. In N. Lambert, & B.L. McCombs (Hrsg.), *How students learn: History, mathematics, and science in the classroom* (S. 351–377). American Psychological Association.
- Mayer, R.E. (2002). Cognitive theory and the design of multimedia instruction: An example of the two-way street between cognition and instruction. *New Directions for Teaching and Learning*, 2002(89), 55–71. DOI: 10.1002/tl.47
- 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. DOI: 10.1002/(SICI)1098-2736(199904)36:4<475::AID-TEA5>3.0.CO;2-O
- Merkt, M., Weigand, S., Heier, A., & Schwan, S. (2011). Learning with videos vs. learning with print: The role of interactive features. *Learning and Instruction*, 21, 687–704.
- Mertens, U., Finn, B., & Lindner, M.A. (2022). Effects of computer-based feedback on lower- and higher-order learning outcomes: A network meta-analysis. *Journal of Educational Psychology*, 114(8), 1743–1772.
- Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97.
- Misra, P.K. (2021). *Learning and teaching for teachers*. Springer Singapore. DOI: 10.1007/978-981-16-3077-4
- 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. DOI: 10.1002/j.2168-9830.2009.tb01011.x

- MSW NRW [Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen] (2013). *Kernlehrplan für die Sekundarstufe II Gymnasium / Gesamtschule in Nordrhein-Westfalen*. MSW NRW.
- Nesbit, J.C., & Adesope, O.O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, 76(3), 413–448. DOI: 10.3102/00346543076003413
- Noetel, M., Griffith, S., Delaney, O., Sanders, T., Parker, P., del Pozo Cruz, B., & Lonsdale, C. (2021). Video improves learning in higher education: A systematic review. *Review of Educational Research*, 91(2), 204–236.
- 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.
- 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/publications/researchpapers/theoryunderlyingconceptmaps.pdf>
- Novak, J.D. & Musonda, D. (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal*, 28, 117–153. DOI: 10.2307/1162881
- O’Day, G.M., & Karpicke, J.D. (2020). Comparing and combining retrieval practice and concept mapping. *Journal of Educational Psychology*, 113(5), 986–997. DOI: 10.1037/edu0000486
- Opitz, S.T., Neumann, K., Bernholt, S., & Harms, U. (2017). How do students understand energy in biology, chemistry, and physics? Development and validation of an assessment instrument. *EURASIA Journal of Mathematics, Science and Technology Education*, 13(7). DOI: 10.12973/eurasia.2017.00703a
- Pähl, V. (2005). Förderung der Empirischen Bildungsforschung durch das Bundesministerium für Bildung und Forschung (BMBF). In H. Mandl, & B. Kopp (Hrsg.), *Impulse für die Bildungsforschung: Stand und Perspektiven Dokumentation eines Expertengesprächs* (S. 118-124). Akademie Verlag. DOI: 10.1524/9783050085036.118
- Plass, J.L., & Kalyuga, S. (2019). Four ways of considering emotion in cognitive load theory. *Educational Psychology Review*, 31(2), 339–359. DOI: 10.1007/s10648-019-09473-5
- Posner, G.J., Strike, K.A., Hewson, P.W., & Gertzog, W.A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227.
- Purpura, J.E. (1998). Investigating the effects of strategy use and second language test performance with high- and low-ability test takers: a structural equation modelling approach. *Language Testing*, 15(3), 333–379. DOI: 10.1191/026553298669984422
- Reeve, J. (2015). *Understanding motivation and emotion* (6. Aufl.). Wiley.
- Reinhold, F., Hoch, S., Werner, B., Richter-Gebert, J., & Reiss, K. (2020). Learning fractions with and without educational technology: What matters for high-achieving and low-achieving students? *Learning and Instruction*, 65(101264). DOI: 10.1016/j.learninstruc.2019.101264

- Rey, G.D., & Buchwald, F. (2011). The expertise reversal effect: Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied*, 17(1), 33–48. DOI: 10.1037/a0022243
- Riewerts, K., Weiß, P., Wimmelmann, S., Saunders, C., Beyerlin, S., Gotzen, S., Linnartz, D., Thiem, J., & Gess, C. (2018). Forschendes Lernen entdecken, entwickeln, erforschen und evaluieren. *die hochschullehre*, 4, 389–406.
- Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, 77(1), 1–15. DOI: 10.1111/j.1467-8624.2006.00852.x
- Rittle-Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives*, 11(3), 184–190. DOI: 10.1111/cdep.12229
- 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. DOI: 10.1002/(sici)1098-2736(199608)33:6<569::aid-tea1>3.0.co;2-m
- Ryan, A., & Deci, E.L. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogische Psychologie*, 39(2), 223–238.
- Ryan, R.M., & Deci, E.L. (2002). Overview of self-determination theory: An organismic-dialectical perspective. In E.L. Deci, & R.M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 3–33). University of Rochester Press.
- Sadava, D., Hillis, D.M., Heller, H.C., & Hacker, S.D. (2019). *Purves Biologie*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-58172-8
- Saye, J.W., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia supported learning environments. *Educational Technology Research and Development*, 50(3), 77–96.
- Schaffner, E., & Schiefele, U. (2002). Die PISA-Studie: Schülerleistungen im internationalen Vergleich. In G. Cramer, H. Schmidt, & W. Wittwer (Hrsg.), *Ausbilder-Handbuch* (S. 1–20). Deutscher Wirtschaftsdienst.
- Schmidt-Weigand, F., Franke-Braun, G., & Hänze, M. (2008). Erhöhen gestufte Lernhilfen die Effektivität von Lösungsbeispielen? *Unterrichtswissenschaft*, 36(4), 365–384.
- Schmidt-Weigand, F., Hänze, M., & Wodzinski, R. (2009). Complex problem solving and worked examples. The role of prompting strategic behavior and fading-in solution steps. *Zeitschrift für Pädagogische Psychologie*, 23(2), 129–138. DOI: 10.1024/1010-0652.23.2.129.
- Schmuckler, M.A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, 2(4), 419–436. DOI: 10.1207/S15327078IN0204_02
- Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, 19(4), 469–508. DOI: 10.1007/s10648-007-9053-4S
- Schroeder, N.L., Nesbit, J.C., Anguiano, C.J., & Adesope, O.O. (2017). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review*, 30(2), 431–455. DOI: 10.1007/s10648-017-9403-9

- Smith III, J.P., diSessa, A.A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Journal of the Learning Sciences*, 3(2), 115–163. DOI: 10.1207/s15327809jls0302_1
- Sosa, G., Berger, D., & Saw, A. (2007). Effectiveness of computer-based instruction in statistics: A meta-analysis [dataset]. *American Psychological Association*. DOI: 10.1037/e655882007-001
- Stanat, P., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Schümer, G., Tillmann, K.J., & Weiß, M. (2003). In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.J. Tillmann, M. Weiß (Hrsg.), *PISA und PISA-E: Zusammenfassung der bereits vorliegenden Befunde. PISA 2000 — Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 51–75). DOI: 10.1007/978-3-322-97590-4_3
- Stäudel, L., Franke-Braun, G., & Schmidt-Weigand, F. (2007). Komplexität erhalten - auch in heterogenen Lerngruppen: Aufgaben mit gestuften Lernhilfen. *CHEMKON*, 14(3), 115–122. DOI: 10.1002/ckon.200710058
- Stiller, C., & Wilde, M. (2021). Einfluss gestufter Lernhilfen als Unterstützungsmaßnahme beim Experimentieren auf den Lernerfolg im Biologieunterricht. *Zeitschrift Für Erziehungswissenschaft*. DOI: 10.1007/s11618-021-01017-4
- Stiller, C., Allmers, T., Habigsberg, A., Stockey, A., & Wilder, M. (2020). Erkenntnisgewinnung in den Naturwissenschaften: Von der Hypothese zur Theorie. *PFLB*, 2(2), 28-39. DOI: 10.4119/pflb-3302
- Streblov, L., & Schiefele, U. (2006). Lernstrategien im Studium. In H. Mandl, & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 352-364). Hogrefe.
- Svicher, A., Gori, A., & Di Fabio, A. (2022). The sustainable development goals psychological inventory: A network analysis in italian university students. *International Journal of Environmental Research and Public Health*, 19(17). DOI: 10.3390/ijerph191710675
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123–138. DOI: 10.1007/s10648-010-9128-5
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2), 261–292. DOI: 10.1007/s10648-019-09465-5
- Sweller, J., van Merriënboer, J.J. G., & Paas, F.G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296. DOI: 10.1023/A:1022193728205
- Szempruch, J. (2022). The goals of modern education in changing society. In B. Pituala, & I. Nowosad (Hrsg.), *Education – Multiplicity of Meanings, Commonality of Goals* (S. 15–32). V&R unipress.
- Taxipulati, S., & Lu, H. (2021). The influence of feedback content and feedback time on multimedia learning achievement of college students and its mechanism. *Frontiers in Psychology*, 7, 1–14.

- Thomas, A.E., Müller, F.H., & Bieg, S. (2018). Entwicklung und Validierung der Skalen zur motivationalen Regulation beim Lernen im Studium (SMR-LS). *Diagnostica*, 64(3), 145–155. DOI: 10.1026/0012-1924/a000201
- Thorpe, M (2022). Digital and online learning. In J. Lord (Hrsg.), *Psychology of education* (S. 383–403). SAGE publishing.
- Tillmann, K. J. (2015, Dezember 3). *PISA & Co. – eine kritische Bilanz. Bundeszentrale für politische Bildung. Dossier Bildung.* <https://www.bpb.de/themen/bildung/dossier-bildung/208550/pisa-co-eine-kritische-bilanz/>
- Tsui, C., & Treagust, D. (2009). Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. *International Journal of Science Education*, 32(8), 1073–1098. DOI: 10.1080/09500690902951429
- Unterbruner, U. (2007). Multimedia-Lernen und Cognitive Load. In *Theorien in Der Biologiedidaktischen Forschung* (S. 153–164). DOI: 10.1007/978-3-540-68166-3_13
- 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. DOI: 10.1007/s10648-010-9127-6
- Van der Kleij, F.M., Feskens, R.C.W., & Eggen, T.J.H.M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research*, 85(4), 474–511.
- Vincent-Lancrin, S. (2022). Educational innovation and digitalisation during the COVID-19 crisis: lessons for the future. *OECD*. DOI: 10.1787/93c3dc5e-en
- Virtbauer, L., Schiffli, I., & Scheicher, E. (2019). Interesse und Motivation. In I. Schiffli, & H. Weiglhofer (Hrsg.), *Biologie kompetent unterrichten: Ein Praxisbuch für Studierende und Lehrkräfte* (S. 49-67). facultas.
- Wadouh, J., Liu, N., Sandmann, A., & Neuhaus, B.J. (2013). The effect of knowledge linking levels in biology lessons upon students' knowledge structure. *International Journal of Science and Mathematics Education*, 12(1), 25–47. DOI: 10.1007/s10763-012-9390-8
- Wai, H.O., & Khine, S.S. (2020). An investigation into the difficulties of students in learning biology. *Art Sci*, 18(9).
- Wang, M., Wu, B., Kirschner, P.A., & Spector, J.M. (2018). Using cognitive mapping to foster deeper learning with complex problems in a computer-based environment. *Computers in Human Behavior*, 87, 450–458. DOI: 10.1016/j.chb.2018.05.021
- Weidlich, J., & Spannagel, C. (2014). Die Vorbereitungsphase im Flipped Classroom. Vorlesungsvideos versus Aufgaben. In K. Rummler (Hrsg.), *Lernräume gestalten -Bildungskontexte vielfältig denken* (S. 237–248). Waxmann.
- Weinert, F.E. (2001). *Leistungsmessungen in Schulen*. Beltz.
- Weinstein, C.E., & Mayer, R.E. (1986). The teaching of learning strategies. In M.C. Wittrock (Hrsg.), *Handbook of research on teaching: A project of the American Educational Research Association* (3. Aufl., S. 315–327). Macmillan.

- Weitzel, H. (2014). Welche Bedeutung haben vorunterrichtliche Vorstellungen für das Lernen? In U. Spörhase (Hrsg.), *Biologie Didaktik. Praxishandbuch für die Sekundarstufe I und II* (S. 62-81). Cornelsen.
- Welter, V., Großschedl, J., & Schlüter, K. (2020). Biologie in der Lehrerinnen- und Lehrerbildung. Perspektiven einer kompetenzorientierten Fachdidaktik. Handbuch Lehrerinnen- Und Lehrerbildung. DOI: 10.35468/hblb2020-04
- Whalley, S. (2022). Special educational needs and disability (SEND) and inclusion. In J. Lord (Hrsg.), *Psychology of education* (S. 302–321). SAGE publishing.
- Wiklund-Hörnqvist, C., Stillesjö, S., Andersson, M., Jonsson, B., & Nyberg, L. (2022). Retrieval practice is effective regardless of self-reported need for cognition - Behavioral and Brain Imaging Evidence. *Frontiers in Psychology, 12*. DOI: 10.3389/fpsyg.2021.797395
- Wild, K.P. (2005). Individuelle Lernstrategien von Studierenden. Konsequenzen für die Hochschuldidaktik und die Hochschullehre. *BzL, 23*(2), 191–206. DOI: 10.36950/bzl.23.2.2005.10089
- 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.
- Wilde, M., Bätz, K., Kovaleva, A., & Urhahne, D. (2009). Überprüfung einer Kurzskala intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften, 15*, 31–45.
- Winn, A. S, DelSignore, L., Marcus, C., Chiel, L., Freiman, E., Stafford, D., & Newman, L. (2019). Applying cognitive learning strategies to enhance learning and retention in clinical teaching settings. *MedEdPORTAL, 15*(10850). DOI: 10.15766/mep_2374-8265.10850
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, 1–14.
- Wulf, C., Thiem, J., Gess, C. (2020). Motivationale Faktoren im Wirkungskontext von Forschendem Lernen. In C. Wulf, S. Haberstroh, M. Petersen (Hrsg.), *Forschendes Lernen* (S. 129–144). Springer VS.
- Zakrajsek, T. D. (2023). Developing your learning strategy. In T.D. Zakrajsek (Hrsg.), *The New Science of Learning* (3. Aufl., S. 46–62). DOI: 10.4324/9781003447986-6
- Zander, S. (2010). *Motivationale Lernervoraussetzungen in der Cognitive Load Theory*. Logos-Verlag.
- Zimbardo, P.G., Gerrig, R.J., & Graf, R. (2008). *Psychologie* (18. Aufl.). Pearson Studium.
- Zoelch, C., Berner, V.D., & Thomas, J. (2019). Gedächtnis und Wissenserwerb. In D. Urhahne, M. Drese, & F. Fischer (Hrsg.), *Psychologie Für Den Lehrberuf* (S. 23–44). DOI: 10.1007/978-3-662-55754-9_2
- Zuo, M., Kong, S., Ma, Y., Hu, Y., & Xiao, M. (2023). The effects of using scaffolding in online learning: A meta-analysis. *Education Sciences, 13*(7). DOI: 10.3390/educsci13070705

Danksagung

Ein besonderer Dank geht an **Prof. Dr. Jörg Großschedl**, der meine Dissertation betreut und mich durch den (teilweise) beschwerlichen Weg begleitet und gelenkt hat. Als besonders hilfreich empfand ich seine fachliche Unterstützung mit wertvollen Anregungen und konstruktiven Rückmeldungen sowie seinen Zuspruch, der mich stets ermutigte und inspirierte. Mein herzlicher Dank gilt auch seiner Bereitschaft, geduldig Antworten auf meine (unzähligen) Fragen zu finden. Auch wenn es nicht einfach war, gelang es ihm, meine Bedenken zu zerstreuen und mich mit seinen scharfsinnigen und sogar amüsanten Bemerkungen zu ermutigen.

Auch möchte ich mich bei **Prof. Dr. Daniela Schmeinck** bedanken, die sich bereit erklärt hat, meine Arbeit als Zweitgutachterin zu betreuen. Der Austausch mit ihr über meine Forschungsprojekte, ihre wertschätzenden Rückmeldungen und ihre wertvollen Anregungen waren besonders zielführend für den Verlauf meiner Promotion.

Ich möchte auch **Prof. Dr. Benjamin Rott** für seine Bereitschaft danken, den Vorsitz für meine Disputation zu übernehmen und **Dr. Fabian Seredszus** für seine Bereitschaft, den Beisitz zu übernehmen.

Zudem möchte ich den Vorsitzenden und Mitgliedern der Graduiertenschule für ihre Unterstützung und den hilfreichen und anregenden Austausch während meiner Promotion danken.

Bedanken möchte ich mich auch bei meinen Kolleg:innen, die mich während jeder Phase des Promotionsvorhabens tatkräftig unterstützt haben. Hier möchte ich mich besonders bei Lea Gussen und Roxanne Gutowski bedanken, mit denen ich im Rahmen der Promotion viel Zeit verbringen durfte und die mich stets unterstützt und bestärkt haben. Ich danke euch für euer offenes Ohr und eure Bereitschaft, fachliche und nicht-fachliche Themen zu eruieren. Auch bei meiner Büro-Kollegin Laura Leiss möchte ich mich für ihren stetigen Zuspruch, ihre Unterstützung und ihre bürointernen Anregungen bedanken. Danken möchte ich auch Dr. Helena Aptyka, die mich tatkräftig unterstützt und bestärkt hat und mit der jede Prüfungsphase eine tolle Erfahrung war. Besonderer Dank gilt auch Dr. Nadine Großmann, die mir durch ihre statistische Expertise und ihre forschungsbezogenen Erfahrungen sehr oft aus schwierigen Situationen verhelfen konnte. Ihre wertschätzende und motivierende Unterstützung war für mich sehr gewinnbringend.

Bedanken möchte ich mich zudem bei allen Kolleg:innen der Biologiedidaktik, die ich während meiner Promotionszeit kennenlernen durfte und die mir hilfreiche Anregungen gaben.

Hier möchte ich mich bei der Arbeitsgruppe von Prof. Dr. Kirsten Schlüter und Dr. Silvia Fränkel bedanken. Auch bei universitätsexternen Kolleg:innen, die ich auf Konferenzen kennenlernen durfte, möchte ich mich bedanken. Besonderer Dank gilt hier Dr. Svea Isabel Kleinert für ihre Bereitschaft, diese Arbeit Korrektur zu lesen.

Ich möchte mich auch bei den Studienteilnehmer:innen für ihre Bereitschaft bedanken, sich an den Studien zu beteiligen.

Zuletzt möchte ich den Menschen danken, die mich nicht nur in den Phasen der Promotion, sondern in jeder Phase meines Lebens unterstützt haben. Besonderer Dank gilt meinen Eltern, die mir jeden steinigen Weg geebnet haben, die als Leitfiguren für mich fungieren und die ich an meiner Seite nie missen musste - *Ju faleminderoj*. Meinen Brüdern möchte ich danken für ihre oft allzu ehrlichen Rückmeldungen, ihre unermüdliche Unterstützung und dass sie ein verlässlicher Ankerpunkt in meinem Leben sind. Auch meinen Freunden, besonders meiner Kindheitsfreundin Sissi, möchte ich danken, dass sie mich stets unterstützt und motiviert haben. Zu diesen Freunden zähle ich nun auch viele meiner neu-gewonnenen Kolleg:innen der Biologiedidaktik.

Vielen Dank Ihnen und euch.

Erklärung zur Verfügbarkeit von Primärdaten

Die im Rahmen dieser Arbeit erhobenen und ausgewerteten Primärdaten wurden am Institut für Biologiedidaktik der Universität zu Köln archiviert. Die Einsichtnahme in das Studienmaterial und der Zugriff auf die digitalisierten Dateien kann bei Prof. Dr. Jörg Großschedl (E-Mail: j.grosssc@uni-koeln.de) erfragt werden.

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:

Mustafa, M., Großmann, N., & Großschedl, J. (2025). *Rückmeldesysteme in interaktiven Lernvideos für angehende Biologielehrkräfte: Einfluss auf die Forschungskompetenz, das forschungsmethodische Wissen und die intrinsische Motivation*. In S. Schaal, A. Lude, M. Krell & K. Kremer (Hrsg.), *Lehr- und Lernforschung in der Biologiedidaktik* (S. 181-195). StudienVerlag.

Mustafa, M., Ioannidis, A., Ferreira González, L., Dabrowski, T., & Großschedl, J. (2021). Fostering learning with incremental scaffolds during chemical experimentation: A study on junior high school students working in peer-groups. *International Journal of Innovation in Science and Mathematics Education*, 29(2), 19-31. DOI: 10.30722/ijisme.29.02.002

Lenski, S., Mustafa, M., & Großschedl, J. (2024). Concept Mapping – Increased Potential as a Retrieval-based Task. *Memory*, 1–12. DOI: 10.1080/09658211.2024.2363463

Erkelenz, den 25.06.2024

Mirlinda Mustafa