

**MULTIPLE EXTERNE REPRÄSENTATIONEN
IN DER BIOLOGIE –
EMPIRISCHE BEFUNDE ZU REDUNDANZ UND
SELBSTWIRKSAMKEITSERWARTUNG**

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Lara Magnus

Kiel, 2022

Erste Gutachterin:

Prof. Dr. Julia Schwanewedel

Zweiter Gutachter:

Prof. Dr. Moritz Krell

Tag der mündlichen Prüfung:

05.04.2022

ZUSAMMENFASSUNG

In den Naturwissenschaften wird mit Texten und vielfältigen Abbildungstypen kommuniziert. Letztere eignen sich etwa zur Darstellung nicht sichtbarer Prozesse oder großer Datenmengen. Abbildungen, Texte und weitere Darstellungsformen werden als externe Repräsentationen (ER) bezeichnet, kombiniert mit Bezug zu einem gemeinsamen Bezugsobjekt als multiple externe Repräsentation (MER, z. B. eine Text-Bild-Kombination). Im naturwissenschaftlichen Unterricht können ER dem Aufbau konzeptuellen Verständnisses dienen. Zudem ist kompetentes Arbeiten mit ihnen Teil naturwissenschaftlicher Grundbildung, die zu gesellschaftlicher Teilhabe befähigen soll, die das Verständnis eines öffentlichen Diskurses zu naturwissenschaftlichen Themen mittels verschiedener ER einschließt. Auch im Sinne wissenschaftspropädeutischer Bildung sollten Schüler:innen lernen, mit verschiedenen (M)ER zurechtzukommen, auch mit fachspezifischen, komplexen Abbildungen. Studien zeigen aber, dass derartige Abbildungen selbst bei Studierenden Schwierigkeiten hervorrufen können. MER verschiedener Medien (z. B. wissenschaftliche Fachartikel, Lehrbücher) können sich zudem in Redundanz – Wiederholung von Informationen in ER einer MER – unterscheiden, sodass Personen nach der Schullaufbahn aus wissenschaftspropädeutischer Sicht mit verschiedenen redundanten MER arbeiten können sollten.

Diese Arbeit untersucht für das Fach Biologie in Studie 1, inwieweit Personen nach dem Abitur (Biologiestudierende) mit unterschiedlich redundanten MER arbeiten können, und in Studie 2, ob Schwierigkeiten mit komplexen Abbildungen in der Biologie auch auf persönliche Überzeugungen (hier: Selbstwirksamkeitserwartung, SWE) zurückzuführen sein könnten, mit solchen Abbildungen arbeiten zu können. Beiden Studien liegt Testmaterial basierend auf zwei Abituraufgaben zugrunde, variiert für drei Experimentalgruppen: zwei MER-Gruppen (*keine Redundanz*, *Redundanz*), eine *nur Text*-Gruppe. Wenn die Aufgabe fachspezifische, komplexe Abbildungen enthielt, erreichte Gruppe *keine Redundanz* die geringste Punktzahl, Gruppe *nur Text* ohne diese Abbildungen die höchste. Auch wurde für Gruppe *keine Redundanz*, die zur Aufgabenlösung Informationen aus derartigen Abbildungen entnehmen musste, ein positiver Effekt genannter SWE auf die Testleistung nachgewiesen. Schüler:innen sollten im Biologieunterricht stärker an fachspezifische, komplexe Abbildungen und unterschiedlich redundante MER herangeführt werden. Eine bewusste Stärkung der SWE bezüglich komplexer Abbildungen könnte gewinnbringend sein. Diese Arbeit leistet einen Beitrag zur evidenzbasierten Forschung zu (M)ER in der Biologie mit potentiell auf weitere naturwissenschaftliche Fächer übertragbaren Ergebnissen.

ABSTRACT

In science, communication takes place using texts and various types of pictures. The latter are useful, for example, for presenting invisible processes or large amounts of data. Pictures, texts and other forms of presentation are referred to as external representations (ERs), and when combined with reference to a common reference object as multiple external representations (MERs; e.g., a text–picture combination). In science lessons, ERs can serve to build conceptual understanding. In addition, being able to work with ERs competently is part of scientific literacy, which should enable social participation, including the understanding of a public discourse on scientific topics by means of various ERs. Further, from a science propaedeutics point of view, students should learn at school to work with different (M)ERs, including subject-specific, complex pictures. Studies show, however, that subject-specific, complex pictures can cause difficulties even for undergraduate students. Moreover, MERs of different media (e.g., scientific articles, textbooks) can differ in redundancy—repetition of information in ERs within the same MER—and therefore, from a science propaedeutics point of view, high school graduates should be able to work with MERs that differ in their redundancy.

This project investigates specifically for the subject of biology in Study 1 to what extent undergraduate students in biology i.e., students that obtained the higher education entrance qualification, the German Abitur, can work with MERs that differ in their redundancy, and in Study 2 whether difficulties with complex pictures in biology could partially be due to an individual's belief (in this study: self-efficacy belief, SE) in their ability to work with such pictures. Both studies are based on test material derived from two Abitur exam tasks, which differed for three experimental groups: two MER groups (*no-redundancy* and *redundancy*) and a *text-only* group. When the task contained subject-specific complex pictures, the *no-redundancy* group achieved the lowest test score, the *text-only* group without these pictures the highest. Moreover, for the *no-redundancy* group, which had to extract information from the subject-specific, complex pictures to solve the task, the above mentioned SE had a positive effect on test performance. These findings suggest that students should be introduced more explicitly to biology-specific, complex pictures and MERs that differ in redundancy in biology lessons at school. Helping students strengthen their SE with respect to complex pictures could be beneficial, as well. This work contributes to evidence-based research on (M)ERs in biology, and findings can potentially be transferred to other science subjects.

INHALTSVERZEICHNIS

ZUSAMMENFASSUNG	I
ABSTRACT.....	III
1 Einleitung.....	1
2 Theoretischer Hintergrund	5
2.1 Repräsentationen als Kommunikationsmittel in den Naturwissenschaften: Bedeutung für die Fachdisziplinen und Lehr-Lernkontexte.....	5
2.2 Klassifikation von Repräsentationen	8
2.2.1 Externe und interne Repräsentationen	9
2.2.2 Deskriptionale und depiktionale Repräsentationen	11
2.3 Multiple externe Repräsentationen	15
2.3.1 Modelle der kognitiven Verarbeitung von Texten und Bildern.....	15
2.3.1.1 <i>Cognitive Theory of Multimedia Learning</i>	16
2.3.1.2 <i>Integrated Model of Text and Picture Comprehension</i>	20
2.3.1.3 <i>Cognitive Load Theory</i>	21
2.3.2 <i>DeFT (Design, Functions, Tasks) Framework</i> nach Ainsworth (2006)	24
2.3.3 Befunde zu Potentialen, Schwierigkeiten und Integrationshilfen bezüglich MER-Einsatz zu naturwissenschaftlichen Inhalten	29
2.3.4 Redundanz in multiplen externen Repräsentationen	32
2.4 Repräsentationskompetenz	36
2.5 Selbstwirksamkeitserwartung und ihre Bedeutung für Leistungen in mathematisch-naturwissenschaftlichen Fächern.....	41
2.5.1 Begriffserläuterung und theoretische Einbettung	42
2.5.2 Charakteristika der Selbstwirksamkeitserwartung	43
2.5.3 Quellen der Selbstwirksamkeitserwartung	46
2.5.4 Empirische Befunde zum Zusammenhang von Selbstwirksamkeitserwartung und akademischer Leistung im mathematisch-naturwissenschaftlichen Bereich.....	51
3 Überblick über die durchgeführten Studien mit ausdifferenzierter Zielsetzung ...	55

4 Studie 1: Challenges solving science tasks with text–picture combinations persist beyond secondary school	61
4.1 Introduction	62
4.2 Theoretical Framework	63
4.2.1 Types of External Representations and their Importance in Biology Contexts	63
4.2.2 Processing MERs	64
4.2.3 Redundancy in MERs	65
4.2.4 Purpose of This Study	67
4.3 Method	68
4.3.1 Participants.....	68
4.3.2 Procedure	68
4.3.3 Material	69
4.3.4 Scoring	73
4.3.5 Control Measures	74
4.3.5.1 Cognitive Abilities	74
4.3.5.2 Working Memory Capacity	75
4.3.5.3 Additional Covariates	75
4.3.6 Data Analysis	75
4.4 Results.....	76
4.4.1 Baseline Equivalence	76
4.4.2 Effects of Group on Biology Test Performance.....	78
4.4.2.1 Effects of Group by Task	78
4.4.2.2 Effects of Group by Task and Type of Representation.....	78
4.4.3 Analyses of Covariance Including Students’ Cognitive Abilities.....	80
4.5 Discussion	81
4.5.1 Effect of the Task Material	81
4.5.2 Comparison of Results of Task A and B.....	85
4.5.3 Strength and Limitations	86
4.5.4 Conclusion and Implications.....	88

4.6 Acknowledgments	89
4.7 Conflict of Interest Statement.....	89
4.8 References	90
4.9 Supplementary Material	95
4.9.1 Supplementary Material S1: Detailed Representations of the Data by Boxplots	95
4.9.2 Supplementary Tables	98
5 Studie 2: Fachspezifische komplexe Abbildungen verarbeiten – (auch) eine Frage der Überzeugung.....	101
5.1 Einleitung	102
5.1.1 Die Rolle und Effekte von Abbildungen in den Naturwissenschaften	103
5.1.2 Die Beziehung zwischen Selbstwirksamkeitserwartung und akademischer Leistung	105
5.1.3 Fragestellung und Hypothese	105
5.2 Methode.....	106
5.2.1 Stichprobe und Studiendurchführung.....	106
5.2.2 Messinstrumente.....	107
5.2.2.1 Selbstwirksamkeitserwartung.....	107
5.2.2.2 Kognitive Fähigkeiten	108
5.2.2.3 Biologienote	108
5.2.3 Material und Aufgabenbewertung	108
5.2.4 Regressionsanalysen.....	111
5.3 Ergebnisse.....	112
5.3.1 Multiple Regression für Aufgabe A	112
5.3.1.1 Berücksichtigung des Repräsentationstyps.....	114
5.3.1.2 Berücksichtigung objektiver Leistungsmaße.....	116
5.3.2 Multiple Regression für Aufgabe B.....	118
5.4 Diskussion	118
5.4.1 Moderatoreffekt.....	119
5.4.2 Limitationen	120
5.4.3 Implikationen.....	121

5.5 Danksagung.....	123
5.6 Literatur.....	124
6 Zusammenfassung der Ergebnisse der Studien	129
7 Diskussion und Ausblick	133
7.1 Studienübergreifende Diskussion der Ergebnisse	133
7.1.1 Potentielle Ursachen für beobachtete Schwierigkeiten mit bildlichen Repräsentationen.....	133
7.1.2 Auswirkungen von Redundanz in MER und MER gegenüber reinem Text.....	141
7.2 Stärken und Limitationen.....	146
7.3 Implikationen für die Praxis.....	150
7.4 Implikationen für zukünftige Forschung.....	155
7.5 Ausblick	159
8 Literatur.....	163
ABBILDUNGSVERZEICHNIS.....	177
TABELLENVERZEICHNIS.....	179
ABKÜRZUNGSVERZEICHNIS	181
DANKSAGUNG	183
ERKLÄRUNG	185

1 Einleitung

Diagramme, Fotografien, Tabellen und Texte sind im alltäglichen Leben in verschiedener Art und Kombination allgegenwärtig. Diese und weitere sogenannte Repräsentationen, textliche wie bildliche, besitzen jeweils bestimmte Eigenschaften und eine damit verbundene Aussagekraft. Auch wenn Folgendes in die Rubrik anekdotische Evidenz fällt, kann bildlichen Repräsentationen geradezu eine weltpolitische Bedeutung beigemessen werden. Um die Aufgabe zu bewältigen, dem 45. Präsidenten der Vereinigten Staaten von Amerika Informationen zu vermitteln, verfolgte der damalige US-Außenminister Tillerson folgende Strategie: „[...] and then I started taking charts and pictures with me because I found that those seemed to hold his attention better. If I could put a photo or a picture in front of him or a map or a piece of paper that had two big bullet points on it, he would focus on that, and I could build on that“ (Bjorklund, 2021). Zusätzliche bildliche Repräsentationen können also etwas leisten, das Wörter allein nicht leisten können. Eine Kombination aus verschiedenen Repräsentationen, zum Beispiel aus einem Text und einer bildlichen Repräsentation, wird als multiple Repräsentation bezeichnet (Ainsworth, 2006). Multiple Repräsentationen sind, wie im Folgenden deutlich wird, auch in den Naturwissenschaften von zentraler Bedeutung. In den Naturwissenschaften sind bildliche Repräsentationen in besonderer Vielzahl und Bandbreite allgegenwärtig (Arsenault et al., 2006) und werden häufig mit Texten kombiniert. Warum bildliche Repräsentationen dabei so bedeutsam sind, welche Funktionen sie beispielsweise übernehmen können bzw. welchen Mehrwert gegenüber Texten sie haben können, lässt sich anhand der öffentlichen Kommunikation zur COVID-19-Pandemie nachzeichnen. Um etwa die Funktionsweise eines mRNA-Impfstoffes darzustellen, werden häufig schematische Abbildungen einer menschlichen Zelle genutzt und eingezeichnet, wie mRNA an Ribosomen in Spike-Protein übersetzt wird. Hier hilft die Abbildung, einen für das menschliche Auge nicht sichtbaren Prozess sichtbar zu machen. Ebenso werden vielfach Achsendiagramme zur Darstellung von Studienergebnissen, zum Beispiel zur Viruslast verschiedener Altersgruppen, eingesetzt. Die bildliche Repräsentation dient hier dazu, große Datenmengen darzustellen und Datenmuster erkennbar werden zu lassen. Letzteres zeigt, dass bildliche Repräsentationen in den Naturwissenschaften nicht nur als Darstellungs-, sondern auch als Erkenntnismittel fungieren können (z. B. Heßler, 2006). Weitere Potentiale verschiedener Repräsentationsarten werden in verschiedenen Abschnitten der vorliegenden Arbeit thematisiert. Verfolgt man die öffentliche Debatte zur

Einleitung

COVID-19-Pandemie, wird unter anderem in Kommentaren in den sozialen Medien deutlich, dass gerade bildliche Repräsentationen, die aus wissenschaftlichen Studien stammen, häufig nicht richtig verstanden werden. Im schlechtesten Fall werden fachlich falsche Aussagen abgeleitet. Informationen zu naturwissenschaftlichen Sachverhalten aus verschiedenen Repräsentationen entnehmen, in Verbindung setzen und interpretieren zu können, sollte im Rahmen einer naturwissenschaftlichen Grundbildung in der Schule erlernt werden, um zu einer aktiven gesellschaftlichen Teilhabe zu befähigen (z. B. Yore et al., 2007).

Naturwissenschaftlicher Unterricht ist stark vom Einsatz verschiedener Repräsentationsarten geprägt (Tsui & Treagust, 2013). Zum Teil sind diese fachspezifisch (für die Biologie z. B. Stammbäume, Genkarten), zum Teil aber auch durch andere Fachdisziplinen oder Präsenz im gesellschaftlichen Leben bekannt (z. B. Fotos, Diagramme, Texte, Formeln). Mit unterschiedlichen Repräsentationen und Kombinationen von Repräsentationen arbeiten zu können, ist in den Naturwissenschaften somit wichtig für die Kommunikation über naturwissenschaftliche Sachverhalte sowie den Aufbau eines konzeptuellen Verständnisses (z. B. Yore & Hand, 2010). Die Bedeutung von Repräsentationen spiegelt sich auch in den normativen Vorgaben in Curricula für naturwissenschaftliche Fächer in vielen Ländern wider, in denen ein kompetenter Umgang mit verschiedenen Repräsentationsarten von Schüler:innen explizit gefordert wird (z. B. KMK, 2005, 2020; NGSS Lead States, 2013). In den deutschen Bildungsstandards für das Fach Biologie, das in der vorliegenden Dissertation fokussiert wird, ist das Arbeiten mit Repräsentationen ein zentraler Aspekt des Kompetenzbereichs Kommunikation (KMK, 2005, 2020). Wie oben bereits beschrieben ist diese Kompetenz auch über die Schulzeit hinaus für eine gesellschaftliche Teilhabe von Bedeutung, besonders aber auch gerade für die Personen relevant, die ein Studium und einen Berufsweg in den Naturwissenschaften, zum Beispiel der Biologie, anstreben. Ziel des Oberstufenunterrichts ist auch wissenschaftspropädeutische Bildung (KMK, 2006, 2021). Die Befähigung zum kompetenten Umgang mit verschiedenen Repräsentationen ist ein wichtiger Teil der Ausbildung zukünftiger Biolog:innen, die für die Bewältigung biologiebezogener gesellschaftlicher Herausforderungen – etwa bezüglich Gesundheit, Nahrungsmitteln, Energie und Umwelt – benötigt werden (Treagust & Tsui, 2013). Schüler:innen sollten also durch Fachunterricht in der Schule in die Lage versetzt werden, mit verschiedenen Kombinationen aus Repräsentationen selbstständig arbeiten zu können und auch fachspezifische Abbildungen zu verstehen. Verschiedene empirische Studien haben allerdings gezeigt, dass Schüler:innen und auch Studierende teilweise Schwierigkeiten mit bildlichen bzw. multiplen Repräsentationen in naturwissenschaftlichen Kontexten haben,

zum Beispiel bei der Informationsentnahme (für Biologie z. B. Novick et al., 2012; Ziepprecht, 2016). Gerade fachspezifische Abbildungen können Schwierigkeiten hervorrufen (z. B. Novick et al., 2012). Die Informationen einzelner Repräsentationen, die gemeinsam einen Sachverhalt darstellen, können sich in unterschiedlichem Maße überschneiden, das heißt, die Redundanz kann unterschiedlich ausgeprägt sein (Ainsworth, 1999). Bezüglich dieses Merkmals von MER können sich auch verschiedene Medien unterscheiden, mit denen Schüler:innen im schulischen Kontext oder im gesellschaftlichen Leben sowie einem potentiellen anschließenden naturwissenschaftlichen Studium konfrontiert sind – z. B. Schulbücher, Arbeitsblätter, Lehrbücher für Studierende, Kursmaterialien, populärwissenschaftliche oder wissenschaftliche Artikel (für Schulbücher und wissenschaftliche Artikel z. B. Roth et al., 1999). Naturwissenschaftlicher Fachunterricht wie der Biologieunterricht in der Schule sollte Schüler:innen auch mit Blick auf eine naturwissenschaftliche Grundbildung und wissenschaftspropädeutische Bildung befähigen, mit verschiedenartigen multiplen Repräsentationen arbeiten zu können. In der vorliegenden Arbeit wird deshalb mit Fokus auf schriftliche Kommunikation untersucht, inwiefern Biologiestudierende, die also die Schule mit Erwerb der Hochschulzugangsberechtigung beendet haben, in der Lage sind, Biologieaufgaben mit redundanten oder nicht-redundanten Text-Bild-Kombinationen zu bearbeiten. Zudem wird in dieser Arbeit als eine potentielle Ursache für Schwierigkeiten, die bei fachspezifischen, komplexen Abbildungen auftreten können, die Selbstwirksamkeitserwartung von Personen, mit komplexen Abbildungen in der Biologie zurechtzukommen, untersucht, da erfolgreiches Arbeiten mit solchen Repräsentationen nicht unbedingt allein auf Wissen über derartige Repräsentationen oder fachliches Vorwissen zurückzuführen sein muss. Selbstwirksamkeitserwartungen – also die Überzeugung, durch eigenes Handeln ein bestimmtes Ergebnis erzielen zu können (Bandura, 2001) – können positiv mit akademischer Leistung in naturwissenschaftlichen Fächern zusammenhängen, wie bereits vielfach empirisch belegt wurde (z. B. Ardura & Pérez-Bitrián, 2019; Britner & Pajares, 2006). Ob eine Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie einen Effekt auf die Leistung hat, wenn Informationen aus derartigen Abbildungen entnommen werden müssen, wurde noch nicht erforscht und wird vor dem Hintergrund der Bedeutung derartiger Abbildungen in der Biologie und einer möglichen Relevanz dieser Selbstwirksamkeitserwartung in Lehr-Lernkontexten im Bereich der Biologie in dem in dieser Dissertation dargelegten Promotionsprojekt untersucht.

Einleitung

Im Folgenden wird zunächst der theoretische Hintergrund dieser Arbeit erläutert (Kapitel 2). Dabei wird zuerst auf die Bedeutung von Repräsentationen für naturwissenschaftliche Fachdisziplinen und Lehr-Lernkontexte eingegangen (Kapitel 2.1). Es folgt die Klassifikation von Repräsentationen, die auf verschiedenen Ebenen vorgenommen werden kann (Kapitel 2.2). Das darauffolgende Kapitel (2.3) nimmt die für diese Arbeit zentralen multiplen Repräsentationen in den Blick: die Verarbeitung von Texten und Bildern nach kognitionspsychologischen Modellen, die Funktionen von multiplen Repräsentationen, empirische Befunde zum Arbeiten von Schüler:innen und Studierenden mit multiplen Repräsentationen zu naturwissenschaftlichen Inhalten und die in Studie 1 dieser Arbeit fokussierte Redundanz in multiplen Repräsentationen. Im Anschluss werden das Konzept der Repräsentationskompetenz sowie zugehörige empirische Befunde aus dem naturwissenschaftlichen Bereich vorgestellt (Kapitel 2.4). Im letzten Kapitel des theoretischen Hintergrunds wird auf das Konzept der Selbstwirksamkeitserwartung eingegangen, auf ihre theoretische Einbettung, Charakteristika und Quellen sowie auf empirische Befunde zum Zusammenhang von Selbstwirksamkeitserwartung und akademischer Leistung im mathematisch-naturwissenschaftlichen Bereich (Kapitel 2.5). Im anschließenden Kapitel 3 folgt ein Überblick über die in diesem Promotionsprojekt durchgeführten Studien mit ausdifferenzierten Zielsetzungen und Forschungsfragen beziehungsweise Hypothese. In Kapitel 4 und 5 sind die zugehörigen Artikel eingefügt. Diese werden in Kapitel 6 zusammengefasst. Zuletzt erfolgen eine studienübergreifende Diskussion und die Ableitung von Implikationen für die Praxis – sowohl für Schulunterricht und Hochschullehre im Fach Biologie als auch die Lehrkräfteaus- und Weiterbildung – und zukünftige Forschung mit abschließendem Ausblick (Kapitel 7).

2 Theoretischer Hintergrund

2.1 Repräsentationen als Kommunikationsmittel in den Naturwissenschaften: Bedeutung für die Fachdisziplinen und Lehr-Lernkontexte

In den Naturwissenschaften erfolgt die Kommunikation mit Hilfe verschiedenartiger Repräsentationen, zum Beispiel mit Texten, Tabellen, Abbildungen oder Gleichungen (z. B. Lemke, 1998). Häufig werden Repräsentationen dabei kombiniert. So nutzen Forschende beispielsweise beim Austausch mit Kolleg:innen spontan auch verschiedene bildliche Repräsentationen (Kozma & Russell, 2005) und kommunizieren ihre Forschungsergebnisse in wissenschaftlichen Fachartikeln unter Nutzung einer Vielzahl verschiedener Repräsentationsarten (Lemke, 1998). Die Anzahl eingesetzter Abbildungen und die Bandbreite verwendeter Abbildungsarten in den Naturwissenschaften übersteigen Anzahl und Bandbreite in anderen Wissenschaften, wie beispielsweise Arsenault et al. (2006) bei der vergleichenden Analyse von Fachjournalen der *Hard-Science*-Disziplinen Physik, Chemie, Biologie und Medizin¹ und der *Soft-Science*-Disziplinen Psychologie, Ökonomie und Soziologie zeigten. In der Disziplin Biologie helfen verschiedenartige Abbildungen unter anderem dabei, mit bloßem Auge nicht erkennbare Strukturen sichtbar zu machen, und bilden Phänomene auf verschiedenen Organisationsebenen, der makroskopischen, mikroskopischen und submikroskopischen Ebene, ab (Schönborn & Anderson, 2006). Schönborn und Anderson (2006) machen dies am Beispiel der Biochemie deutlich. Hier könnten Fotos einer Mikrotiterplatte aus dem Labor, Mikroskopaufnahmen, schematische Darstellungen von Proteinstrukturen oder Graphen, die eine Immunreaktion anhand einer Farbstoffabsorption sichtbar machen, eingesetzt werden. Die Beispiele zeigen, dass die Abbildungen nicht nur hinsichtlich der Organisationsebene des repräsentierten Phänomens variieren, sondern sich auch hinsichtlich des Abstraktionsgrads unterscheiden: Sie befinden sich auf einem Kontinuum von abstrakt (z. B. Graph) bis hin zu realistisch (z. B. Mikroskopaufnahme; Schönborn & Anderson, 2006). Im Vergleich zu Abbildungen, die allgemein im gesellschaftlichen Leben präsent sind (z. B. bildliche Beschreibungen, Fotos oder Graphen in der Zeitung), sind Abbildungen in den Naturwissenschaften häufig komplexer, da oft

¹ Arsenault et al. begründen die Zuordnung von Medizin, einer angewandten Wissenschaft, zu den *Hard-Science*-Disziplinen in ihrer Studie durch die enge Verbindung von Medizin zu den Disziplinen Biologie und Chemie.

eine Vielzahl spezifischer grafischer Konventionen genutzt wird oder sie sehr viele Datenpunkte oder viele Variablen abbilden (Lowe, 2000; Trafton et al., 2002). Der Einsatz von beispielsweise Graphen macht deutlich, dass Wissenschaftsbilder auch dem Erkenntnisgewinn dienen (Heßler, 2006). Denn zum Teil werden erst durch die graphische Darstellung von Messdaten oder durch bestimmte visuelle Verfahren (z. B. Röntgen) (Daten-)Muster oder Strukturen erkennbar, die neue Erkenntnisse ermöglichen (Heßler, 2006). Historisch betrachtet lässt sich festhalten, dass neu entwickelte Repräsentationsarten (z. B. Strukturformeln) oder durch die Einführung von Computern ermöglichte Visualisierungsformen (z. B. 3D-Modelle auch großer Moleküle mit Möglichkeit zur Analyse von Bindungslängen und -winkeln) neue Denkansätze liefern und somit auch zur Entwicklung einer Disziplin beitragen können (Kozma & Russell, 2005).

Nicht nur in der Kommunikation zwischen Wissenschaftler:innen oder zwischen Wissenschaftler:innen und der Öffentlichkeit (z. B. zum Klimawandel, Kommunikation zur COVID-19-Pandemie, s. Kapitel 1) spielen verschiedene Repräsentationsarten eine große Rolle. Sie sind auch ein zentrales Element in Lehr-Lernkontexten in den Naturwissenschaften (z. B. Tsui & Treagust, 2013) und tragen unter anderem zum Aufbau konzeptuellen Wissens bei. Ein Ziel schulischer Bildung ist zudem, dass Schüler:innen eine *Scientific Literacy* bzw. naturwissenschaftliche Grundbildung aufbauen, wobei dem kompetenten Umgang mit verschiedenen Repräsentationen eine wichtige Rolle zukommt (z. B. Norris & Phillips, 2003; Stäudel et al., 2008; Yore et al., 2007). Stäudel et al. (2008) merken an, dass Naturwissenschaften heutzutage eine hohe Bedeutung für „Gesellschaft, Technik und Wirtschaft, wie auch für das Leben der Einzelnen“ (S. 4) haben und eine „[n]aturwissenschaftliche Grundbildung für alle [...] daher eine selbstverständliche Forderung, sowohl von Seiten der Wirtschaft, um über hinreichend gebildete Arbeitskräfte verfügen zu können, wie auch für den Einzelnen als Grundlage einer aktiven politischen Teilhabe und für reflektierte, ganz praktische Entscheidungen im Alltag“ (S. 4) sei, und betonen die Wichtigkeit des Aufbaus einer domänenspezifischen Lesefähigkeit, um dieses Ziel zu erreichen (Stäudel et al., 2008). Norris und Phillips (2003) sehen die Fähigkeit zum Lesen und Schreiben naturwissenschaftlicher Texte, die neben Wörtern auch zum Beispiel Abbildungen, Tabellen und Gleichungen enthalten können, als grundlegenden Aspekt (*fundamental sense*) von Scientific Literacy an. Dabei vertreten Norris und Phillips die Auffassung, dass Lesen nicht die pure Dekodierung von Wörtern oder Entnahme von Informationen sei, sondern ein konstruktiver Prozess, nämlich die Interpretation des Dargestellten in

Abhängigkeit vom Hintergrundwissen der Lesenden und unter Berücksichtigung der Intention der Autor:innen. Auf diesem grundlegenden Aspekt von Scientific Literacy basieren abgeleitete Aspekte (*derived sense*) wie die Erkenntnisfähigkeit bezüglich der Naturwissenschaften (*scientific knowledgeability*; Norris & Phillips, 2003). Schüler:innen sollten somit mit Charakteristika und Bandbreite bildlicher Repräsentationen vertraut gemacht werden, sie für das Verständnis von Phänomenen heranziehen und auch lernen, sie so zu nutzen, wie es Naturwissenschaftler:innen tun (Eilam & Gilbert, 2014). Die Fähigkeit, verschiedene Repräsentationen interpretieren, konstruieren, vergleichen, auswählen, und zueinander in Beziehung setzen oder ineinander umwandeln zu können, wird als repräsentationale Kompetenz bezeichnet (Kozma & Russell, 1997). Diese wird in Kapitel 2.4 genauer in den Blick genommen.

Die Befähigung von Schüler:innen, mittels verschiedener Repräsentationen Informationen im naturwissenschaftlichen Kontext wie beispielsweise der Biologie zu erschließen und darzustellen, ist auch in verschiedenen nationalen Bildungsstandards verankert (z. B. für Deutschland: KMK, 2005, 2020; für UK: Department for Education, 2015; für die USA: NGSS, 2013). In den deutschen Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss sowie für die Allgemeine Hochschulreife wird der Umgang mit Repräsentationen als Teil der Kommunikationskompetenz aufgeführt, die mit den weiteren Kompetenzbereichen Sach-², Erkenntnisgewinnungs- und Bewertungskompetenz die Fachkompetenz abbildet (KMK, 2005, 2020). Die Kommunikationskompetenz wird wie folgt beschrieben:

„Die Kommunikationskompetenz der Lernenden zeigt sich in der Kenntnis von Fachsprache, fachtypischen Darstellungen und Argumentationsstrukturen und in der Fähigkeit, diese zu nutzen, um fachbezogene Informationen zu erschließen, adressaten- und situationsgerecht darzustellen und auszutauschen.“ (KMK, 2020, S. 10)

Schüler:innen sollen also sowohl Rezipient:innen sein, die Informationen aus verschiedenen Darstellungsformen erschließen, als auch zur Konstruktion verschiedener Darstellungsformen befähigt sein, um Informationen aufzubereiten und auszutauschen (KMK, 2020). Während in den Bildungsstandards im Fach Biologie für den Mittleren Schulab-

² Der Kompetenzbereich *Sachwissen* (KMK, 2020) wird in den Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss (KMK, 2005) noch als *Fachwissen* bezeichnet.

schluss bereits auf die Diversität der zu verwendenden Repräsentationen (z. B. Texte, Schemazeichnungen, Graphen, Tabellen) hingewiesen wird, wird in den Bildungsstandards für die Allgemeine Hochschulreife explizit auch die häufige Komplexität von in der Biologie eingesetzten Darstellungsformen thematisiert (KMK, 2005, 2020). Laut Bildungsstandards ist die Kommunikationskompetenz Teil der von Schüler:innen zu erwerbenden naturwissenschaftlichen Kompetenz. Letztere „bietet Orientierung in der durch Naturwissenschaften und Technik geprägten Lebenswelt, eröffnet Perspektiven für die berufliche Orientierung und schafft Grundlagen für selbstgesteuertes, lebenslanges, globales und soziales Lernen“ (KMK, 2020, S. 10). Lehrende in den Naturwissenschaften sollten Lernende also explizit darin unterrichten, wie naturwissenschaftliche Fachsprache mit den verschiedenen Repräsentationstypen zu lesen ist, wie Verbindungen zwischen Repräsentationen hergestellt werden können und wie sie selbst Repräsentationen anfertigen können (Lemke, 2004). Ob dieses normative Ziel in der Unterrichtspraxis tatsächlich erreicht wird und Lernende dementsprechend in der Lage sind, mit verschiedenartigen Repräsentationen kompetent zu arbeiten, war und ist in verschiedenen Kontexten und mit unterschiedlichen Schwerpunktsetzungen Untersuchungsgegenstand wissenschaftlicher Studien, so auch der vorliegenden Arbeit.

2.2 Klassifikation von Repräsentationen

Wie bereits beschrieben sind in den Naturwissenschaften wie der Biologie verschiedene Repräsentationsarten omnipräsent und in der Kommunikation³ über naturwissenschaftliche Sachverhalte elementar. Dies gilt sowohl für die einzelnen Fachdisziplinen als auch für Lehr-Lernkontexte zum Beispiel im Bereich der Biologie und so sind Repräsentationen auch in der biologiedidaktischen bzw. naturwissenschaftsdidaktischen Forschung ein wichtiger Forschungsgegenstand. Im Folgenden werden der Begriff der Repräsentation und die Klassifikation von Repräsentationen betrachtet.

Der Begriff der Repräsentation umfasst, dass ein zu repräsentierender Sachverhalt durch eine Repräsentation abgebildet wird, wobei eine Abbildungsrelation bestimmt, welche Charakteristika des Sachverhalts mit welchen Charakteristika der Repräsentation kor-

³ Die vorliegende Dissertation fokussiert auf Formen schriftlicher Kommunikation (z. B. geschriebener Text, Abbildungen). Aus Gründen der Vollständigkeit wird in diesem und nachfolgenden Kapiteln auch mündliche Kommunikation (d. h. gesprochene Wörter, Sätze) eingeschlossen, wenn diese Bestandteil von Klassifikationen oder Modellen ist.

respondieren (Schnotz, 1996). Die Klassifikation von Repräsentationen erfolgt auf mehreren Ebenen. Zunächst wird zwischen internen, mentalen Repräsentationen und externen Repräsentationen unterschieden: Ausgehend von externen Repräsentationen (z. B. Texten, Diagrammen), die verschiedene Sinnesmodalitäten ansprechen können (zumeist die visuelle oder auditive), konstruiert das Individuum interne, mentale Repräsentationen (Schnotz, 2001). Für eine weitere Einteilung hat sich eine aus der Kognitionspsychologie stammende Klassifikation nach Zeichensystemen in der fachdidaktischen Forschung etabliert (Ziepprecht et al., 2017): Basierend auf einer auf Peirce (1906) zurückgehenden Unterscheidung von Symbolzeichen und ikonischen Zeichen wird zwischen deskriptionalen und depiktionalen Repräsentationen unterschieden (Schnotz, 2001). Die genannten Einteilungen von Repräsentationen sowie das Kombinieren mehrerer Repräsentationen zu sogenannten multiplen externen Repräsentationen werden in den folgenden Unterkapiteln genauer erläutert.

2.2.1 Externe und interne Repräsentationen

Um über Sachverhalte zu kommunizieren, nutzen Menschen Sprachäußerungen (z. B. gesprochenen oder geschriebenen Text) oder verschiedene Arten von Bildern wie Fotografien, Graphen und Diagramme (Schnotz, 2001, 2003). In der Disziplin Biologie sowie Lehr-Lernkontexten im Fach Biologie werden neben Text zum Beispiel Schemata von Prozessen wie der Zellteilung oder Zellatmung, Mikroskopbilder, Fotos bestimmter Pflanzen- oder Tierarten, Stammbäume oder Strukturformeln von Molekülen eingesetzt sowie Achsendiagramme zur Darstellung einer Beziehung zwischen Variablen (z. B. Räuber und Beute). Auf Basis dieser externen Repräsentationen entwickeln Rezipierende interne mentale Repräsentationen des dargestellten Sachverhalts (Schnotz, 2003). Dies wird in Kapitel 2.3.1 genauer erläutert. Externe und interne Repräsentationen interagieren miteinander: Auf Basis externer Repräsentationen erfolgt die Bildung interner Repräsentationen, die durch bestehende interne Repräsentationen beeinflusst werden kann. Auf Grundlage interner Repräsentationen können wiederum externe Repräsentationen konstruiert werden (Krey & Schwanewedel, 2018). So konstruieren Schüler:innen zum Beispiel bei der Bearbeitung von Testaufgaben ausgehend von internen Repräsentationen zu einem Sachverhalt, die sie mittels externer Repräsentationen aufgebaut haben, in ihrer Antwort wiederum externe Repräsentationen wie Texte oder auch bildliche Repräsentationen (z. B. Zeichnungen, Diagramme). Neben der Informationsentnahme aus und Interpretation von gegebenen

textlichen und bildlichen Repräsentationen (z. B. in Schulbüchern und weiterem Lernmaterial) ist auch die Konstruktion von bildlichen Repräsentationen für das Lernen im naturwissenschaftlichen Unterricht bedeutsam (Ainsworth et al., 2011). Im empirischen Teil dieser Arbeit liegt der Fokus allerdings auf Informationsentnahme und Interpretation von Text-Bild-Material.

Externe Repräsentationen können sowohl statisch als auch dynamisch sein. Der vielseitige Einsatz statischer Repräsentationen in der Biologie und im Biologieunterricht (z. B. Mikroskopaufnahmen, Fotos, Schemazeichnungen, Stammbäume, Achsendiagramme) und deren Nutzen wurden bereits herausgearbeitet (vgl. auch Kapitel 1 und 2.1). Dynamische Repräsentationen wie Animationen sind geeignet, um sich verändernde, also dynamische Sachverhalte darzustellen und sind dabei realitätsnäher als es statische Repräsentationen wären (Schnotz, 2001). Während der COVID-19-Pandemie wurde beispielsweise das Öfteren die Entwicklung des Infektionsgeschehens über die Zeit auf einer Deutschlandkarte als Animation dargestellt. Allerdings wird bei Animationen die genaue Analyse der Repräsentation erschwert, da die Reizquelle flüchtig ist. Dem kann zum Beispiel entgegengewirkt werden, indem die Möglichkeit zum Anhalten der Animation gegeben wird (Schnotz, 2001). Trotz der weltweit voranschreitenden Digitalisierung und damit einhergehender Möglichkeiten zum Einsatz dynamischer Repräsentationen, auch zum Beispiel in elektronischen Lehrbüchern, sind statische Repräsentationen sowohl im gesellschaftlichen Leben (z. B. in Zeitungen, als Beschilderungen) als auch in der wissenschaftlichen Kommunikation und der Lehre (z. B. in Fachpublikationen und herkömmlichen Schul-/Lehrbüchern, Arbeitsblättern, Testmaterialien) allgegenwärtig. Bezüglich gedruckter versus elektronischer Lehrbücher ergab zum Beispiel eine Befragung von Studierenden von Millar und Schrier (2015), dass die Mehrheit der Befragten gedruckte Lehrbücher elektronischen Lehrbüchern vorzieht. Die Autor:innen gehen allerdings davon aus, dass sich dies in der nächsten Generation, die bereits ab der Einschulung mit elektronischen Lehrbüchern konfrontiert wird und daher mit ihnen vertraut ist, ändern wird. In der vorliegenden Arbeit wurde die Rezeption statischer externer Repräsentationen untersucht. Sowohl bei externen als auch bei internen Repräsentationen kann zwischen deskriptionalen und depiktionalen Repräsentationen unterschieden werden, wie nachfolgend erläutert wird.

2.2.2 Deskriptionale und depiktionale Repräsentationen

Eine weitere Einteilung von Repräsentationen basiert auf der auf Peirce (1906) zurückgehenden Kategorisierung von Zeichensystemen in Symbolzeichen und ikonische Zeichen (Schnotz, 2001). Peirce (1906) führt des Weiteren noch Indexzeichen auf, die direkt mit einem individuellen Objekt verbunden sind, wie zum Beispiel Rauch als Zeichen für Feuer (Schnotz, 1994), auf die im Folgenden aber nicht weiter eingegangen werden soll. Symbolzeichen weisen eine willkürliche Struktur auf, sehen dem Gegenstand, den sie bezeichnen, nicht ähnlich, sondern sind über Konventionen mit ihm verbunden (Schnotz, 2001). Dies gilt zum Beispiel für gesprochene oder geschriebene Wörter, Sätze und Texte sowie Gleichungen (Schnotz, 2001; Schnotz & Bannert, 2003; Tab. 2.2.1). So hat zum Beispiel das Wort *Buch* keine Ähnlichkeit mit dem von ihm bezeichneten Gegenstand. Zu geschriebenem Text wird auch die Brailleschrift gezählt (Schnotz, 2005), die von Personen mit Beeinträchtigung des Sehvermögens tastend wahrgenommen wird. Wird ein Sachverhalt mittels Symbolzeichen repräsentiert, bezeichnet Schnotz (2001) dies als deskriptionale Repräsentation (vgl. Abb. 2.2.1).

Anders als Symbolzeichen haben ikonische Zeichen im engeren Sinne keine willkürliche Struktur, sondern weisen eine Ähnlichkeit mit dem repräsentierten Sachverhalt oder Gegenstand auf (Schnotz, 2001). Dies trifft beispielsweise auf Fotografien oder (Schema-)Zeichnungen zu, sogenannte realistische Bilder (Schnotz, 2001). Realistische Bilder in der Disziplin Biologie oder Lehr-Lernkontexten im Fach Biologie sind zum Beispiel Fotografien oder Schemazeichnungen von Tieren, Pflanzen oder Pflanzenkrankheiten, Genkarten oder Landkarten mit eingezeichneter Verbreitung bestimmter Arten. Wird der Begriff der ikonischen Zeichen etwas weiter gefasst, nämlich als Zeichen, die mit dem repräsentierten Sachverhalt auf Basis gemeinsamer Strukturmerkmale verbunden sind, können auch sogenannte logische Bilder wie Diagramme zu den ikonischen Zeichen gezählt werden (Schnotz, 2001; Tab. 2.2.1). Schnotz (2001) bezeichnet realistische und logische Bilder als depiktionale Repräsentationen (vgl. Abb. 2.2.1). Logische Bilder sind durch bestimmte Analogierelationen mit dem repräsentierten Sachverhalt verknüpft, was eine abstrakte Form struktureller Gemeinsamkeit darstellt, während die Ähnlichkeit, die realistische Bilder zum repräsentierten Sachverhalt aufweisen, einer konkreteren Form struktureller Gemeinsamkeit entspricht (Schnotz, 2001). Genauer betrachtet lassen sich depiktionale Repräsentationen bezüglich ihres Abstraktionsgrads entlang eines Kontinuums anordnen: Von *konkret* zu *abstrakt* angeordnet würde beispielsweise auf eine Fotografie einer

Rose eine Schemazeichnung einer Rose folgen, gefolgt von einem Graphen, in dem die Höhe der Rose gegen die Jahreszeit aufgetragen ist (Pozzer-Ardenghi & Roth, 2010)⁴.

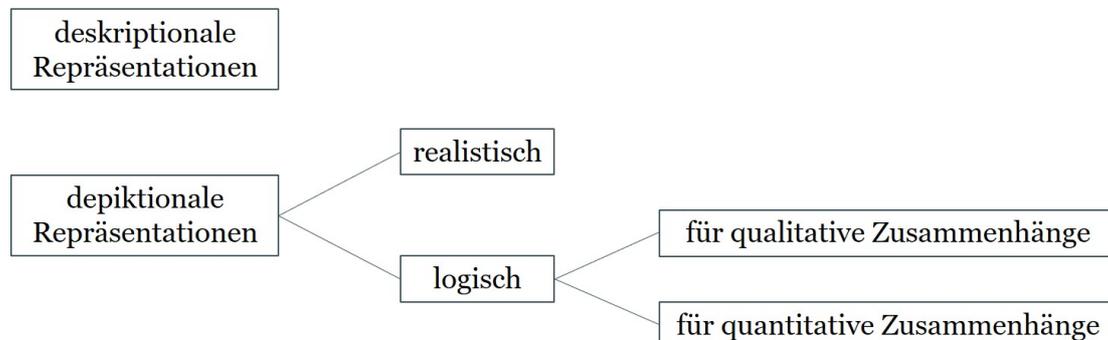


Abbildung 2.2.1. Klassifikation von Repräsentationen basierend auf verschiedenen Zeichensystemen (nach Schnotz 1994, 2001).

Logische Bilder visualisieren Zusammenhänge zwischen Variablen, wobei diese Zusammenhänge qualitativer (z. B. durch Fluss- und Baumdiagramme visualisiert) oder quantitativer (z. B. durch Linien-, Säulen- oder Streudiagramme visualisiert) Natur sind (Lachmayer et al., 2007; Schnotz, 1994; vgl. Abb. 2.2.1). Konkrete Beispiele im Bereich der Biologie sind Stammbäume, die qualitative Zusammenhänge darstellen, und das besagte Räuber-Beute-Achsendiagramm, das einen quantitativen Zusammenhang zwischen Variablen visualisiert. Bei der Darstellung qualitativer Zusammenhänge müssen beispielsweise Pfeile oder Verbindungslinien eines Diagramms kontextbezogen interpretiert werden. So repräsentieren Pfeile in einem Nahrungsnetz eine Nahrungsbeziehung (Igel → Spinne = Igel frisst Spinne), während Pfeile im Citratzyklus für die Umwandlung eines Moleküls in ein anderes Molekül stehen. Darstellungen quantitativer Zusammenhänge sind hingegen nach festgelegten Regeln zu interpretieren oder zu konstruieren (Lachmayer et al., 2007). Zum Beispiel werden in Linien- und Streudiagrammen Datenpunkte durch ihre relative Lage zu den Achsen des Diagramms charakterisiert (Schnotz, 1994).

Mayer (2011) zählt Tabellen zu den Bildern, die er als räumliche Repräsentation charakterisiert. Auch Winn (1987) zählt zumindest Tabellen, die Beziehungen zwischen

⁴ Pozzer-Ardenghi und Roth (2010) benutzen für bildliche Darstellungen den Begriff *Inscriptions*. Dieser umfasst alle Repräsentationstypen außer Text. Anders als depiktionale Repräsentationen nach Schnotz schließen *Inscriptions* also auch Gleichungen ein. Deren Abstraktionsgrad fügt sich auf dem genannten Kontinuum über dem von Graphen ein.

kategorialen Variablen darstellen (z. B. verschiedene Charakteristika unterschiedlicher Vitamine), zu graphischen Formen, die zwar abstrakt sind, aber eine sinntragende räumliche Anordnung besitzen. Während also in Linien- und Streudiagrammen Beziehungen durch die relative Lage der Datenpunkte zu den Achsen ausgedrückt werden, werden in Tabellen Beziehungen durch die relative Lage von Elementen in den Spalten und Reihen dargestellt (Winn, 1987). Aufgrund dieser Überlegungen von Mayer (2011) und Winn (1987) und einer entsprechenden Zuordnung in der biologiedidaktischen Arbeit von Ziepprecht et al. (2017) werden in der vorliegenden Arbeit Tabellen den logischen depiktionalen Repräsentationen zugeordnet, auch wenn Schnotz (1994) selbst diese Einordnung nicht vornimmt. In Tabelle 2.2.1 sind Charakteristika von deskriptionalen und depiktionalen Repräsentationen zusammengefasst sowie für die Disziplin Biologie relevante Beispiele aufgeführt.

Tabelle 2.2.1

Charakteristika und Beispiele deskriptionaler und depiktionaler Repräsentationen (basierend auf Schnotz, 2001)

Repräsentation	deskriptional	depiktional	
Zugrunde liegendes Zeichensystem	Symbolzeichen	ikonische Zeichen	
Verknüpfung mit bezeichnetem Gegenstand/ Sachverhalt	durch Konvention	aufgrund gemeinsamer Strukturmerkmale	
(Form struktureller Gemeinsamkeit/ Depiktionstyp)	–	konkret/ realistische Depiktion	abstrakt/ logische Depiktion
Für die Biologie relevante Beispiele ^a	<ul style="list-style-type: none"> • Wort^b/ Text^b • Brailleschrift • (Reaktions-) Gleichung • Summenformel 	<ul style="list-style-type: none"> • Fotografie • Mikroskopbild • Gelbild • Schemazeichnung • (Gen-)Karte 	<ul style="list-style-type: none"> • Flussdiagramm • Baumdiagramm • Balken-/Säulendiagramm • Box-Plot • Heatmap • Linien-/Flächendiagramm • Tabelle^c

^aDie Beispiele wurden teilweise Lachmayer et al. (2007) und Schnotz (2005) entnommen sowie durch Analyse von Fachpublikationen in der Disziplin Biologie zusammengetragen. ^bgesprochen oder geschrieben. ^cDie Einordnung von Tabellen als logische Depiktionen in dieser Arbeit beruht auf Arbeiten von Mayer (2011) und Winn (1987), die Tabellen aufgrund der räumlichen Anordnung ihrer Elemente zu Bildern bzw. graphischen Formen zählen.

Im Folgenden wird wie von Schnotz (2005) aufgeführt das übergreifende jeweilige Potential von deskriptionalen und depiktionalen Repräsentationen gegenübergestellt. Durch Deskriptionen lassen sich abstrakte Konzepte, zum Beispiel *Säugetiere* oder *Reptilien*, einfacher ausdrücken als durch Depiktionen, durch die in diesem Fall jeweils nur ein bestimmtes Säugetier oder ein bestimmtes Reptil dargestellt werden könnte. Auch Disjunktionen wie die Aussage *Säugetiere oder Reptilien* sind textlich besser darstellbar als durch ein Bild. Ein Charakteristikum von Depiktionen ist, dass sie informationell vollständig sind, d. h. es werden unweigerlich Informationen über die Form von Elementen, die Größe und die Orientierung im Raum o. Ä. vermittelt (Schnotz, 2005). Eine realistische Zeichnung von einem Baum auf einer Wiese zeigt zwangsläufig, wie hoch er ist, wo auf der Wiese er lokalisiert ist und gegebenenfalls auch dessen Färbung, während diese Informationen in derselben Genauigkeit durch Text viel schwerer zu vermitteln wären.⁵ Besonders offensichtlich wird dies auch bei Land- oder Genkarten. Wie in Kapitel 2.1 bereits beschrieben wurde liegt ein weiteres Potential von wissenschaftlichen Bildern, also von Depiktionen, darin, dass sie einen Erkenntnisgewinn herbeiführen können (z. B. Röntgenbilder).

In der Kommunikation in naturwissenschaftlichen Fachdisziplinen und naturwissenschaftlichem Unterricht ist es üblich, Kombinationen aus mehreren der genannten Repräsentationen zu nutzen (vgl. Kapitel 2.1). Diese Kombinationen mehrerer Repräsentationen zu einem gemeinsamen Bezugsobjekt werden als multiple externe Repräsentationen (MER) bezeichnet (Ainsworth, 2006; Krey & Schwanewedel, 2018). Eine MER kann aus mehreren Texten bzw. Deskriptionen, mehreren Depiktionen oder Text-Bild-Kombinationen bestehen (Krey & Schwanewedel, 2018). Die Informationspräsentation unter Verwendung unterschiedlicher Zeichensystemen, wie es bei Text-Bild-Kombinationen der Fall ist, versteht Schnotz (2001) als semiotischen Aspekt multimedialer (Lern-)Umgebungen. Die kognitive Verarbeitung von MER sowie Charakteristika, Funktionen, deren gewinnbringendes Potential wie auch durch sie hervorgerufene Schwierigkeiten sind Gegenstand des nächsten Kapitels.

⁵ Anmerken möchte ich hier, dass durch Vereinfachungen oder Auslassungen in – auch realistischen – Depiktionen durchaus Informationen über den repräsentierten Gegenstand (in diesem Beispiel etwa die genaue Baumart) fehlen können, weshalb der Ausdruck *informationell vollständig* womöglich irreführend sein könnte.

2.3 Multiple externe Repräsentationen

MER können wie beschrieben aus mehreren Deskriptionen, Depiktionen oder Text-Bild-Kombinationen bestehen und sind in den Naturwissenschaften und beim Lehren und Lernen naturwissenschaftlicher Inhalte allgegenwärtig. Im Folgenden werden zunächst Modelle der kognitiven Verarbeitung von Texten und Bildern und daraus abgeleitete Gestaltungsprinzipien multimedialen Materials vorgestellt (Kapitel 2.3.1). Im Anschluss wird ein Rahmenwerk zum Lernen mit MER eingeführt, das auch Funktionen von MER berücksichtigt (Kapitel 2.3.2). Es folgt die Vorstellung von Befunden zu Potentialen, Schwierigkeiten und Integrationshilfen bezüglich MER zu naturwissenschaftlichen Themen (Kapitel 2.3.3). Abschließend wird aufgrund der Relevanz für diese Dissertation ein Charakteristikum von MER, die Redundanz, detailliert betrachtet (Kapitel 2.3.4).

2.3.1 Modelle der kognitiven Verarbeitung von Texten und Bildern

Bevor auf unterschiedliche Modelle der kognitiven Verarbeitung von Texten und Bildern näher eingegangen wird, soll zunächst der Aufbau des menschlichen Gedächtnisses in den Blick genommen werden. Ein gängiges Modell ist das von Atkinson und Shiffrin (1968) vorgeschlagene Mehrspeichermodell mit drei interagierenden Speichersystemen (Abb. 2.3.1): Vereinfacht gesagt werden Stimuli aus der Umwelt im sensorischen Register registriert und ausgewählte Informationen an einen Kurzzeitspeicher, das Arbeitsgedächtnis, weitergegeben. Dort werden Informationen verarbeitet und dazu gegebenenfalls Informationen aus dem Langzeitspeicher hinzugezogen. Nach der Verarbeitung können Informationen vom Arbeitsgedächtnis in den Langzeitspeicher transferiert werden. Während die Informationen im sensorischen Register und im Arbeitsgedächtnis nur sehr kurz verbleiben, können sie im Langzeitspeicher langfristig gespeichert werden, wobei sie allerdings noch modifiziert werden können (Atkinson & Shiffrin, 1968).

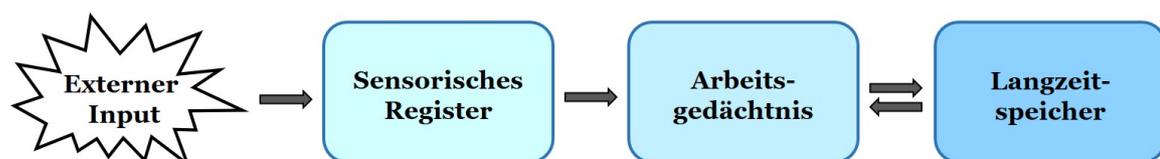


Abbildung 2.3.1. Informationsfluss durch das menschliche Gedächtnis. Vereinfachte Darstellung basierend auf dem von Atkinson und Shiffrin (1968) vorgeschlagenen Mehrspeichermodell.

Die oben dargelegte Struktur des Mehrspeichermodells liegt auch den nachfolgend erläuterten Modellen zur Verarbeitung von Texten und Bildern zugrunde.

2.3.1.1 *Cognitive Theory of Multimedia Learning*

Ab den späten 1980er Jahren entwickelten Mayer und Kolleg:innen eine evidenzbasierte Theorie zum Lernen mit Wörtern und Bildern, die sie letztendlich als *Cognitive Theory of Multimedia Learning* (CTML) bezeichneten (Mayer, 2014). In der CTML wird ähnlich wie im oben dargestellten Modell von Atkinson und Shiffrin (1968) zwischen sensorischem Gedächtnis, Arbeits- und Langzeitgedächtnis unterschieden (vgl. Abb. 2.3.2) und darüber hinaus werden weitere kognitionswissenschaftliche Annahmen berücksichtigt, wie im Folgenden erläutert wird. Nach dem Multimediaprinzip (Mayer, 2001) verläuft das Lernen mit Material, in dem Wörter und Bilder kombiniert sind, erfolgreicher als das Lernen mit Wörtern allein, wobei das Kombinieren von Wörtern und Bildern nicht willkürlich erfolgen sollte, damit es effektiv ist. Die CTML berücksichtigt die Prozesse der Informationsverarbeitung im menschlichen Gedächtnis. Grundlegend für die CTML sind drei kognitionswissenschaftliche Annahmen (Mayer, 2014):

- Verarbeitung von visuell/bildlich und auditiv/verbal präsentierten Informationen in zwei verschiedenen Kanälen des Gedächtnisses
- jeder dieser Kanäle besitzt limitierte Kapazität
- aktive Verarbeitung der Informationen nötig

Die erste Annahme ist, dass der Mensch visuell/räumlich präsentierte Informationen und auditiv/verbal präsentierte Informationen in zwei verschiedenen Kanälen des Gedächtnisses verarbeitet. Diese Annahme geht auf die *Dual Coding Theory* (DCT) von Paivio und Baddeleys Modell des Arbeitsgedächtnisses zurück: Nach der seit den 1960er Jahren entwickelten DCT werden Informationen in zwei übergeordneten Subsystemen verarbeitet, einem für verbale und einem für nonverbale Informationen (Paivio, 1991). Dabei bestehen diese Systeme jeweils aus mehreren Subsystemen, da unter anderem verschiedene Sinnesmodalitäten berücksichtigt werden (Paivio, 1986). Verbale Informationen umfassen zum Beispiel visuell oder auditiv präsentierte Wörter oder Schrift, die haptisch wahrgenommen wird. Nonverbale Informationen können zum Beispiel visuell präsentierte Ob-

jekte, Umgebungsgeräusche, das Fühlen von Objekten, Geschmacks- oder Geruchserinnerungen sein (Paivio, 1986). Besonders dominante Subsysteme sind laut Paivio das *auditory-vocal* System im übergeordneten verbalen System und das visuelle System im nonverbalen System. Alle Subsysteme können unabhängig voneinander funktionieren, sie können aber auch Subsysteme aus demselben oder dem anderen übergeordneten System (verbal oder nonverbal) aktivieren (Paivio, 1986). Verbale und nonverbale Informationen werden nach Paivio in repräsentationalen Einheiten abgebildet, den sogenannten Logogenen bzw. Imagenen. Über diese geschieht auch der Austausch zwischen den (Sub)Systemen (Paivio, 1986). Die duale, also nonverbale und verbale Codierung eines Objekts kann zu einem additiven Effekt beim späteren Erinnern führen (Paivio, 1991). Mayer (2014) verweist zudem auf das Modell des Arbeitsgedächtnisses von Baddeley. Baddeley (1992) beschrieb ein Modell des Arbeitsgedächtnisses, das auf Arbeiten mit Graham Hitch zurückgeht, wonach sich das Arbeitsgedächtnis aus drei Komponenten zusammensetzt: Einer sogenannten phonologischen Schleife für sprachbasierte Informationen, einem zweiten Untersystem zur Verarbeitung visuell-räumlicher Informationen sowie der zentralen Exekutive. Letztere koordiniert die Informationen aus beiden Untersystemen und dient der Aufmerksamkeitssteuerung (Baddeley, 1992). In der CTML werden hinsichtlich der sensorischen Modalität die zwei Kanäle in auditiv und visuell unterschieden. Aus den auditiv oder visuell präsentierten Informationen können zudem, bezugnehmend auf den *Representation-mode* (also verbal vs. nonverbal), im Arbeitsgedächtnis *verbally based* oder *pictorially based models* konstruiert werden (Mayer, 2014; Abb. 2.3.2). Mayer (2014) merkt allerdings an, dass weitere Forschung nötig ist, um die Art der Unterschiede zwischen den Kanälen genauer zu untersuchen.

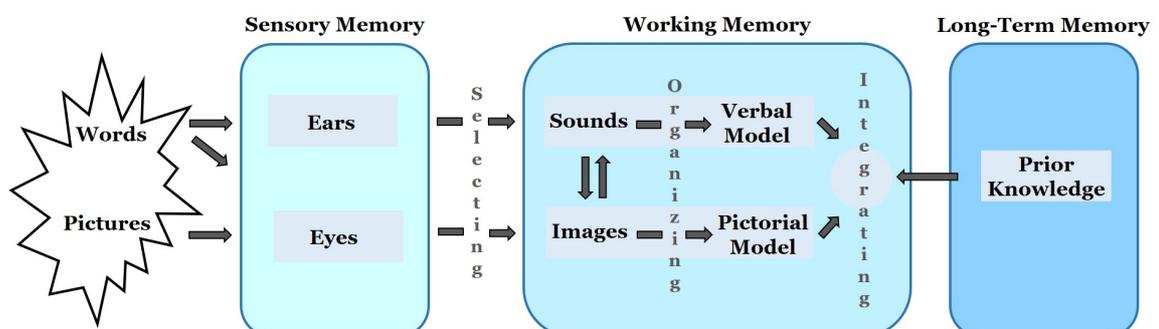


Abbildung 2.3.2. *Cognitive Theory of Multimedia Learning* verändert nach Mayer (2014, S. 52).

Im Folgenden soll die in der CTML (Mayer, 2014) dargelegte Verarbeitung von multimedial präsentierten Wörtern und Bildern genauer erläutert werden, wobei direkt auch eine weitere der CTML zugrunde liegende Annahme, die aktive Verarbeitung, die für effektives Lernen nötig ist, adressiert wird.

Die drei Hauptschritte der aktiven Verarbeitung sind die Selektion, Organisation und Integration (Mayer, 2014; Abb. 2.3.2). Zunächst werden Wörter und Bilder über die Augen oder Ohren aufgenommen und gelangen in das sensorische Gedächtnis. Gesprochene Wörter und andere Geräusche werden hier für kurze Zeit als auditives Bild gespeichert, Bilder und geschriebene Wörter als visuelles Bild. Nun erfolgt als aktiver Prozess im auditiven und visuellen Kanal jeweils die Selektion der relevanten Informationen zur Erstellung eines spezifischeren Klangbildes bzw. visuellen Bildes im Arbeitsgedächtnis. Dort können Klangbilder und visuelle Bilder ineinander umgewandelt werden. So kann aus dem Klangbild *Katze* das mentale visuelle Bild einer Katze gebildet werden, umgekehrt kann aus dem mentalen visuellen Bild einer Katze das mentale Klangbild *Katze* entstehen (Mayer, 2014). Ebenfalls im Arbeitsgedächtnis findet getrennt voneinander im jeweiligen Kanal die Organisation ausgewählter Klangbilder zu als verbale Modelle bezeichneten Wissensstrukturen statt bzw. die Organisation ausgewählter visueller Bilder zu Wissensstrukturen, die als bildliche Modelle bezeichnet werden und eine strukturierte, schematische räumliche Repräsentation ausgewählter visueller Bilder darstellen (Mayer, 2014). Es werden im jeweiligen Kanal also sinnhafte Verbindungen zwischen ausgewählten Wörtern/Phrasen (auditiver Kanal) oder verschiedenen Teilen der bildlichen Informationen (visueller Kanal) hergestellt (Mayer, 2014). Beim multimedialen Lernen werden anschließend weiterhin im Arbeitsgedächtnis Informationen aus den beiden Kanälen zusammengeführt: Verbale und bildliche Modelle werden im Schritt der Integration mit relevantem Vorwissen aus dem Langzeitgedächtnis zu einem integrierten Modell zusammengesetzt. Die beschriebenen kognitiven Prozesse laufen während des Arbeitens mit multimedialen Material wiederholt ab, nämlich mit verschiedenen Abschnitten des Materials, und konstruierte integrierte Modelle werden zur dauerhaften Speicherung dem Langzeitgedächtnis übermittelt, aus dem sie wiederum als Vorwissen für neue Lernprozesse abgerufen werden können (Mayer, 2014).

Die letzte Annahme, die der CTML zugrunde liegt, besagt, dass jeder der genannten Kanäle eine limitierte Kapazität besitzt, dass also gleichzeitig pro Kanal nur eine bestimmte Menge an Informationen verarbeitet werden kann (Mayer, 2014). Die Vorstellung einer limitierten Verarbeitungskapazität ist unter anderem sowohl in Baddeleys (1992) Modell

des Arbeitsgedächtnisses als auch in der *Cognitive Load Theory* (CLT) von Sweller und Kolleg:innen (z. B. Paas & Sweller, 2014) dargelegt. Die CLT adressiert die limitierte Kapazität des Arbeitsgedächtnisses und beschreibt drei Kategorien von *Cognitive Load* (*Intrinsic*, *Extraneous*, *Germane*; z. B. Paas & Sweller, 2014)⁶, die nachfolgend noch erläutert werden (Kapitel 2.3.1.3). Die limitierte Kapazität steht in direktem Zusammenhang mit Selektieren, Organisieren und Integrieren von Informationen. Denn eine limitierte Verarbeitungskapazität zwingt dazu, nur relevante Informationen aus dem Material zu selektieren, nur relevante Verbindungen zwischen Elementen herzustellen und nur in bestimmtem Maß Verbindungen mit dem Vorwissen herzustellen (Mayer, 2014). Während des Lernens müssen nach der CTML drei Arten kognitiver Anforderungen bewältigt werden, die bei der Gestaltung multimedialen Materials berücksichtigt werden sollten. Dies ist erstens das *Extraneous Processing*, das notwendig ist, wenn das Material schlecht gestaltet ist, und das das Erreichen des Lernziels erschwert. Diese Anforderung wird als analog zum *Extraneous Cognitive Load* der CLT betrachtet (Mayer, 2014). Des Weiteren muss ein *Essential Processing* erfolgen, das das Selektieren und Organisieren von Informationen aus dem Material beinhaltet und durch die Komplexität dessen bestimmt wird. Es gilt als analog zum *Intrinsic Cognitive Load* der CLT (Mayer, 2014). Als drittes dient ein *Generative Processing* dem Verstehen des präsentierten Materials durch das Erstellen eines integrierten mentalen Modells. Es wird durch die Motivation der Lernenden beeinflusst und wird als analog zum *Germane Cognitive Load* der CLT betrachtet (Mayer, 2014). Wie die genannten Anforderungen bewältigt werden, hängt auch von den Lernenden selbst ab. So können sich zum Beispiel die Arbeitsgedächtniskapazität und das Vorwissen unterscheiden, was sich auf die Informationsverarbeitung auswirkt (Mayer, 2014). Diese drei Verarbeitungsarten teilen sich nach Mayer die kognitive Gesamtkapazität, sodass in Hinblick auf deren Limitation das multimediale Material so gestaltet werden sollte, dass das *Essential Processing* kontrolliert und das *Generative Processing* gefördert wird, während das *Extraneous Processing* reduziert werden sollte (2014).

Letzteres kann nach Mayer und Fiorella (2014) zum Beispiel geschehen, indem irrelevantes Material entfernt wird (Kohärenzprinzip), Text in die Nähe der zugehörigen Abbildung platziert wird (Prinzip der räumlichen Kontiguität) oder auf geschriebenen Text zusätzlich zu gesprochenem Text zu einer Animation verzichtet wird (Redundanzprinzip;

⁶ In dieser Arbeit werden in diesem Fall die englischen Originalbegriffe nicht übersetzt, um inhaltliche Verzerrungen durch die Übersetzung ins Deutsche zu vermeiden. Gerade das im Deutschen negativ konnotierte Wort *Belastung* könnte hier teilweise irreführend sein, wie in Kapitel 2.3.1.3 deutlich wird.

vgl. Kapitel 2.3.4). Dem *Essential Processing* kann zum Beispiel durch die Anwendung des Modalitätsprinzips Rechnung getragen werden, das besagt, dass in multimedialem Material gesprochener Text lernförderlicher ist als geschriebener Text (Mayer, 2014). Allerdings dürfte dies im schulischen Kontext nicht immer möglich sein. In Prüfungskontexten wird zum Beispiel in aller Regel geschriebener Text (sowie Abbildungen) eingesetzt. Das *Generative Processing* kann beispielsweise gefördert werden, indem das bereits beschriebene Multimediaprinzip berücksichtigt wird (Mayer, 2014). Selbst wenn das multimediale Material so gestaltet ist, dass genug Arbeitsgedächtniskapazität für *Generative Processing* gegeben wäre, wird die Informationsverarbeitung aber nur dann erfolgreich sein, wenn die Person, die das Material bearbeitet, motiviert ist, diese Kapazität auch zu nutzen (Mayer, 2020). Das heißt, die Verarbeitungsschritte Selektion, Organisation und Integration können durch die Motivation der Lernenden beeinflusst sein (Mayer, 2014, 2020).

2.3.1.2 *Integrated Model of Text and Picture Comprehension*

Ein alternatives Modell zur Verarbeitung von Texten und Bildern, dem ebenfalls kognitionspsychologische Annahmen zugrunde liegen, ist das *Integrated Model of Text and Picture Comprehension* (ITPC-Modell) von Schnotz und Bannert (2003), das sowohl die einzelne Verarbeitung von Texten oder Bildern als auch die Verarbeitung von Texten und Bildern abbildet (Schnotz, 2014). Auch nach diesem Modell werden Texte und Bilder separat verarbeitet, Schnotz und Bannert (2003) nutzen die Bezeichnungen deskriptionaler und depiktionaler Zweig (*descriptive branch, depictive branch*). Im deskriptionalen Zweig wird aus Texten zunächst eine mentale Repräsentation der Textoberflächenstruktur und anschließend eine propositionale Repräsentation des semantischen Inhalts im Arbeitsgedächtnis gebildet (Schnotz, 2014; Schnotz & Bannert, 2003). Propositionale Repräsentationen sind deskriptionale Repräsentationen, die die Ideen des dargestellten Textes aufgreifen, ohne exakte Formulierung oder Syntax beizubehalten (Schnotz, 2005). Einhergehend mit einem Wechsel zum depiktionalen Zweig wird anhand einer propositionalen Repräsentation ein mentales Modell konstruiert oder ausgearbeitet (Schnotz, 2014; Schnotz & Bannert, 2003). Von diesem mentalen Modell können im Vorgang der Modellinspektion unter Einfluss von Vorwissen wiederum Informationen abgelesen werden und zur Ausarbeitung der propositionalen Repräsentation dienen (Schnotz & Bannert, 2003). Bei der Bildverarbeitung im depiktionalen Zweig wird zunächst eine visuelle mentale Repräsentation gebildet und anschließend wird im Arbeitsgedächtnis sowohl ein mentales Modell als auch beim

Vorgang der Modellinspektion, einhergehend mit einem Wechsel in den deskriptionalen Zweig, eine propositionale Repräsentation konstruiert (Schnotz, 2014; Schnotz & Bannert, 2003). Mentale Modelle haben eine strukturelle oder funktionale Analogie zu ihrem Bezugsobjekt, wobei irrelevante Informationen entfernt wurden, während andere Informationen durch Vorwissen ergänzt worden sein können (Schnotz & Bannert, 2003). Nach dem ITPC-Modell werden also sowohl externe deskriptionale als auch depiktionale Repräsentationen intern in beiden Zweigen verarbeitet. Die Konstruktion eines mentalen Modells verläuft wie oben beschrieben beim Bildverstehen allerdings direkter als beim Textverstehen. In einer Erweiterung des ITPC-Modells werden verschiedene Sinnesmodalitäten berücksichtigt (Schnotz, 2014). Auditiv (z. B. gesprochener Text, Klangbilder) oder visuell (z. B. geschriebener Text, visuelle Bilder) präsentierte Informationen gelangen über verschiedene sensorische Register in das Arbeitsgedächtnis. Unabhängig von der initialen Sinnesmodalität wird sowohl geschriebener als auch gesprochener Text wie für das Modell aus 2003 beschrieben zunächst im deskriptionalen und anschließend im depiktionalen Zweig verarbeitet, visuell präsentierte Bilder und Klangbilder zunächst im depiktionalen und danach im deskriptionalen Zweig (Schnotz, 2014). Auch Schnotz geht von einer Limitation der Kapazität des Arbeitsgedächtnisses aus.

2.3.1.3 Cognitive Load Theory

Die bereits angesprochene CLT befasst sich damit, wie mit Blick auf Limitationen des Arbeitsgedächtnisses kognitive Ressourcen beim Lernen und Problemlösen eingesetzt werden (z. B. Chandler & Sweller, 1991; Sweller et al., 2011). Nach Sweller wird der Großteil der kognitiven Aktivität eines Menschen durch Informationen gelenkt, die im Langzeitgedächtnis gespeichert sind. Diese Informationen gelangen durch den Prozess des Lernens in das Langzeitgedächtnis, weswegen Lernen als Änderung des Langzeitgedächtnisses angesehen wird (Sweller, 2005a). Lernmaterial sollte deshalb so gestaltet sein, dass es eine angemessene Änderung im Langzeitgedächtnis bewirkt (Sweller, 2005a). Personen, die viele Informationen zu einem bestimmten Thema im Langzeitgedächtnis gespeichert haben, können als Expert:innen in diesem Thema betrachtet werden. Informationen werden im Langzeitgedächtnis in Form von Schemata gespeichert (Sweller, 2005a). Dies sind kognitive Konstrukte, die eine große Menge an Informationen zu einer Einheit zusammengefügt enthalten können (Kirschner, 2002; Sweller, 2005a). Expert:innen auf einem Gebiet

verfügen also über eine hohe Anzahl entsprechender Schemata (Cook, 2006). Die Schemakonstruktion erfolgt meist bewusst und umfasst auch die Konstruktion von komplexeren Schemata aus weniger komplexen Schemata (van Merriënboer & Kester, 2005). Wird ein Schema wiederholt erfolgreich angewendet und kann fortan unbewusst genutzt werden, wird von Schemaautomatisierung gesprochen, die – wie im Folgenden deutlich wird – wie die Schemakonstruktion die Auslastung des Arbeitsgedächtnisses reduzieren kann (van Merriënboer & Kester, 2005). Als Beispiel führt Sweller (2005a) das Lesen an, das nach dem Erlernen von Buchstaben und ihrer Kombination zu Wörtern ohne eine bewusste Beachtung einzelner Buchstaben oder Wörter erfolgen kann. Schemata können sowohl visuell als auch auditiv sein (Sweller, 2005a).

Das Arbeitsgedächtnis ist anders als das Langzeitgedächtnis in der Anzahl der gleichzeitig speicherbaren Informationen und der Dauer der Speicherung stark limitiert. Auch wenn die genaue Anzahl zur Diskussion steht (nach bisherigen Befunden weniger als 10), steht fest, dass im Arbeitsgedächtnis zeitgleich sehr wenige Elemente gespeichert oder verarbeitet werden können, wenn es mit neuen Informationen konfrontiert wird (Sweller et al., 2011). Zudem gehen Informationen, die nicht wiederholt werden, nach einigen Sekunden verloren (Sweller et al., 2011). Als Alltagsbeispiel führen Sweller et al. (2011) das Merken einer Telefonnummer an. Wird einer Person eine neue Telefonnummer gesagt, muss sie diese bis zum Wählen wiederholen, da sie die Nummer sonst nach wenigen Sekunden vergisst. Wiederholen kann zudem die Informationsübertragung in das Langzeitgedächtnis unterstützen (Sweller et al., 2011).

Den Limitationen des Arbeitsgedächtnisses wird durch seine Verbindungen zum Langzeitgedächtnis entgegengewirkt: Während neue Informationen, die über das sensorische Gedächtnis in das Arbeitsgedächtnis eintreten, von den genannten Limitationen in Kapazität und Dauer betroffen sind, ist dies für bereits vorhandene Informationen, die als Schemata aus dem Langzeitgedächtnis abgerufen werden, nicht der Fall (Sweller, 2005a). Schemata können wie beschrieben große Informationsmengen bündeln und werden als einzelnes Element im Arbeitsgedächtnis genutzt (van Merriënboer & Kester, 2005). Sweller (2005a) beschreibt ein Kontinuum, wonach bei Konfrontation mit gänzlich neuen Informationen die Limitationen des Arbeitsgedächtnisses stark ins Gewicht fallen, bei zunehmender Vertrautheit und dem Nutzen von Informationen aus dem Langzeitgedächtnis die Limitationen weniger kritisch sind bis hin zur Irrelevanz der Limitationen, wenn mit gänzlich vertrauten Informationen aus automatisierten Schemata gearbeitet wird. Ein Sachverhalt

kann nur dann verstanden werden, wenn alle relevanten Informationen gleichzeitig im Arbeitsgedächtnis verarbeitet werden können (Sweller, 2005a). Aufgrund der Limitationen müssen bei umfangreichem Material mit vielen neuen Informationen zunächst einzelne Elemente zu Schemata organisiert und kombiniert und im Langzeitgedächtnis gespeichert werden, aus dem sie für den Verständnisprozess abgerufen werden können (Sweller, 2005a). Im Langzeitgedächtnis gespeicherte Schemata nehmen die Rolle einer zentralen Exekutive ein, die die Informationsverarbeitung im Arbeitsgedächtnis steuert (Sweller, 2005a). Das Fehlen von Schemata zu bestimmten Sachverhalten – wie es bei Schüler:innen häufig zu erwarten ist – kann durch ein bedachtes Design multimedialen Materials ausgeglichen werden. Dabei sollte unter anderem den bereits angesprochenen Kategorien von *Cognitive Load* Rechnung getragen werden (Sweller, 2005a):

Der *Intrinsic Cognitive Load* wird durch die Komplexität des zu erfassenden Inhalts bestimmt, die nach Sweller wiederum auf die sogenannte Element-Interaktivität zurückgeht. Wenn viele einzelne Elemente der dargestellten Information interagieren und gemeinsam verarbeitet werden müssen, die Element-Interaktivität also hoch ist, ist der *Intrinsic Cognitive Load* hoch und das Material kann als komplex angesehen werden (Sweller, 2005a). Bei Personen mit hohem Vorwissen, die mehrere Elemente des Materials zu einem kombinieren können, kann der *Intrinsic Cognitive Load* erniedrigt sein (Paas & Sweller, 2014).

Der *Extraneous Cognitive Load* geht aus einem ungeeigneten Design des Materials hervor. Zum Beispiel können überflüssige Informationen oder eine unzulängliche Strukturierung des Materials Kapazitäten des Arbeitsgedächtnisses beanspruchen, die dann nicht mehr für die effektive Verarbeitung der Informationen, also zur Konstruktion von Schemata und deren Automatisierung zur Verfügung stehen (Sweller, 2005a). Der *Extraneous Cognitive Load* sollte deshalb reduziert werden. Dies kann zum Beispiel durch Berücksichtigung des Redundanzeffekts geschehen, der in Kapitel 2.3.4 erläutert wird. Wenn multimediales Material so gestaltet wird, dass es Noviz:innen bei der Informationsverarbeitung besser unterstützt als eine alternative Gestaltung, mit zunehmender Expertise aber den Vorteil verliert oder sogar nachteilig ist, liegt der *Expertise Reversal Effect* vor (Kalyuga et al., 2003). Dieser kann nach Kalyuga und Sweller (2014) beispielsweise so erklärt werden, dass etwa ein erklärender Text zu einer Abbildung für Noviz:innen essentiell sein könnte und deshalb nah an der Abbildung platziert werden sollte, während dieser Text für Expert:innen überflüssig ist. Für Expert:innen wäre eine Trennung von Abbildung und Text

besser, da sie den Text nicht benötigen und er den *Extraneous Load* erhöht (Kalyuga & Sweller, 2014).

Den ***Germane Cognitive Load*** bezeichnet Sweller als effektiv, da er zur Konstruktion und Automatisierung von Schemata führt (2005a). Der *Germane Cognitive Load* bezieht sich auf die Ressourcen des Arbeitsgedächtnisses, die, statt für den Umgang mit *Extraneous Cognitive Load* benötigt zu werden, dem Umgang mit *Intrinsic Cognitive Load* gewidmet sind (Paas & Sweller, 2014). Der *Extraneous Cognitive Load* und *Intrinsic Cognitive Load* gelten als additiv (Paas & Sweller, 2014)⁷. Wird Arbeitsgedächtniskapazität durch Reduktion von *Extraneous Cognitive Load* frei, könnten mehr Ressourcen für den Umgang mit *Intrinsic Cognitive Load* bereitstehen, das heißt, der *Germane Cognitive Load* wird erhöht (Paas & Sweller, 2014). Bei Gestaltung multimedialen Materials sollte daher versucht werden, den *Extraneous Load* zu reduzieren (Paas & Sweller, 2014). Ist der *Intrinsic Cognitive Load* allerdings gering, dürften geringe Ressourcen des Arbeitsgedächtnisses für einen Umgang mit diesem *Load* ausreichen, also ein geringer *Germane Cognitive Load*, das heißt, die Informationsverarbeitung kann bei geringem *Intrinsic Cognitive Load* vermutlich auch dann erfolgreich sein, wenn der *Extraneous Cognitive Load* hoch sein sollte (Paas & Sweller, 2014).

In Kapitel 2.3.2 werden einige, auch das Lehren und Lernen von Biologie betreffende empirische Studien zum Einsatz von Text-Bild-Kombinationen (d. h. MER) angeführt, deren Ergebnisse mit den hier erläuterten Modellen der kognitiven Verarbeitung von Texten und Bildern erklärt werden können. Zunächst wird aber ein Rahmenwerk von Ainsworth (2006) vorgestellt, das auch pädagogische Funktionen von MER berücksichtigt.

2.3.2 *DeFT (Design, Functions, Tasks) Framework* nach Ainsworth (2006)

An die im vorangegangenen Unterkapitel behandelten Erkenntnisse zur kognitiven Verarbeitung von Text-Bild-Kombinationen und damit MER sowie die Prinzipien zur Gestaltung von MER soll angeschlossen werden, indem das *DeFT Framework* (*DeFT: De-*

⁷ In vielen Ausführungen werden *Extraneous*, *Intrinsic* und *Germane Cognitive Load* als additiv dargestellt. Kalyuga (2011) hält hingegen *Intrinsic* und *Germane Cognitive Load* nicht für trennbar und argumentiert dafür, wie hier auch nach Paas und Sweller (2014) beschrieben den *Germane Cognitive Load* als die Ressourcen des Arbeitsgedächtnisses zu begreifen, die dem Umgang mit *Intrinsic Cognitive Load* statt mit *Extraneous Cognitive Load* gewidmet sind.

sign, Functions, Tasks) von Ainsworth (2006) erläutert wird. Das *DeFT Framework* integriert „research on learning, the cognitive science of representation and constructivist theories of education“ (Ainsworth, 2006, S. 183), bezieht also teilweise Aspekte der bereits vorgestellten kognitiven Verarbeitung von MER mit ein. Darüber hinaus wird besonderes Augenmerk auf die Funktionen gelegt, die Kombinationen von mehreren Repräsentationen erfüllen können bzw. sollen (Ainsworth, 2006). Das *DeFT Framework* benennt drei grundlegende Aspekte des Lernens mit MER: *Design Parameters*, pädagogische Funktionen von MER (bereits in Ainsworth, 1999 beschrieben) und kognitive Aufgaben, die bei der Interaktion mit MER bewältigt werden müssen (Ainsworth, 2006). Diese Aspekte und ihre Unterpunkte sind in Abbildung 2.3.3 übersichtswise dargestellt und werden nachfolgend erläutert.

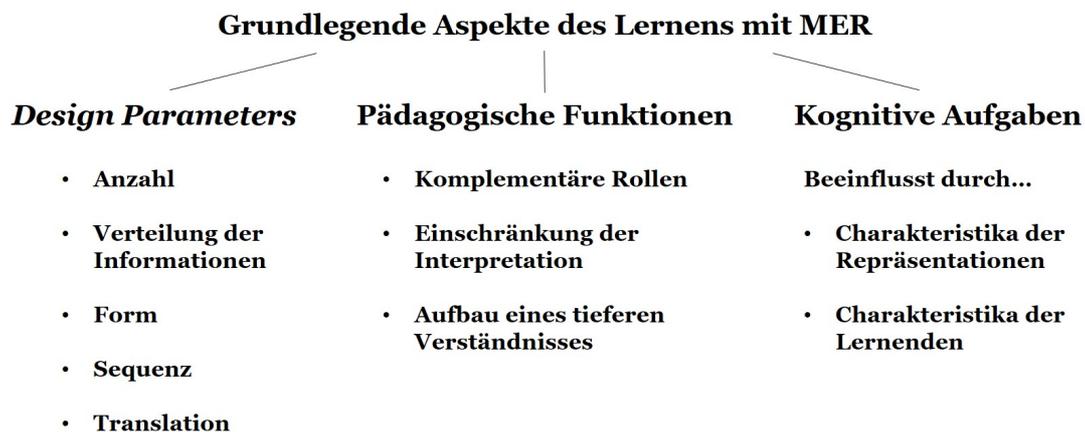


Abbildung 2.3.3. Grundlegende Aspekte des Lernens mit MER, die das *DeFT (Design, Functions, Tasks) Framework* von Ainsworth (2006) umfasst. Die pädagogischen Funktionen wurden bereits in Ainsworth (1999) vorgestellt.

Die folgende Erläuterung der drei genannten grundlegenden Aspekte und ihrer Unterpunkte beziehen sich, wenn nicht anders gekennzeichnet, auf die Erläuterung von Ainsworth (2006):

Design. Bei der Gestaltung von MER sind bestimmte Aspekte zu beachten, die die Effektivität für das Lernen beeinflussen können. Die Anzahl von Repräsentationen einer MER kann variieren und sollte der Funktion gerecht werden, die die MER erfüllt. Die Verteilung von Informationen auf die einzelnen Repräsentationen einer MER kann so gestaltet sein, dass einzelne Repräsentationen komplett unterschiedliche Informationen enthalten,

also nicht redundant sind, was die Komplexität der Repräsentationen mindert, aber eine hohe Integrationsleistung erfordert. Die Informationen können sich teilweise überschneiden (partielle Redundanz) oder komplett identisch bzw. redundant sein, so dass sich nur die Art der Repräsentation unterscheidet. Redundanz in MER ist ein Hauptaspekt der vorliegenden Arbeit und wird ausführlich in Kapitel 2.3.4 behandelt. Repräsentationen einer MER können sich in ihrer Form bzw. Art unterscheiden und zwar bezogen auf verschiedene Zeichensysteme (z. B. Text oder Bild), unterschiedliche angesprochene Sinnesmodalitäten (z. B. visuell oder auditiv) oder auch die Dynamik und Dimensionalität (vgl. auch Kapitel 2.2). Dabei sind in einer MER Auswirkungen der Interaktion verschiedener Repräsentationsformen zu beachten. Die Sequenz von einzelnen Repräsentationen einer MER ist ein weiterer Aspekt, der beim Design von MER eine Rolle spielt. Die Repräsentationen können simultan oder nacheinander präsentiert werden. Ein weiterer Aspekt des Designs von MER betrifft die Translation⁸ und zwar das Ausmaß oder die Art (eher oberflächlich oder tiefergehend), wie Lernende aktiv darin unterstützt werden, Zusammenhänge zwischen Repräsentationen einer MER herzustellen. Dies bezieht Ainsworth (2006) besonders auf computergestützte MER-Umgebungen.

Pädagogische Funktionen. Das Design von MER sollte je nach Ziel an verschiedene Funktionen angepasst sein, die MER erfüllen können. Diese Funktionen wurden bereits in Ainsworth (1999) beschrieben und von Ainsworth (2006) in das DeFT-Rahmenwerk integriert. Repräsentationen in MER können komplementäre Funktion besitzen mit dem Ziel, dass Vorteile der einzelnen Repräsentationen den Lernenden helfen. Diese komplementäre Funktion kann sich auf Prozesse oder Informationen beziehen. Wenn verschiedene Repräsentationen (z. B. Diagramm vs. Text) einer MER die gleichen Informationen enthalten, kann dies die Informationen Lernenden mit verschiedenen Präferenzen bezüglich Repräsentationsarten zugänglich machen oder Prozesse unterschiedlicher Strategienutzung anstoßen, die wiederum zusammen zu einer erfolgreicherer Bearbeitung führen können. In Bezug auf Informationen ergänzen sich nicht oder partiell redundante Repräsentationen, was hilfreich sein kann, wenn eine Repräsentation alleine zu komplex wäre oder für die Darstellung aller Informationen ungeeignet wäre (Ainsworth, 1999). Als Beispiel für eine

⁸ Ainsworth äußert sich zum Begriff *Translation* in ihrem früheren Artikel wie folgt: „Translation throughout this paper is used to refer to **all** cases when a learner must see the relation between two representations. It is used to refer both to the cases when a learner must comprehend the relation between two representations and also when they must act to reproduce this relation. It is neutral about whether translation occurs through direct mapping between the symbols or whether it is mediated through domain understanding” (1999, S.132).

MER mit komplementären Informationen nennt Ainsworth (1999) die Kombination aus einer Karte eines U-Bahn-Netzes und einer Straßenkarte. Um den schnellsten Weg zwischen zwei U-Bahn-Stationen zu ermitteln, reicht es nicht zwingend, die Karte des U-Bahn-Netzes zu betrachten, wenn sie lediglich die Haltestellen, aber nicht die genaue Entfernung zwischen selbigen angibt. Werden Informationen der U-Bahn-Netz-Karte mit Informationen zu tatsächlichen Entfernungen aus der Straßenkarte integriert, kann der schnellste Weg ermittelt werden, was möglicherweise auch bedeutet, dass Strecken zu Fuß zurückgelegt werden sollten, wenn dort die U-Bahn-Strecke viele zeitintensivere Umstiege beinhalten würde. Durch das Aufteilen der Informationen auf zwei Karten ist jede einzelne weniger komplex (Ainsworth, 1999).

Eine einschränkende Funktion von MER liegt dann vor, wenn eine Repräsentation die Interpretation einer anderen einschränkt. Wenn Lernende zum Beispiel mit einer Repräsentationsart der MER vertraut sind, kann dies bei der Interpretation einer weniger vertrauten Repräsentationsart helfen. Oder die Eigenschaften einer Repräsentationsart (z. B. Spezifität) helfen bei der Interpretation einer weiteren Repräsentation. Krey und Schwaneveld (2018) führen als Beispiel im Bereich der Biologie an, dass ein Text, der Spaltöffnungen von Pflanzen und deren Funktion beschreibt, durch eine Abbildung von Spaltöffnungen an einer bestimmten Pflanze ergänzt werden könnte, die also das Aussehen und die Position von Spaltöffnungen spezifiziert. In Bezug auf die einschränkende Funktion von MER führt Ainsworth (2014) an, dass mit dem *Expertise Reversal Effect* (Kapitel 2.3.1.3) einhergehend die Einschränkung der Interpretation einer Repräsentation durch eine andere für Noviz:innen vorteilhaft sein kann, während sie für Expert:innen, die mit beiden Repräsentationsarten vertraut sind, ihren Nutzen verlieren kann.

Das Integrieren von Informationen aus verschiedenen Repräsentationen einer MER kann den Aufbau eines tieferen Verständnisses bewirken, als wenn die Information in einer einzelnen Repräsentation dargeboten würde. Dies kann durch den Vorgang der Abstraktion erfolgen, wenn Bezüge zwischen verschiedenen Repräsentationen einer MER hergestellt werden und so die der repräsentierten Domäne zugrunde liegende Struktur erkannt wird. Auch kann es durch Erweiterung erfolgen. Erweiterung bedeutet hier, dass Lernende bestehendes Fachwissen von einer bekannten Repräsentation auf eine andere übertragen, ohne dass sich die Art dieses Wissens bedeutsam verändert. Bei Einsatz von MER kann zudem durch Lehrende adressiert werden, wie einzelne Repräsentationen zueinander in Beziehung stehen, was wiederum auch bei der Abstraktion helfen könnte. Eine MER kann

eine der genannten Funktionen, aber auch eine Kombination verschiedener Funktionen erfüllen.

Kognitive Aufgaben. Damit MER die genannten Funktionen erfüllen können, müssen Lernende in der Lage sein, die mit dem Gebrauch von MER einhergehenden kognitiven Aufgaben zu bewältigen. Gerade das Herstellen von Verbindungen zwischen einzelnen Repräsentationen und die Integration der Informationen gelingen nicht immer. Um Faktoren zu identifizieren, die die Schwierigkeit des Herstellens von Verbindungen zwischen Repräsentationen einer MER beeinflussen, werden Charakteristika von Repräsentationen und von Lernenden betrachtet. Je mehr sich einzelne Repräsentationen unterscheiden, desto schwieriger kann es für Lernende sein, diese Repräsentationen in Beziehung zu setzen. Unterschiede können zum Beispiel darin bestehen, ob die Repräsentation die Informationen auditiv oder visuell präsentiert, textlich oder bildlich, und wenn textlich und bildlich, inwiefern Text und Bild physisch integriert präsentiert werden, ob Repräsentationen statisch oder dynamisch sind, welches Abstraktionslevel und welche Dimensionalität einzelne Repräsentationen besitzen. Auch Charakteristika der Lernenden selbst können beeinflussen, wie gut Repräsentationen einer MER in Beziehung gesetzt werden können. Wenn sie zum Beispiel mit einzelnen Repräsentationen bereits vertraut sind, sollten mehr kognitive Ressourcen für das Herstellen von Verbindungen zwischen Repräsentationen verbleiben. Auch die positiven Auswirkungen von fachlichem Vorwissen auf das Verstehen von tieferen Verbindungen zwischen Repräsentationen führt Ainsworth (2006) an. Das von ihr vorgeschlagene Rahmenmodell soll bei weiterer Forschung zum Lernen mit MER zur Orientierung dienen (Ainsworth, 2006).

Auch wenn Ainsworth sich meist auf das Lernen mit MER bezieht, betreffen oben genannte Aspekte generell MER-Material und die Bearbeitung dessen. Die im Rahmenmodell genannten Design-Aspekte gelten also sowohl für Lehr- und Lernmaterial als auch für Testmaterial. Ebenso können MER die genannten Funktionen sowohl in Lehr- und Lernmaterial als auch in Testmaterial erfüllen. Die kognitiven Aufgaben, die sich durch Charakteristika der Repräsentationen und der bearbeitenden Person ergeben, müssen sowohl beim Lernen mit MER als auch bei der Bearbeitung von MER in einem Test bewältigt werden. Lehrende können somit auch über die Wahl von Repräsentationen und ihrer Kombination die Schwierigkeit von Testmaterial steuern, wofür sie sich potentieller Effekte und Funktionen bestimmter Kombinationen von Repräsentationen bewusst sein sollten. Letzteres gilt auch für die Gestaltung von Lehr-Lernmaterial in naturwissenschaftlichen Fächern.

Im Verlauf dieser Arbeit wird diesbezüglich deutlich, dass das Ziel der MER-Gestaltung nicht primär in einer möglichst vereinfachten Darstellung von Fachinhalten liegen, sondern bei der Gestaltung die Kommunikation in der Fachdisziplin berücksichtigt werden sollte. Das Wissen über Potentiale und Schwierigkeiten beim Einsatz von MER kann Lehrenden helfen, Lernende beim Arbeiten mit verschiedenen MER zu unterstützen. Im folgenden Kapitel werden empirische Befunde zu Potentialen, Schwierigkeiten und Integrationshilfen bezüglich MER zu naturwissenschaftlichen Inhalten vorgestellt.

2.3.3 Befunde zu Potentialen, Schwierigkeiten und Integrationshilfen bezüglich MER-Einsatz zu naturwissenschaftlichen Inhalten

Im Folgenden werden Studien vorgestellt, die mit Bezug zu den in den vorangegangenen Kapiteln erläuterten Theorien und Modellen Potentiale, Schwierigkeiten und Integrationshilfen bezüglich MER im naturwissenschaftlichen Bereich untersucht haben. Dabei variieren die Altersgruppen der Studienteilnehmenden von Grundschüler:innen bis zu Studierenden. Da in dem in der vorliegenden Dissertation dargelegten Promotionsprojekt die Testleistung beim Arbeiten mit vorgegebenen MER und nicht die Konstruktion von MER untersucht wird, wird auch in diesem Unterkapitel der Fokus auf die Informationsentnahme aus beziehungsweise Interpretation von MER gelegt.

Positive Effekte von MER in Naturwissenschaften verglichen mit Text zeigten beispielsweise Mason et al. (2013). Sie setzten Material in drei Variationen ein: Eine Gruppe von Elftklässler:innen bekam einen Text, zwei weitere Gruppen arbeiteten mit Text-Bild-Material zur schiefen Ebene, wobei in einem MER-Material eine konkrete, im anderen eine abstrakte Abbildung enthalten war. In Einklang mit dem Multimediateprinzip (Mayer, 2001; vgl. Kapitel 2.3.1.1) erzielten Schüler:innen in den MER-Gruppen ein höheres Faktenwissen nach Lesen des Materials als die Text-Gruppe. Zudem zeigten sie direkt sowie bei einem Posttest 1,5 Monate später eine bessere Transferleistung (Mason et al., 2013). Eilam und Poyas (2008) entwickelten Material zu Mobiltelefonen. Studierende einer Experimentalgruppe lösten Aufgaben mit Hilfe von Informationen, die in Form von Texten oder verschiedenen Abbildungen dargeboten wurden (MER-Bedingung), eine weitere Gruppe bekam die Informationen rein textlich dargeboten. Die MER-Gruppe erzielte bessere Leistungen bei der Bearbeitung der Aufgaben und zeigte auch – ganz im Sinne der Dual Coding Theory – in einem Posttest bessere Behaltensleistungen sowie bessere Transferleistungen (Eilam & Poyas, 2008). Herrlinger et al. (2017) ließen Grundschüler:innen mit Material zu

Herz und Blutkreislauf arbeiten. In einem anschließenden Multiple-Choice-Test zeigten Schüler:innen einer Bedingung mit geschriebenem Text und Bildern zwar insgesamt keine bessere Leistung als Schüler:innen einer Bedingung mit geschriebenem Text. Eine Gruppe, der Bilder mit gesprochenem Text präsentiert wurden, schnitt aber besser ab als eine Gruppe, die nur gesprochenen Text hörte, und auch als die Gruppe mit Bildern und geschriebenem Text, was mit dem Modalitätseffekt begründet wurde (vgl. Kapitel 2.3.1.1).

Teilweise zeigte Forschung zu MER, dass bildliche Repräsentationen in MER weniger beachtet werden als Text. In einer Studie mit Grundschüler:innen, die mit Material eines Biologieschulbuchs und einer Version, in der die Abbildungen aus dem Material entfernt worden waren, arbeiteten, stellten Hannus und Hyönä (1999) zwar einen positiven Effekt von Abbildungen im Material auf die Testleistung bezüglich Fragen zu illustrierten Textinformationen fest. In einem zweiten Experiment mit Biologieschulbuch-Abschnitten zeigten sie mittels Eye-Tracking aber, dass die Schüler:innen deutlich länger Texte fixierten als Abbildungen (Hannus & Hyönä, 1999) und somit möglicherweise weiteres gewinnbringendes Potential der Abbildungen nicht ausnutzten. Schüler:innen mit höheren intellektuellen Fähigkeiten, die durch einen nonverbalen Intelligenztest erfasst wurden, fokussierten die relevanten Segmente in Texten und Abbildungen länger und zeigten eine höhere Anzahl an Blickwechseln zwischen zusammengehörigen Segmenten beider Repräsentationsarten als Schüler:innen mit niedrigeren intellektuellen Fähigkeiten. Sie erreichten zudem eine höhere Testleistung (Hannus & Hyönä, 1999). Cromley et al. (2010) führten eine Studie mit Studierenden in einem Biologiekurs durch, die unter Anwendung der Thinking-Aloud-Methode eine Passage aus ihrem Biologie-Lehrbuch lasen. Dabei wurde unter anderem festgestellt, dass die Studierenden selten mehr als einen Textabsatz übersprungen haben, Abbildungen hingegen häufig. Ziepprecht (2016) setzte im Rahmen der Entwicklung eines Kompetenzstrukturmodells einen Kompetenztest zum Erschließen von biologischen Informationen aus Texten, Bildern und Bild-Text-Kombinationen ein. Der Test wurde von Schüler:innen der Sekundarstufe I bearbeitet. Die Aufgaben enthielten Text-Bild-Kombinationen zu biologischen Themen mit komplementär dargebotenen Informationen (vgl. Kapitel 2.3.2). Es zeigte sich eine zweidimensionale Kompetenzstruktur. Dabei war es für die Schüler:innen leichter, Informationen aus Texten zu erschließen als aus Bildern und Bild-Text-Kombinationen (Ziepprecht, 2016). Dies könnte zum einen auf Schwierigkeiten mit fachspezifischen Abbildungen, zum anderen auf Schwierigkeiten bei der Integration von Text und Bild zurückzuführen sein (Ziepprecht, 2016).

Seufert (2003) nahm den Integrationsprozess beim Lernen mit MER genauer in den Blick. Damit die Konstruktion einer mentalen Repräsentation gelingen kann, die aus verschiedenen – sich beispielsweise ergänzenden oder sich in der Interpretation gegenseitig einschränkenden (vgl. Kapitel 2.3.2) – Repräsentationen aufgebaut wird, müssen zwei Formen der Kohärenzbildung erfolgen: In den einzelnen Repräsentationen müssen relevante Elemente erkannt und in Beziehung gesetzt werden (lokale Kohärenzbildung), anschließend müssen relevante Elemente der Repräsentationen der MER in Beziehung gesetzt (globale Kohärenzbildung; Seufert, 2003). Auch mit Blick auf eine potentielle Überlastung des Arbeitsgedächtnisses besonders von Noviz:innen beim Arbeiten mit MER untersuchte Seufert (2003) – wenn auch mit eher kleiner Stichprobengröße von $N = 86$ – mögliche Hilfen für die Kohärenzbildung. Es zeigte sich, dass Studierende mit mittlerem Vorwissensniveau sowohl von direktiven Hilfen (explizite Hinweise auf relevanten Inhalt, der allen ER der MER zugrunde liegt) als auch nicht-direktiven Hilfen (Aufforderung, nach Gemeinsamkeiten und Unterschieden in ER einer MER zu suchen) profitierten, auf Personen mit geringem oder hohem Vorwissen traf das nicht zu. Corradi et al. (2012) stellten zunächst fest, dass zu Texten hinzugefügte bildliche Repräsentationen zur Darstellung chemischer Konzepte und Prinzipien bei Studierenden mit geringem Vorwissen nicht zu einer besseren Lernleistung führten. Sie beobachteten aber, dass Prompts (hier in Form von Fragen zum Material, deren Beantwortung die Integration von Text und bildlichen Repräsentationen erforderte) für die Gruppe mit geringem Vorwissen die Lernleistung steigerten, was für Gruppen mit mittlerem und hohem Vorwissen nicht zutraf (Corradi et al., 2012). Die Kohärenzbildung bezüglich MER sollte, gerade aufgrund bekannter Schwierigkeiten, im naturwissenschaftlichen Unterricht also explizit berücksichtigt werden.

Ein Designaspekt von MER ist wie bereits beschrieben der Grad der Redundanz von Informationen einzelner Repräsentationen der MER (Ainsworth, 2006). Auch wenn dieser Aspekt notwendigerweise bereits bei den in diesem Unterkapitel vorgestellten Studien berührt wurde, soll aufgrund der zentralen Rolle in dieser Dissertation im Folgenden der Aspekt Redundanz in MER separat beleuchtet werden. Dabei werden vorrangig Forschungsarbeiten angeführt, die gezielt Effekte von Redundanz untersucht haben.

2.3.4 Redundanz in multiplen externen Repräsentationen

Wie in vorstehenden Unterkapiteln beschrieben ist eine wichtige Eigenschaft von MER der Grad der Redundanz der dargestellten Informationen, also inwieweit sich Informationen in den einzelnen Repräsentationen einer MER überschneiden. Dieses Merkmal von MER ist auch aus biologische- bzw. naturwissenschaftsdidaktischer Perspektive bedeutsam. Schüler:innen werden sowohl im schulischen Fachunterricht als auch darüber hinaus – im gesellschaftlichen Leben und in einem potentiellen naturwissenschaftlichen Studium mit anschließender Berufstätigkeit in diesem Feld – mit MER in einer Vielzahl von Medien konfrontiert. Dazu zählen Schulbücher, Arbeitsblätter, Lehrbücher für Studierende, von Dozierenden angefertigte Kursmaterialien, populärwissenschaftliche oder wissenschaftliche Artikel. Die Art der in diesen Medien eingesetzten depiktionalen Repräsentationen, aber auch, wie diese mit Text kombiniert werden – dies betrifft auch den Grad der Redundanz von Informationen in einzelnen ER einer MER – variiert dabei (z. B. Roth et al., 1999). So stellten beispielsweise Roth et al. (1999) bei der Analyse von Graphen in auf Ökologie bezogenen Kapiteln aus Highschool-Schulbüchern sowie wissenschaftlichen Fachartikeln aus Fachzeitschriften aus dem Bereich der Ökologie Unterschiede bezüglich in Haupttext oder Abbildungsunterschrift wiederholten Informationen aus den Graphen fest. In den wissenschaftlichen Fachartikeln war die Redundanz von Informationen zwischen Graphen und Texten insgesamt höher, was mit einer Einschränkung des Interpretationsspielraums in Bezug auf die Aussage der Graphen einherging (Roth et al., 1999). Bowen und Roth (2002) nahmen zusätzlich depiktionalen Repräsentationen und ihre Einbettung in den Text in Ökologie-Lehrbüchern für Studierende in den Blick. Dabei stellten sie sowohl Unterschiede zu den Highschool-Schulbüchern als auch zu den wissenschaftlichen Fachartikeln fest (Bowen & Roth, 2002). Gerade auch bezüglich Lehr-Lernmaterialien, die von Dozierenden gestaltet werden, ist von einer Variation von Redundanz in MER auszugehen, die vermutlich nicht immer bewusst erfolgt. Aufgrund der Vielzahl und Unterschiedlichkeit verwendeter Medien und ihrer auch bezüglich Redundanz verschiedenartigen MER, sollten Schüler:innen im naturwissenschaftlichen Fachunterricht lernen, MER, die sich im Grad ihrer Redundanz unterscheiden, nutzen zu können. Dies ist, wie mit Blick auf die genannten Medien deutlich wird, nicht nur für Anforderungen im schulischen Kontext, sondern auch hinsichtlich des Ziels einer naturwissenschaftlichen Grundbildung und wissenschaftspropädeutischer Bildung relevant. Damit im Unterricht zum einen bei der Gestaltung von MER das Merkmal Redundanz bewusst berücksichtigt werden kann, zum anderen Lehrende auch

für möglicherweise notwendige Unterstützungsmaßnahmen für Schüler:innen sensibilisiert werden, die sich durch Schwierigkeiten mit (partiell) redundanten oder nicht-redundanten MER ergeben können, sollten potentielle Effekte des Einsatzes von MER, die sich in ihrer Redundanz unterscheiden, bekannt sein. Im Folgenden werden zunächst Funktionen verschieden redundanter MER erläutert sowie Redundanz aus kognitionspsychologischer Sicht in den Blick genommen, bevor empirische Befunde zu Effekten von Redundanz in vorrangig naturwissenschaftlichen Kontexten vorgestellt werden.

Ainsworth (1999) unterscheidet zwischen nicht-redundanten MER, in denen die einzelnen Repräsentationen einer MER unterschiedliche Informationen enthalten, und partiell redundanten MER, in denen sich die Informationen in einzelnen Repräsentationen zum Teil überschneiden, wobei jede Repräsentation zusätzlich noch weitere Informationen enthält (vgl. Kapitel 2.3.2). Wie in Kapitel 2.3.2 anhand des anschaulichen U-Bahn-Beispiels von Ainsworth geschildert, kann der Einsatz nicht-redundanter oder – wie im Beispiel – partiell redundanter MER zum Beispiel hilfreich sein, wenn der darzustellende Inhalt nicht in einer einzelnen Repräsentation wiedergegeben werden kann oder diese durch Einfügen aller relevanten Informationen zu kompliziert würde (Ainsworth, 1999). Partielle Redundanz kann auch, wie beim Nebeneinanderlegen der U-Bahn-Netz-Karte und der Straßenkarte, zu einer neuen Interpretation des Dargestellten anregen (Ainsworth, 1999). Zudem können auch alle Repräsentationen einer MER die gleichen Informationen enthalten (komplette Redundanz), sodass also nur die Darstellungsart (z. B. Schemazeichnung und Text) variiert (Ainsworth, 2006; Kapitel 2.3.2). (Partiell) redundante MER können auch eine einschränkende Funktion einnehmen oder zum Aufbau eines tieferen Verständnisses dienen (Ainsworth, 2006; vgl. Kapitel 2.3.2). Mit der Berücksichtigung der Funktionen von MER geht Ainsworth über die kognitionspsychologische Perspektive hinaus, die bei der MER-Gestaltung primär die Limitation der Arbeitsgedächtniskapazität und damit den *Cognitive Load* in den Blick nimmt, der durch Redundanz erhöht werden kann (vgl. Kapitel 2.3.1).

Mayers Multimediaprinzip, nach dem das Lernen mit Kombinationen aus Text und Bild dem Lernen mit Text allein überlegen ist (Mayer, 2001), wurde bereits in Kapitel 2.3.1.1 erläutert. Die CTML nimmt aber auch die limitierte Kapazität des Arbeitsgedächtnisses in den Blick und benennt Möglichkeiten, das *Extraneous Processing* zu reduzieren, wie zum Beispiel durch Berücksichtigung des Redundanzprinzips, um eine kognitive Überlastung zu verhindern und ausreichend Kapazität für das *Essential* und *Generative Processing* sicherzustellen (Mayer, 2005, 2014). Werden Animationen oder Abbildungen mit gesprochenem Text und redundantem geschriebenem Text kombiniert, kann es nach Mayer

durch den Versuch des Abgleichs von gesprochenem und geschriebenem Text sowie durch die gleichzeitige Verarbeitung der Animation bzw. Abbildung und des geschriebenen Texts im visuellen Kanal des Arbeitsgedächtnisses zu einer kognitiven Überlastung kommen. Der redundante geschriebene Text sollte deshalb entfernt werden (Mayer, 2005). Im Biologieunterricht oder allgemein im Lehr-Lernkontext, gerade auch in Prüfungssituationen, dürfte im Regelfall geschriebener Text in Kombination mit Bildern eingesetzt werden. In Bezug auf Redundanz ist hier ein Blick auf den Redundanzeffekt nach Sweller (2005b) hilfreich.

Sweller fasst den Begriff des Redundanzeffekts etwas weiter. Aufgrund der in der CLT beschriebenen begrenzten Kapazität des Arbeitsgedächtnisses bei Konfrontation mit neuen Informationen tritt nach Sweller der Redundanzeffekt dann auf, wenn zusätzliche Informationen durch die zusätzliche Belastung des Arbeitsgedächtnisses das Lernen erschweren. Dies kann zum einen der Fall sein, wenn die gleichen Informationen in verschiedenen Formen präsentiert werden, z. B. als Bild und als Text oder als geschriebener und gesprochener Text, und das Entfernen einer der Formen der Informationsdarbietung zu einem größeren Lernerfolg führt (Sweller, 2005b). Zum anderen kann das Entfernen von zusätzlichem erklärendem Material, das für das eigentliche Verständnis überflüssig ist, zu einem größeren Lernerfolg führen (Sweller, 2005b), was Mayer als Kohärenzprinzip bezeichnet (2005). Den Redundanzeffekt bei dem Design von multimedialem Material zu berücksichtigen, ist der CLT folgend eine Möglichkeit, den *Extraneous Cognitive Load* zu verringern und mehr freie Arbeitsgedächtniskapazität für das Lernen zu schaffen (Kalyuga & Sweller, 2014). Dies betrifft generell Situationen, in denen MER dargeboten werden, die verarbeitet werden müssen. Denn auch wenn nicht das Lernen, sondern beispielsweise die Aufgabenlösung das Ziel ist, müssen (relevante) Informationen aus dem Material entnommen und im Arbeitsgedächtnis verarbeitet werden, wofür genügend Kapazität vorhanden sein muss.

Experimentelle Studien, die das Redundanzprinzip in Bezug auf gesprochenen und geschriebenen Text zu Abbildungen bestätigen, wurden zum Beispiel anhand von Material zur Entstehung von Blitzen von Mayer et al. (2001) durchgeführt. Studierende, die Animationen mit gesprochenem Text erhielten, zeigten einen größeren Lernerfolg als Studierende, denen zusätzlich auf dem Bildschirm geschriebener Text angezeigt wurde. Redundanz im Sinne von überflüssigem erläuterndem Material wurde zum Beispiel von Reder und Anderson (1980, 1982) untersucht mit dem Ergebnis, dass es besser war, wenn Personen auf die Hauptaussagen eines Originaltexts gekürzte Texte zu einem Thema erhielten, als wenn sie mit dem Originaltext lernten. Da in der vorliegenden Arbeit die Redundanz in Text-

Bild-Kombinationen in der Biologie im Fokus steht, werden im Folgenden empirische Arbeiten vorgestellt, die Redundanz in solchen Kombinationen meist in naturwissenschaftlichen Kontexten untersucht haben. Hier gab es sowohl experimentelle Studien, die positive Effekte von Redundanz gezeigt haben, als auch solche, die negative Effekte aufzeigten. Zunächst sollen Studien mit positiven Effekten von Redundanz vorgestellt werden.

Lindner et al. (2018) untersuchten anhand einer Stichprobe aus Fünft- und Sechstklässler:innen die Effekte von Abbildungen im Itemstamm von Multiple-Choice-Items zu naturwissenschaftlichen Themen. Es zeigte sich, dass redundante Bilder, die dem Itemstamm hinzugefügt worden waren, die Itemschwierigkeit im Vergleich zu rein textlichen Items reduzierte, wenn die im Itemstamm dargestellten Informationen lösungsrelevant waren. Saß et al. (2012) untersuchten ebenfalls Effekte von Abbildungen in Multiple-Choice-Items in Physik anhand einer Stichprobe aus Viertklässler:innen. Das Hinzufügen redundanter Abbildungen zum Itemstamm, zu den Antwortmöglichkeiten oder zu Stamm und Antwortmöglichkeiten wirkte sich positiv auf die Testleistung der Schüler:innen aus. Hegarty und Just (1993) führten eine Studie mit $N = 47$ Studierenden zum Verständnis von Flaschenzugsystemen durch, in der entweder mit rein textlichem Material, mit Abbildungen oder mit redundanten Text-Bild-Kombinationen gearbeitet wurde. Das Arbeiten mit redundanten Text-Bild-Kombinationen führte zum besten Verständnis der Kinematik der Flaschenzugsysteme. Corradi et al. (2014) führten eine Studie mit $N = 92$ Studierenden mit geringem Vorwissen im Fach Chemie durch und variierten in vier Experimentalgruppen Material zu chemischen Themen. Die Studierenden arbeiteten nacheinander mit einem Text, dann mit einer submikroskopischen Abbildung und anschließend mit Text und Abbildung kombiniert. In den vier Gruppen variierte der Grad der Redundanz zwischen Text und Abbildung (0, 25, 50 oder 100 %). Studierende der 25 %-Redundanz-Gruppe steigerten ihr konzeptuelles Verständnis am meisten.

Ein negativer Effekt von Redundanz wurde unter anderem von Bobis et al. (1993) beobachtet. Sie führten vier Experimente mit 20–45 Viertklässler:innen durch, in denen diese verschiedenartig gestaltete Anleitungen zum Falten eines Papiers erhielten. Das Arbeiten mit allein verständlichen Abbildungen der auszuführenden Faltschritte führte zu größerem Erfolg als das Arbeiten mit zusätzlichem redundantem, allein aber nicht zu verstehendem Text. Auch gegenüber dem Arbeiten mit allein verständlichem Text oder der Kombination aus diesem und allein verständlichen Abbildungen (d. h. Redundanz) führte das Arbeiten mit Abbildungen zu einem besseren Abschneiden, während sich zwischen den beiden anderen Gruppen kein signifikanter Unterschied zeigte. Der Einsatz redundanter

Abbildungen zusätzlich zu allein verständlichen Abbildungen führte ebenfalls zu einer schwächeren Leistung der Schüler:innen (Bobis et al., 1993). Bobis et al. (1993) erklärten die Ergebnisse mit einem erhöhten *Extraneous Cognitive Load* bei Anweisungen mit redundantem Material. Pociask und Morrison (2008) untersuchten in einer Studie mit kleiner Stichprobe aus $N = 41$ Physiotherapie-Studierenden unter anderem den Redundanzeffekt. Die Studierenden bearbeiteten Material ihres Fachgebiets, das Anweisungen enthielt, wie die Ursachen von Symptomen von Patient:innen ausfindig gemacht werden können. Das Arbeiten mit Text-Bild-Material, in dem Redundanz vermieden wurde, führte verglichen mit redundantem Text-Bild-Material zu einem besseren Abschneiden der Studierenden. Dies wurde durch eine Verminderung des *Extraneous Cognitive Load* und somit einer Erhöhung des *Germane Cognitive Load* erklärt (Pociask & Morrison, 2008). Durch das Entfernen von redundantem Material war der Umfang des Materials der entsprechenden Gruppe verglichen mit dem der Kontrollgruppe allerdings geringer.

Die Studienergebnisse zeigen die verschiedenen Auswirkungen, die Redundanz in MER-Material haben kann, gleichzeitig ist aber davon auszugehen, dass Lehrende bei dem Design von Lern- und Testmaterialien bezüglich Redundanz von Informationen meistens keinen klaren Konstruktionsprinzipien folgen. Wie eingangs dieses Unterkapitels erläutert sollten Schüler:innen auch mit Blick auf das Ziel einer naturwissenschaftlichen Grundbildung und wissenschaftspropädeutischer Bildung durch naturwissenschaftlichen Fachunterricht in der Schule befähigt werden, mit MER, die sich in ihrer Redundanz unterscheiden, zurechtzukommen. Helfen kann dabei eine Stärkung der Repräsentationskompetenz von Schüler:innen, die Gegenstand des folgenden Unterkapitels ist.

2.4 Repräsentationskompetenz

In Kapitel 2.1 wurde bereits das Konzept der Scientific Literacy vorgestellt, deren grundlegender Aspekt die Fähigkeit zum Lesen und Schreiben naturwissenschaftlicher Texte (inklusive bildlicher Repräsentationen) ist (Norris & Phillips, 2003). Beim Erreichen dieses Aspekts von Scientific Literacy spielt die Repräsentationskompetenz eine große Rolle. In dieser Arbeit werden unter dem Begriff Repräsentationskompetenz zwei Begriffe subsumiert: die *representational competence* (Kozma et al., 2000; Kozma & Russell, 1997, 2005) und die *metarepresentational competence* (diSessa, 2004; diSessa & Sherin, 2000). Trotz Unterschieden, auf die nachfolgend eingegangen wird, haben diese beiden Konzepte

viel gemeinsam. Beide beziehen sich auf Fähigkeiten zur Interpretation, Translation, Konstruktion und zum Vergleichen und Kritisieren von Repräsentationen sowie Fähigkeiten, ihre Funktionsweise verstehen zu können (Nitz, 2012). Zunächst wird nun auf die *representational competence* näher eingegangen.

Der Begriff *Repräsentation* wurde in Kapitel 2.2 detailliert eingeführt. Kozma und Russell (2005) nutzen ihn zwar in erster Linie für das, was nach der Kategorisierung von Schnotz (s. Kapitel 2.2.2) als depiktionale Repräsentation eingeordnet würde. Zusätzlich schließen sie Gleichungen mit ein und verweisen darauf, dass sie mit dem Begriff Repräsentation das bezeichnen, was unter anderem auch *Inscription* (s. Kapitel 2.2.2 dieser Arbeit, z. B. Pozzer-Ardenghi & Roth, 2010) genannt wird. In einem älteren Artikel bezeichnen sie aber explizit auch Verbales als Repräsentation (Kozma & Russell, 1997). Die *representational competence* wurde basierend auf Arbeiten im Bereich der Chemie – unter anderem einer Untersuchung der Nutzung von Repräsentationen durch Chemiker:innen bei ihrer Laborarbeit (Kozma et al., 2000) sowie dem Vergleich von Expert:innen und Noviz:innen aus der Chemie bezüglich des Gruppierens von Repräsentationen und des Transformierens von einer Repräsentationsart in eine andere (Kozma & Russell, 1997) – definiert als „a set of skills and practices that allow a person to reflectively use a variety of representations or visualizations, singly and together, to think about, communicate, and act on chemical phenomena in terms of underlying, a perceptual physical entities and processes [...]“ (Kozma & Russell, 2005, S. 131). Die Fähigkeiten, die die *representational competence* umfasst, listen Kozma und Russell (2005) in folgenden – hier teilweise gekürzt dargestellten – Punkten auf:

Die Fähigkeit, ...

- 1.) Repräsentationen zur Beschreibung beobachtbarer chemischer Phänomene und deren zugrunde liegender Prozesse zu nutzen.
- 2.) eine bestimmte Repräsentation für einen bestimmten Zweck konstruieren oder auswählen sowie die Wahl der Repräsentation begründen zu können.
- 3.) verbal bestimmte Merkmale oder Muster von Merkmalen in verschiedenen Repräsentationen identifizieren und analysieren zu können.
- 4.) Ähnlichkeiten und Unterschiede in der Aussage(kraft) verschiedener Repräsentationen beschreiben zu können.

- 5.) Verbindungen zwischen Repräsentationen herstellen und Beziehungen zwischen ihnen erklären zu können sowie Merkmale einer Repräsentation auf eine andere übertragen zu können.
- 6.) die Korrespondenz, aber auch den Unterschied zwischen Repräsentation und beobachtetem Phänomen erkennen zu können.
- 7.) in sozialen Situationen mit Hilfe von Repräsentationen Behauptungen untermauern und Schlüsse ziehen sowie Vorhersagen über chemische Phänomene treffen zu können.

Personen mit geringer *representational competence* fokussieren eher Oberflächenmerkmale von Repräsentationen, während Personen mit hoher *representational competence* Phänomene anhand verschiedener Repräsentationen erklären und sie zur Problemlösung, für Vorhersagen oder zum Untermauern von Behauptungen im fachlichen Diskurs heranziehen können (Kozma & Russell, 2005). Kozma und Russell (2005) sehen die *representational competence* in fünf Level mit ansteigender Schwierigkeit strukturiert, die von der Nutzung von Repräsentationen als einfaches Abbild eines physischen Phänomens, über den zusätzlichen Gebrauch symbolischer Elemente (z. B. Pfeile für dynamische Prozesse), einen syntaktischen und semantischen Gebrauch verschiedener Repräsentationen bis hin zu einer reflektierten Nutzung von Repräsentationen im fachlichen Diskurs reichen. Je nach Kontext, Zeitpunkt oder Repräsentationsart kann das Level der Repräsentationskompetenz einer Person variieren, es wird aber davon ausgegangen, dass sich mit der Zeit und verschiedenen Berührungspunkten mit Repräsentationen die repräsentationalen Fähigkeiten verbessern (Kozma & Russell, 2005). Auch über den Kontext Chemie hinaus wurden die für die Repräsentationskompetenz zentralen Fähigkeiten untersucht. Halverson (2011) beobachtete die Entwicklung des Lesens und Konstruierens phylogenetischer Bäume durch Studierende über ein Semester hinweg und identifizierte gerade für das Lesen phylogenetischer Bäume ähnliche zentrale Fähigkeiten zum Aufbau einer diesbezüglichen Repräsentationskompetenz. Einen Teilaspekt der Repräsentationskompetenz stellt auch die Repräsentationskompetenz in Bezug auf Achsendiagramme dar, die in den Naturwissenschaften und im naturwissenschaftlichen Unterricht häufig genutzt werden. Lachmayer et al. (2007) schlugen ein Modell der kognitiven Fähigkeiten für den Gebrauch von Achsendiagrammen vor, d. h. zur Diagrammkompetenz, das die Komponenten Informationsentnahme (Identifizierung und Ablesen), Konstruktion (Aufbau des Rahmens und Dateneintrag) und In-

tegration enthält. Dieses Kompetenzstrukturmodell wurde durch Einsatz eines Testinstruments mit Aufgaben zu Linien- und Säulendiagrammen, die Neunt- und Zehntklässler:innen bearbeiteten, empirisch überprüft und bestätigt (Lachmayer, 2008).

Die *metarepresentational competence* nach diSessa (2004) bzw. diSessa und Sherin (2000), die im Kontext der Mathematik und Naturwissenschaften beschrieben wird, geht in erster Linie in dem Sinne über die von Kozma und Russell beschriebene *representational competence* hinaus, als hier auch das Generieren neuer Repräsentationen eine wichtige Fähigkeit darstellt. Die *metarepresentational competence* umfasst folgende Komponenten (diSessa, 2004):

- 1.) Neue Repräsentationen entwerfen oder erfinden zu können
- 2.) Zur Kritik und zum Vergleich der Angemessenheit bestimmter Repräsentationen fähig zu sein sowie ihre Eignung für verschiedene Zwecke beurteilen zu können
- 3.) Verstehen zu können, welche Zwecke Repräsentationen erfüllen (allgemein und in speziellen Kontexten) und wie sie dies tun
- 4.) Repräsentationen erklären zu können
- 5.) Schnell und möglichst selbstständig neue Repräsentationen erlernen zu können

Die Nutzung des Präfixes *meta* begründet diSessa (2004) damit, dass eine Kompetenz beschrieben werden soll, die über das bloße Nutzen und Konstruieren von Repräsentationen hinausgeht und führt an, dass es dazu dienen solle „to warn against what we consider limitations of narrower views of representational competence that have dominated the study of representation in science and mathematics education, specifically focusing on production and interpretation of a limited set of sanctioned representations“ (diSessa, 2004, S. 294). Die von Kozma und Russell (1997, 2005) beschriebene *representational competence* geht meines Erachtens über die im Zitat beschriebene sehr enge Sicht des Konzepts hinaus (z. B. umfasst sie auch die Fähigkeit zur Begründung der Auswahl einer bestimmten Repräsentation und zum Vergleich der Aussagekraft verschiedener Repräsentationen) und weist wie oben erwähnt eine deutliche Überschneidung mit der *metarepresentational competence* auf.

Des Weiteren weisen diSessa und Sherin (2000) darauf hin, dass selbst junge Schüler:innen schon durch Alltagsaktivitäten eine *metarepresentational competence* besitzen, die zum Beispiel durch das Ausüben von Tätigkeiten wie Zeichnen entsteht. Auf diese

grundlegend vorhandene Kompetenz kann dann im naturwissenschaftlichen und mathematischen Unterricht hinsichtlich des Lernens über Repräsentationen aufgebaut werden (diSessa & Sherin, 2000). Die *metarepresentational competence* betreffende Aufgaben wie das Entwickeln neuer Repräsentationen erfordern Kreativität und künstlerische Fähigkeiten und könnten bei entsprechender Berücksichtigung im Unterricht auch dazu beitragen, Interesse für Naturwissenschaften und Mathematik bei solchen Schüler:innen zu wecken, die sich sonst weniger für mathematisch-naturwissenschaftliche Fächer interessieren (diSessa & Sherin, 2000). *Metarepresentational competence* könnte nach diSessa (2004) zudem dazu führen, dass Schüler:innen einen Sinn im Erlernen von Repräsentationen erkennen, und könnte so ihre Motivation steigern. Auch sollte *metarepresentational competence* Schüler:innen unter anderem in die Lage versetzen, Repräsentationen in naturwissenschaftlichen Kontexten in den Medien, egal ob durch Wissenschaftler:innen, Reporter:innen oder andere Personen erstellt, kritisch beurteilen zu können (diSessa, 2004).

Neben den in diesem Unterkapitel bereits erwähnten Studien haben weitere empirische Arbeiten die Repräsentationskompetenz Studierender in naturwissenschaftlichen Kursen untersucht. Während in Kapitel 2.3 bereits Studien vorgestellt wurden, die das Arbeiten mit MER zu naturwissenschaftlichen Themen und damit das Verbinden und Integrieren von Repräsentationen untersucht haben, werden im Folgenden Studien angeführt, die das Arbeiten mit bestimmten Repräsentationsarten in den Naturwissenschaften in den Blick genommen haben. Dabei wird auf Studien zur Informationsentnahme und Interpretation fokussiert, da auch in diesem Promotionsprojekt die Informationsentnahme und Interpretation aus bildlichen Repräsentationen beziehungsweise MER untersucht wird und nicht deren Konstruktion. Beispielsweise konnten Schönborn et al. (2002) verschiedene Arten von Schwierigkeiten Biochemiestudierender im zweiten oder dritten Jahr identifizieren, wenn diese mit einer schematischen Abbildung der Tertiärstruktur von Immunglobulin G mit gebundenen Antigenen konfrontiert waren. Getätigte Aussagen wiesen darauf hin, dass es ihnen teilweise nicht gelang, das zugrunde liegende Konzept (Antigen-Antikörper-Bindung) zu identifizieren, und sie stattdessen die Darstellung komplexer Prozesse vermuteten oder dass es ihnen teilweise nicht gelang, die Bedeutung einzelner dargestellter Elemente zu verstehen (z. B. ein kurzer schwarzer Strich bedeutete Disulfidbindung; Schönborn et al., 2002). Kottmeyer et al. (2020) untersuchten, wie gut Studierende

Aufgaben zu verschiedenen, aus Highschool-Lehrbüchern entnommenen Abbildungen⁹ lösen konnten. Der Großteil der Abbildungen behandelte Themen aus der Biologie und der restliche Teil Themen der Geowissenschaften. Die Studierenden befanden sich in den ersten beiden Jahren ihres meist biologiebezogenen Studiums und nahmen an einem Humanphysiologiekurs teil. Für jede Abbildung wurde die Art der genutzten Konventionen bestimmt und der Effekt verschiedener Konventionen auf die Testleistung untersucht. Im Durchschnitt wurde eine als schwach eingeordnete Leistung (nur ca. 70 % der Fragen korrekt beantwortet) der Studierenden bei der Aufgabenbearbeitung festgestellt. Bezüglich der Konventionen stellte sich heraus, dass – in Kombination mit weiteren Konventionen – Schattierungen in Abbildungen mit der besten Testleistung einhergingen, Nummerierungen durch Ziffern oder Buchstaben mit der geringsten (Kottmeyer et al., 2020). Farheen und Lewis (2021) untersuchten in einem Chemiekurs mit Studierenden des zweiten Semesters, wie sich die Repräsentationsart (Summenformel, Lewis-Struktur, Bild eines Kugel-Stab-Modells, Bild eines Kalottenmodells) zur Darstellung eines Schwefeldioxid-Moleküls darauf auswirkt, wie gut Studierende chemische Eigenschaften (z. B. relative Bindungslänge, Polarität) des Moleküls einschätzen konnten. Teilweise zeigten sich Unterschiede in der korrekten Einschätzung je nach Repräsentationsart. Je nach chemischer Eigenschaft eigneten sich zudem unterschiedliche Repräsentationsarten am besten zur Vorhersage der Eigenschaft (Farheen & Lewis, 2021). Durch Lerngelegenheiten mit verschiedenen Repräsentationsarten könnte die Repräsentationskompetenz gestärkt werden. Ein weiterer Faktor, der bei Konfrontation mit Repräsentationen eine Rolle spielen könnte, wird im folgenden Kapitel beleuchtet.

2.5 Selbstwirksamkeitserwartung und ihre Bedeutung für Leistungen in mathematisch-naturwissenschaftlichen Fächern

Das erfolgreiche Arbeiten mit verschiedenen Repräsentationsarten muss nicht allein durch das Wissen über diese Repräsentationsarten oder fachliches Vorwissen geprägt sein, sondern könnte auch auf die persönliche Überzeugung zurückgehen, mit den jeweiligen Repräsentationsarten arbeiten zu können. Gerade in Bezug auf komplexe Abbildungen zum Beispiel aus der Biologie könnte eine solche Überzeugung eine Rolle spielen. Dies wird im

⁹ Der von Kottmeyer et al. (2020) eingesetzte Test enthält 34 Items zu verschiedenen Abbildungstypen „including schematic, photographic, and representational“ (S. 174). Als Beispiele sind drei schematische Abbildungen dargestellt, die den Vorgang der Bestäubung, der Meiose und verschiedene Wnt-Signalwege zeigen, weshalb anzunehmen ist, dass dieser Abbildungstyp im Test dominiert.

Rahmen dieser Arbeit untersucht, weshalb im Folgenden das Konstrukt Selbstwirksamkeitserwartung beleuchtet wird, das einen zentralen Aspekt in Banduras *Social Cognitive Theory* darstellt (Bandura, 1986, 2001). Anschließend werden empirische Studien zum Zusammenhang von Selbstwirksamkeitserwartung und akademischer Leistung im mathematisch-naturwissenschaftlichen Bereich vorgestellt.

2.5.1 Begriffserläuterung und theoretische Einbettung

Bandura nimmt in der *Social Cognitive Theory* (SCT) unter anderem die Bedeutung der Fähigkeit zur Selbstreflexion für das menschliche Handeln in den Blick (Bandura, 2001). Zunächst sind drei Arten des Handelns zu unterscheiden: persönliches Handeln, Handeln durch Stellvertreter:innen und kollektives Handeln (*personal, proxy, collective agency*; Bandura, 2001). Da niemand Zeit, Ressourcen und Expertise hat, Aspekte aller Lebensbereiche selbst zu bewältigen, sind Personen teilweise auf Stellvertreter:innen angewiesen, die in ihrem Sinne handeln sollen (Bandura, 2001). Beispielsweise wenden sich Kinder für das Ausführen bestimmter Handlungen an ihre Eltern und Bürger:innen an Repräsentant:innen der Legislative, damit diese für sie die gewünschten Ergebnisse erzielen (Bandura, 2001). Kollektives Handeln bezieht sich beispielsweise auf Familien, Sportmannschaften, Organisationen oder politische Gruppen und basiert auf der geteilten Überzeugung, gemeinsam erwünschte Ergebnisse erbringen zu können (Bandura, 1997, 2001). Hauptmerkmale persönlichen menschlichen Handelns sind nach Bandura (2001) die Intentionalität, vorausschauendes Denken, Selbstreaktivität bzw. Selbstregulation und Selbstreflexion. Letztere beinhaltet Selbstwirksamkeitserwartungen – Überzeugungen, durch eigenes Handeln bestimmte Ergebnisse erzielen zu können –, die als grundlegend für menschliches Handeln angesehen werden (Bandura, 2001).

Die Selbstwirksamkeitserwartung war und ist Forschungsgegenstand in verschiedenen Bereichen, zum Beispiel in der Forschung zu Phobien, bestimmten Erkrankungen und Gesundheitsaspekten (z. B. Bandura et al., 1977; Benight & Bandura, 2004; Grembowski et al., 1993) und spielt seit einigen Jahrzehnten auch in der Bildungsforschung eine große Rolle, wie beispielsweise Usher und Pajares (2008) in ihrem Übersichtsartikel zeigen. Sie gehört wie unter anderem das Fähigkeitsselbstkonzept und das Selbstwertgefühl zu den sogenannten Self-Beliefs, die mit akademischer Leistung assoziiert sein können (Valentine et al., 2004). Für die Selbstwirksamkeitserwartung bedeutet das zum Beispiel,

dass von Schüler:innen mit gleicher Fähigkeit in einem Aufgabengebiet trotzdem diejenigen eine bessere Leistung erbringen könnten, die eine höhere Selbstwirksamkeitserwartung bezüglich dieses Aufgabengebiets besitzen, also überzeugt sind, in diesem Aufgabengebiet höhere Leistungen erbringen zu können (Valentine et al., 2004), worauf in Kapitel 2.5.4 anhand von empirischen Arbeiten näher eingegangen wird.

2.5.2 Charakteristika der Selbstwirksamkeitserwartung

Im Folgenden wird in den Blick genommen, durch welche Charakteristika die Selbstwirksamkeitserwartung gekennzeichnet ist. Hinsichtlich der genannten Self-Beliefs besteht der Unterschied zwischen Selbstwirksamkeitserwartung und Selbstwertgefühl vor allem darin, dass sich die Selbstwirksamkeitserwartung anders als das Selbstwertgefühl auf die Beurteilung persönlicher Fähigkeiten und nicht auf die Beurteilung des Werts der eigenen Person bezieht (Bandura, 1986). Selbstwirksamkeitserwartung und Fähigkeitsselbstkonzept haben viele Gemeinsamkeiten (z. B. prädiktive Bedeutung in Bezug auf Leistung in einer Domäne) und können miteinander korrelieren, sind aber konzeptuell und empirisch trennbare Konstrukte (z. B. Jansen et al., 2015; Marsh et al., 2019). Das Fähigkeitsselbstkonzept bezieht sich auf die Wahrnehmung des Selbst in Bezug auf eine bestimmte Domäne und basiert stark auf dem sozialen Vergleich mit anderen (z. B. Mitschüler:innen), wohingegen die Selbstwirksamkeitserwartung sich wie beschrieben auf die Überzeugung bezieht, mit jeweiligen Fähigkeiten eine konkrete Aufgabe bewältigen zu können, und meist weniger stark aus dem Vergleich mit anderen abgeleitet wird (Bong & Skaalvik, 2003; Marsh et al., 2019). Die Selbstwirksamkeitserwartung ist somit anders als das Fähigkeitsselbstkonzept nicht nur domänen-, sondern auch aufgabenspezifisch (Pajares, 1996). Die Messung einer Selbstwirksamkeitserwartung, die stark auf die untersuchte Aufgabe oder Tätigkeit zugeschnitten ist, hat eine höhere Vorhersagekraft in Bezug auf das Resultat als die Messung einer allgemeineren Selbstwirksamkeitserwartung (Bandura, 1986; Pajares, 1996). Bandura (1986) geht aber dennoch davon aus, dass sich eine entwickelte Selbstwirksamkeitserwartung für eine bestimmte Aktivität auch auf ähnliche Aktivitäten auswirken und somit in gewissem Maß generalisierbar sein könnte. Eine empirische Arbeit von Bong (1997) zur Allgemeingültigkeit der Selbstwirksamkeitserwartung im schulischen Kontext zeigte, dass die Selbstwirksamkeitserwartung über spezifische Aufgaben und Schulfächer hinweg generalisierbar ist, wenn sich diese ähneln. Mittels konfirmatorischer Faktorenanalyse konnten bei der Untersuchung der Selbstwirksamkeitserwartung von

Theoretischer Hintergrund

Highschool-Schüler:innen bezüglich verschiedener Schulfächer zwei übergeordnete Selbstwirksamkeitserwartungen identifiziert werden: Eine quantitative Selbstwirksamkeitserwartung für die Fächer Geometrie, Algebra und Chemie und eine verbale Selbstwirksamkeitserwartung für die Fächer Englisch und US-amerikanische Geschichte (Bong, 1997). Die Hypothese, dass dies möglicherweise auf die wahrgenommene Ähnlichkeit der zu bewältigenden Aufgaben zurückzuführen sei, wurde beispielhaft anhand von Algebra- und Physikaufgaben geprüft, indem Schüler:innen ihre Selbstwirksamkeitserwartung bezüglich dieser Aufgaben einschätzten und anschließend die Ähnlichkeit der Aufgaben beurteilten. Je höher die Ähnlichkeit eingeschätzt wurde, desto näher lag die Stärke der zu diesen Aufgaben angegebenen Selbstwirksamkeitserwartungen beieinander, wobei dieser Zusammenhang schwächer war als erwartet (Bong, 1997). Trotz der Hinweise auf eine hierarchische Struktur der Selbstwirksamkeitserwartung mit der Generalisierbarkeit über Aufgaben und Schulfächer hinweg ist laut Bong (1997) eine Ableitung einer aufgaben- oder fachspezifischen Selbstwirksamkeitserwartung aus einer übergeordneten Selbstwirksamkeitserwartung nicht zwingend möglich. Sie empfiehlt zudem auch, dass die Spezifität der eingeschätzten Selbstwirksamkeitserwartung (z. B. Selbstwirksamkeitserwartung bezüglich einer spezifischen Aufgabe, eines Kurses wie Algebra oder einer Domäne wie Mathematik) der Spezifität der Leistung, die vorhergesagt werden soll, entsprechen sollte. So sei zum Beispiel die Vorhersage der Leistung im Schulfach Englisch durch die übergeordnete verbale Selbstwirksamkeitserwartung nicht unbedingt vergleichbar mit der Vorhersage durch die Selbstwirksamkeitserwartung bezüglich des Fachs Englisch selbst (Bong, 1997). Bei der Operationalisierung der Selbstwirksamkeitserwartung sind also oben genannte Punkte zu beachten.

Die im letzten Absatz genannte Generalisierbarkeit ist eine von drei Dimensionen (*generality, level, strength*), in denen sich Selbstwirksamkeitserwartungen nach Bandura (1986) unterscheiden. Sie bezieht sich laut Bandura darauf, ob eine Person ihre Selbstwirksamkeitserwartung nur auf bestimmte Domänen oder ein breiteres Spektrum an Situationen oder Tätigkeiten bezieht. Unterschiede im Level von Selbstwirksamkeitserwartungen bedeuten, dass sich die Überzeugungen von Personen, Aufgaben innerhalb einer Domäne erfolgreich bewältigen zu können, hinsichtlich der Schwere der als bewältigbar empfundenen Aufgaben unterscheiden können (Bandura, 1986). Manche Personen könnten also überzeugt sein, leichte Aufgaben einer bestimmten Domäne bewältigen zu können, während andere Personen überzeugt sind, mittelschwere oder schwere Aufgaben derselben Domäne bewältigen zu können. Darüber hinaus variiert die Selbstwirksamkeitserwartung in Bezug

auf eine bestimmte Aufgabe oder Tätigkeit in ihrer Stärke, kann also bei einer Person schwächer oder stärker ausgeprägt sein (Bandura, 1986).

Selbstwirksamkeitserwartungen können sich in verschiedener Weise auf das Verhalten von Personen auswirken (Pajares, 1996). Die Selbstwirksamkeitserwartung in einem bestimmten Bereich kann sich zum Beispiel auf die aufgebrachte Anstrengung bei der Ausübung einer Tätigkeit oder Aufgabe auswirken, darauf, wie widerstandsfähig eine Person in schwierigen Situationen ist, wie lange eine Person durchhält, wenn Hindernisse auftreten (Persistenz), ob eine Person eine bestimmte Aufgabe überhaupt angeht oder direkt meidet und kann sich letztendlich auf die Leistung auswirken (Bandura, 1977; Pajares, 1996). Eine höhere Selbstwirksamkeitserwartung geht mit höherer Anstrengung, Widerstandsfähigkeit, höherem Durchhaltevermögen, einer höheren Wahrscheinlichkeit, eine Aufgabe oder Tätigkeit anzugehen, und besserer Leistung einher (Bandura, 1977; Pajares, 1996). Für Letzteres müssen allerdings auch benötigte Fähigkeiten vorhanden sein (Bandura, 1977). Empirische Befunde zu Effekten der Selbstwirksamkeitserwartung auf die akademische Leistung in Naturwissenschaften und Mathematik werden in Kapitel 2.5.4 vorgestellt. In Bezug auf Anstrengung und Durchhaltevermögen bzw. Persistenz zeigten Galla et al. (2014) in einer Studie mit relativ jungen Schüler:innen, dass deren akademische Selbstwirksamkeitserwartung ihr von Lehrkräften eingeschätztes Engagement (u. a. erfasst als Anstrengung und Persistenz) vorhersagte und darüber indirekt auch ihre Leistung in den Bereichen Lesen und Mathematik. Lent et al. (1986) wiesen ebenfalls in Bezug auf die Persistenz nach, dass eine höhere Selbstwirksamkeitserwartung Studierender bezogen auf Natur- und Ingenieurwissenschaften mit einer längeren Verweildauer (betrachtet wurde die Anzahl absolvierter Quartale im Studienjahr nach Teilnahme an einem Kurs zur Karriereplanung) an einem technischen College einherging und hier als Prädiktor fungierte. Zudem konnte die Selbstwirksamkeitserwartung zur Vorhersage der Bandbreite der von den Studierenden in Erwägung gezogenen Karriereoptionen im genannten Bereich herangezogen werden (Lent et al., 1986) und wirkte sich somit auf die Entscheidung von Personen aus, eine bestimmte Tätigkeit (hier die Ausübung eines bestimmten Berufs) eher anzustreben oder zu meiden. Neben den genannten motivationalen Auswirkungen – wie Anstrengung, Widerstandsfähigkeit und Wahl einer Aufgabe oder Tätigkeit – sowie Auswirkungen auf die Leistung kann die Selbstwirksamkeitserwartung unter anderem auch die Selbstregulation einer Person beeinflussen (Schunk & Usher, 2012). Dies betrifft beispielsweise das Setzen von Zielen und die

Anwendung effektiver Lernstrategien, was bei Personen mit höherer auf das Lernen bezogener Selbstwirksamkeitserwartung stärker ausgeprägt sein dürfte (Schunk & DiBenedetto, 2016; Schunk & Usher, 2012).

Nachdem die Effekte von Selbstwirksamkeitserwartungen erläutert wurden, werden im Folgenden die Informationsquellen für die Bildung von Selbstwirksamkeitserwartungen betrachtet. Die Selbstwirksamkeitserwartung wird wie oben beschrieben weniger stark aus dem Vergleich mit anderen Personen abgeleitet als das Fähigkeitsselbstkonzept, wobei dieser Vergleich durchaus einfließen kann, wie im nächsten Unterkapitel erläutert wird.

2.5.3 Quellen der Selbstwirksamkeitserwartung

Bandura (1977) benennt vier Hauptinformationsquellen, aus denen Personen ihre Selbstwirksamkeitserwartung hauptsächlich ableiten¹⁰:

- Erfolgserfahrungen
- stellvertretende Erfahrungen
- verbale Überzeugungsarbeit
- emotionale Erregung

Auf diese vier Quellen wird im Folgenden absatzweise eingegangen. Zunächst soll jede Quelle primär anhand von Banduras Ausführungen, in die er verschiedene empirische Befunde einbezieht, vorgestellt werden. Im Anschluss werden empirische Befunde zu diesen Informationsquellen für die Selbstwirksamkeitserwartung von Schüler:innen oder Studierenden bezüglich Naturwissenschaften oder Mathematik angeführt.

Erfolgserfahrungen werden als die einflussreichste Informationsquelle der Selbstwirksamkeitserwartung betrachtet (Bandura, 1977). Wenn zum Beispiel Schüler:innen eine bestimmte Leistung in der Schule erbringen sollen, beurteilen sie anschließend basierend auf ihren erzielten Ergebnissen ihre Fähigkeiten bezüglich derartiger Aufgaben (Usher &

¹⁰ In der deutschsprachigen Literatur zur Selbstwirksamkeitserwartung werden verschiedene Übersetzungen für die vier Begriffe genutzt. In der vorliegenden Arbeit wurde angestrebt, die deutschen Begriffe möglichst stark an die englischen Begriffe von Bandura (1977) anzulehnen: *performance accomplishments*, *vicarious experience*, *verbal persuasion*, *emotional arousal*. Bandura selbst nutzt allerdings auch unterschiedliche englische Begriffe (vgl. z. B. Bandura, 1986).

Pajares, 2008). Dies könnte zum Beispiel das Arbeiten oder Lösen von Aufgaben mit komplexen oder wenig vertrauten Abbildungen im naturwissenschaftlichen Unterricht betreffen. Häufige Erfolgserfahrungen in Bezug auf ein bestimmtes Gebiet bzw. eine Problemstellung führen zu einer gestärkten Selbstwirksamkeitserwartung, diese Aufgaben (auch zukünftig) bewältigen zu können (Bandura, 1977). Zwar könnten wiederholte Misserfolge diese Überzeugung schmälern, ist die positive Selbstwirksamkeitserwartung aber erst einmal stark ausgeprägt, wird davon ausgegangen, dass gelegentliche Misserfolge auf dem jeweiligen Gebiet bezüglich der Ausprägung der zugehörigen Selbstwirksamkeitserwartung nicht mehr stark ins Gewicht fallen (Bandura, 1977). Vielmehr würden diese Misserfolge dann vermutlich etwa auf situative Faktoren oder fehlerhafte Strategien anstatt auf eigenes Unvermögen zurückgeführt werden (Bandura, 1986).

Stellvertretende Erfahrungen. Eine weitere Quelle ist das Beobachten einer Aufgabenbearbeitung durch andere Personen, sogenannte stellvertretende Erfahrungen (Bandura, 1977, 1986). Bewältigt eine als ähnlich empfundene Person die Aufgabe erfolgreich, kann die eigene Zuversicht bzw. Erwartung einer Person steigen, derartige Aufgaben ebenfalls erfolgreich bewältigen zu können (Bandura, 1986). Schwarzer und Jerusalem (2002) führen beispielhaft eine potentiell effektive Nachhilfe- oder Tutoriumssituation an, bei der ein:e Schüler:in durch eine:n andere:n Schüler:in unterrichtet wird, die oder der selbst gerade erst die zu bewältigende Aufgabe unter Schwierigkeiten erfolgreich gelöst hat. Scheitert die als vergleichbar kompetent wahrgenommene Person trotz Anstrengung, kann die auf die jeweilige Aufgabe bezogene Selbstwirksamkeitserwartung der beobachtenden Person aber auch sinken (I. Brown & Inouye, 1978). I. Brown und Inouye (1978) konnten aber gleichzeitig zeigen, dass die Selbstwirksamkeitserwartung einer Person, die eine andere beim Scheitern an einer Aufgabe beobachtet, aber die scheiternde Person von vornherein als weniger kompetent als sich selbst wahrnimmt, höher ist, als wenn die scheiternde Person als ähnlich kompetent wahrgenommen wird, das heißt, die wahrgenommene Ähnlichkeit in der Kompetenz spielt hier eine entscheidende Rolle. Die stellvertretende Erfahrung ist also eine auf einen sozialen Vergleich gestützte Informationsquelle für eigene Fähigkeiten, die aber weniger verlässlich ist als die oben genannten eigenen Erfolgserfahrungen (Bandura, 1977).

Verbale Überzeugungsarbeit durch andere kann sich ebenfalls auf die Selbstwirksamkeitserwartung einer Person auswirken (Bandura, 1977). Allerdings wird angenommen, dass die verbale Überzeugungsarbeit allein keinen starken bzw. stabilen Anstieg der

Selbstwirksamkeitserwartung bewirken kann, sondern beispielsweise weitere Unterstützungsmaßnahmen gewinnbringend ergänzen kann, indem sie etwa die Anstrengung einer Person erhöht (Bandura, 1977). Damit verbale Überzeugungsarbeit die Selbstwirksamkeitserwartung stärken kann, sollte die zu bewältigende Aufgabe für die jeweilige Person tatsächlich machbar sein, da die ansonsten wahrscheinlich auftretenden Misserfolge zum einen die Glaubwürdigkeit der Person, die zu überzeugen versucht, untergraben und zum anderen die Selbstwirksamkeitserwartung sogar schwächen könnten (Bandura, 1977, 1986). Bandura (1986, 1997) vermutet, dass es schwieriger ist, die Selbstwirksamkeitserwartung einer Person allein durch verbale Überzeugungsarbeit dauerhaft zu stärken, als sie zu untergraben. Wenn eine Person durch Äußerungen einer anderen zu der Überzeugung gelangt ist, bestimmte Fähigkeiten nicht zu besitzen, kann dies dazu führen, dass sie zum Beispiel bei Schwierigkeiten schneller aufgibt (Bandura, 1997). Im schulischen Kontext kann verbale Überzeugungsarbeit zur Stärkung der Selbstwirksamkeitserwartung von Schüler:innen zum Beispiel durch Lehrkräfte, aber auch durch Eltern oder Gleichaltrige, denen vertraut wird, geleistet werden (Usher & Pajares, 2008). Schwarzer und Jerusalem (2002) schlagen vor, dass eine Lehrkraft einer/m Schüler:in zum Beispiel explizit rückmelden könnte, dass diese:r eine schwierige Aufgabe aufgrund hoher Kompetenz gelöst hat, also Kompetenz als Ursache für gute Leistung konkret benennt, um so die Selbstwirksamkeitserwartung zu stärken (Schwarzer & Jerusalem, 2002).

Emotionale Erregung. Wenn eine Person einer belastenden oder stressigen Situation ausgesetzt ist, löst dies häufig eine emotionale bzw. körperliche Erregung aus, aus der die Person unter Umständen Informationen über die eigene Kompetenz ableitet (Bandura, 1977, 1997). Ein Zustand starker Erregung, der sich zum Beispiel durch körperliche Reaktionen wie Schwitzen, Herzklopfen, Magenschmerzen oder Zittern äußern kann, kann die Leistung verringern und so kann es sein, dass das Erfahren der Erregung zu einer geringeren Selbstwirksamkeitserwartung führt, als wenn diese Erregung ausbleibt (Bandura, 1997). Allerdings haben verschiedene Faktoren – beispielsweise zurückliegende Erfahrungen zur Beeinflussung der eigenen Leistung durch Erregung – Einfluss darauf, wie die wahrgenommene Erregung kognitiv verarbeitet und beurteilt und somit die Selbstwirksamkeitserwartung beeinflusst wird (Bandura, 1986). Die Beurteilung von Erregungszuständen variiert zwischen Personen: So tendieren manche Personen dazu, Erregung auf eigene Schwäche zurückzuführen, was zur Reduktion ihrer Selbstwirksamkeitserwartung führen kann, während andere ihre Erregung als normale Reaktion hinnehmen, die auch kompetenten Menschen widerfahren kann (Bandura, 1986). Die Wahrnehmung erhöhter Erregung muss nicht

zwingend zu einer Verminderung der Selbstwirksamkeitserwartung führen (Usher & Pajares, 2008).

Verschiedene empirische Studien, die die oben erläuterten Informationsquellen für die Selbstwirksamkeitserwartung bezüglich Naturwissenschaften oder Mathematik von Schüler:innen verschiedener Altersstufen oder Studierenden untersucht haben, konnten (teilweise) einen prädiktiven Charakter dieser Quellen nachweisen (z. B. Britner & Pajares, 2006; Joët et al., 2011; Klassen, 2004; Lent et al., 1991; Matsui et al., 1990). Britner und Pajares (2006) konnten für alle vier Quellen eine statistisch signifikante Korrelation mit der Selbstwirksamkeitserwartung von Schüler:innen im Bereich der Naturwissenschaften nachweisen, bei gleichzeitigem Einfügen der vier Quellen in ein Regressionsmodell konnten nur Erfolgserfahrungen als signifikanter Prädiktor der Selbstwirksamkeitserwartung identifiziert werden, die anderen drei Quellen klärten darüber hinaus keine Varianzanteile auf. Webb-Williams (2018) führte neben quantitativen Messungen von Selbstwirksamkeitserwartung und Leistung im naturwissenschaftlichen Bereich (signifikante positive Korrelation) Interviews mit 10- bis 12-jährigen Schüler:innen durch, um die Quellen der Selbstwirksamkeitserwartung zu bestimmen. Die Antworten der Schüler:innen enthielten Aussagen zu allen vier Quellen, wenn auch in verschiedener Häufigkeit. Auch hier wurden Erfolgserfahrungen als wichtige Quelle der Selbstwirksamkeitserwartung identifiziert, die allerdings für Jungen eine größere Rolle spielte als für Mädchen. Mädchen nannten am häufigsten stellvertretende Erfahrungen bzw. den Vergleich mit anderen als Quelle ihrer Selbstwirksamkeitserwartung (Webb-Williams, 2018). Auch in quantitativen Studien wurde beobachtet, dass sich die Stärke einzelner von Schüler:innen wahrgenommener Quellen der Selbstwirksamkeitserwartung zwischen den Geschlechtern unterschied (z. B. Britner & Pajares, 2006; Joët et al., 2011), wobei zum Beispiel Matsui et al. (1990) in ihrer Studie mit Studierenden diesen Unterschied nicht feststellten.

Wie sich die genannten Informationsquellen tatsächlich auf die Selbstwirksamkeitserwartung auswirken, hängt allerdings davon ab, wie sie kognitiv bewertet werden (Bandura, 1977). Die Stärke des Einflusses zurückliegender Erfolgserfahrungen auf die Selbstwirksamkeitserwartung kann zum Beispiel von der Schwierigkeit der Aufgaben oder externer Hilfe beim Bewältigen der Aufgaben, bei denen die Erfolgserfahrungen gemacht wurden, abhängen (Bandura, 1986). Wurde viel Hilfe in Anspruch genommen, werden die Erfolge nicht unbedingt in hohem Maße auf persönliche Fähigkeiten zurückgeführt und tragen folglich genauso wie Erfolge bei als leicht wahrgenommenen Aufgaben wenig oder gar nicht zu einer Steigerung der Selbstwirksamkeitserwartung bei (Bandura, 1986, 1997).

Auch die Anstrengung, die für den Erfolg aufgebracht werden musste, kann in die Bewertung einfließen, wobei Erfolg bei geringer Anstrengung trotz hoher Schwierigkeit der Aufgabe meist als Indikator für hohe Fähigkeiten gewertet wird (Bandura, 1986). Bei einer solchen Bewertung ist der positive Effekt von Erfolgserfahrungen auf die Selbstwirksamkeitserwartung als höher einzuschätzen, als wenn viel Anstrengung aufgebracht werden musste (Bandura, 1986). Der Effekt von stellvertretenden Erfahrungen auf die Selbstwirksamkeitserwartung kann wie oben bereits beschrieben unter anderem von der durch die beobachtende Person wahrgenommenen Ähnlichkeit zur beobachteten Person abhängen. Diese Ähnlichkeit kann sich sowohl auf vorangegangene Leistungen als auch auf für die jeweilige Fähigkeit als relevant eingestufte persönliche Eigenschaften beziehen (Bandura, 1986). Verbale Überzeugungsarbeit kann dann erfolgreich sein und sich positiv auf die Selbstwirksamkeitserwartung auswirken, wenn diejenigen, die die Überzeugungsarbeit leisten, als glaubwürdig und kenntnisreich auf dem jeweiligen Gebiet wahrgenommen werden (Bandura, 1986). Wie ebenfalls schon oben beschrieben können verschiedene Faktoren eine Rolle bei der Beurteilung emotionaler bzw. körperlicher Erregung spielen und dadurch die daraus resultierende Selbstwirksamkeitserwartung beeinflussen (Bandura, 1986). Neben vorherigen Erfahrungen zur Beeinflussung der eigenen Leistung durch Erregung gehört zu diesen Faktoren laut Bandura unter anderem die Beurteilung der Quelle der Erregung. Ob Schwitzen etwa auf die Raumbedingungen zurückgeführt wird oder als Ausdruck persönlicher Schwäche interpretiert wird, dürfte unterschiedliche Auswirkungen auf die Selbstwirksamkeitserwartung haben (Bandura, 1986).

Im schulischen Kontext könnten die Lehrkräfte eine positive Entwicklung der Selbstwirksamkeitserwartung ihrer Schüler:innen unterstützen, indem sie beispielsweise durch an die Fähigkeiten der Schüler:innen angepasste Aktivitäten Gelegenheiten für Erfolgserfahrungen bezüglich Aufgaben bzw. Tätigkeiten generieren, die während der weiteren Schullaufbahn relevant sind (Britner & Pajares, 2006). Konkret schlagen Britner und Pajares (2006) vor, dass Naturwissenschaftslehrkräfte in *Middle Schools* ihre Schüler:innen Erfolgserfahrungen bei der Durchführung forschungsorientierter Untersuchungen sammeln lassen sollten, damit diese mit einer gestärkten Selbstwirksamkeitserwartung in die Highschool wechseln, in der naturwissenschaftlicher Unterricht häufig laborbasiert stattfindet. Stellvertretende Erfahrungen als Informationsquelle für die Bildung der Selbstwirksamkeitserwartung könnten im Unterricht unter anderem durch eine überlegte Gruppenzusammensetzung berücksichtigt werden (Britner & Pajares, 2006). Lernende verbal von ihren Fähigkeiten zu überzeugen, könnte unter anderem beispielsweise durch Steigerung des

Durchhaltevermögens oder der Anstrengung zum Erleben von Erfolgserfahrungen beitragen, die wiederum die Selbstwirksamkeitserwartung stärken (Britner & Pajares, 2006). Wie oben bereits angeklungen ist, muss aber darauf geachtet werden, dass die verbale Überzeugungsarbeit glaubwürdig ist (Britner & Pajares, 2006). Bezüglich emotionaler Erregung schlagen Britner und Pajares (2006) zur Förderung der Selbstwirksamkeitserwartung unter anderem vor, dass Lehrkräfte in naturwissenschaftlichen Fächern versuchen könnten, ihren Schüler:innen bei der Kontrolle von Ängsten zu helfen und ihnen deutlich machen könnten, dass negative Erregung nicht unbedingt mit schlechter Leistung einhergeht.

Genau wie in schulischen Lehr-Lernkontexten haben auch in Lehr-Lernkontexten im Studium Selbstwirksamkeitserwartungen einen Einfluss auf die Leistung (Bartimote-Aufflick et al., 2016). Eine auch in der Schulzeit entwickelte Selbstwirksamkeitserwartung kann sich somit auf Leistungen im Studium auswirken. Inwiefern die Förderung der Selbstwirksamkeit im Studium in den Blick genommen wird, könnte zum einen davon abhängen, inwieweit Dozierende eine Studierfähigkeit voraussetzen und gezielte Förderungen von persönlichen Überzeugungen der Studierenden für nötig befinden, zum anderen auch von Gruppengrößen (z. B. Vorlesungen), die zumindest persönliche Rückmeldungen nicht immer zulassen. Dennoch ist es auch im Studium durch bestimmte Lehrmethoden möglich, Selbstwirksamkeitserwartungen von Studierenden zu stärken (Bartimote-Aufflick et al., 2016). Nachfolgend werden Effekte der Selbstwirksamkeitserwartung auf die akademische Leistung im Bereich der Naturwissenschaften anhand empirischer Arbeiten erläutert.

2.5.4 Empirische Befunde zum Zusammenhang von Selbstwirksamkeitserwartung und akademischer Leistung im mathematisch-naturwissenschaftlichen Bereich

Während Erfolgserfahrungen, also vorangegangene Leistungen, wie bereits beschrieben eine wichtige Informationsquelle für die Selbstwirksamkeitserwartung sind, haben Selbstwirksamkeitserwartungen wiederum häufig einen Effekt auf nachfolgende Leistungen. Zahlreiche empirische Studien haben sich der Erforschung des Zusammenhangs von Selbstwirksamkeitserwartung und akademischer Leistung für verschiedene Bereiche gewidmet (Honicke & Broadbent, 2016). Nachstehend wird aufgrund der thematischen Verortung dieser Arbeit auf Studien fokussiert, die den Zusammenhang von Selbstwirksamkeitserwartung und Leistungen im Gebiet der Naturwissenschaften oder Mathematik in den Blick genommen haben. Dieser Zusammenhang wurde sowohl in Studien mit Schüler:innen als auch mit Studierenden untersucht und vielfach belegt.

So wurde eine spezifische Selbstwirksamkeitserwartung von Schüler:innen im Bereich der Naturwissenschaften oder Mathematik als Prädiktor der Leistung identifiziert, die in Form von Leistungstests oder Schulnoten erfasst wurde (z. B. Ardura & Pérez-Bitrián, 2019; Britner & Pajares, 2006; Klassen, 2004; Schöber et al., 2018; Uçar & Sungur, 2017). Höhere Selbstwirksamkeitserwartungen gingen dabei mit besserer Leistung einher als geringer ausgeprägte Selbstwirksamkeitserwartungen. Die Studie von Schöber et al. (2018) ist eine Längsschnittstudie, in der die Selbstwirksamkeitserwartungen in der Domäne Mathematik sowie der Domäne Lesen und die jeweilige Leistung an zwei Messzeitpunkten innerhalb eines Schuljahrs erhoben wurden. Anders als für das Lesen (unter Beachtung des Migrationshintergrunds) wurde für Mathematik kein reziproker Zusammenhang zwischen Selbstwirksamkeitserwartung und Leistung ermittelt, aber ein positiver Effekt der Selbstwirksamkeitserwartung auf die Mathematikleistung (Schöber et al., 2018). Williams und Williams (2010) hatten bei ihrer Modellierung anhand von PISA-Daten von 2003 in 24 von 33 in ihre Studie eingeschlossenen Ländern Hinweise auf einen reziproken Effekt im Bereich der Mathematik gefunden, einschränkend aber auf den querschnittlichen Charakter der Daten und eine ausstehende Prüfung mit längsschnittlichen Daten verwiesen. Eine statistisch signifikante Korrelation einer naturwissenschaftsspezifischen Selbstwirksamkeitserwartung und naturwissenschaftlicher Kompetenz 15-Jähriger haben Schütte et al. (2007) anhand von PISA-Daten von 2006 nachgewiesen, Korrelationen der Selbstwirksamkeitserwartung mit den Noten im Fach Biologie, Chemie und Physik waren schwächer ausgeprägt, aber dennoch statistisch signifikant.

Verschiedene Studien, die Selbstwirksamkeitserwartungen Studierender bezüglich Naturwissenschaften (und Technik) oder Mathematik und deren entsprechende Leistung untersuchten, identifizierten die Selbstwirksamkeitserwartung oder eine in einem bestimmten Zeitraum stattfindende Änderung der Selbstwirksamkeitserwartung als Prädiktor für die Leistung (z. B. DiBenedetto & Bembenuddy, 2013; Lent et al., 1986; Obrentz, 2012; Pajares & Miller, 1995; Partin & Haney, 2012; Zusho et al., 2003). Pajares und Miller (1995) setzten drei Selbstwirksamkeitserwartungssubskalen bezüglich unterschiedlicher mathematischer Fähigkeiten ein und konnten zeigen, dass bestimmte Leistungen bzw. Resultate am besten durch diejenige Selbstwirksamkeitserwartung vorhergesagt wurden, die die beste Passung zu der spezifischen Leistung bzw. dem Resultat hatte. Zum Beispiel sagte die Selbstwirksamkeitserwartung zum Lösen von Mathematikproblemen das tatsächliche Lösen von Mathematikproblemen besser vorher, als dies die Selbstwirksamkeitserwartung zu Erfolg in mathematikbezogenen Kursen oder dem Durchführen von mathematikbezogenen

Aufgaben taten (Pajares & Miller, 1995). Lawson et al. (2007) untersuchten unter anderem eine auf spezifische naturwissenschaftliche Aufgaben bezogene Selbstwirksamkeitserwartung Studierender und konnten am zweiten Messzeitpunkt eines Semesters eines Biologiekurses eine statistisch signifikante Korrelation mit verschiedenen Leistungsmaßen feststellen. Am zweiten Messzeitpunkt korrelierte die Selbstwirksamkeitserwartung auch mit einer *Reasoning Ability* bezüglich des Testens von Hypothesen. Diese war ein starker Prädiktor für die Leistungsmaße, für die diese Fähigkeit benötigt wurde. Die Selbstwirksamkeitserwartung konnte über den Anteil hinaus, der durch *Reasoning Ability* (und die Punktzahl in einem Hochschulzulassungstest) aufgeklärt wurde, nur einen geringen Anteil der Varianz in diesen Leistungsmaßen aufklären.

Der Effekt der Selbstwirksamkeitserwartung auf die akademische Leistung im mathematisch-naturwissenschaftlichen Bereich kann aber auch durch weitere Variablen moderiert werden. Zum Beispiel zeigten Villavicencio und Bernardo (2013) einen moderierenden Effekt von negativen Emotionen (Wut, Angst, Hoffnungslosigkeit und Scham) auf den Effekt einer Selbstwirksamkeitserwartung bezüglich Trigonometrie auf die akademische Leistung. Bei starker Ausprägung der negativen Emotionen wirkte sich die Selbstwirksamkeitserwartung nicht oder kaum auf die Leistung aus, bei geringer Ausprägung der erhobenen negativen Emotionen ging eine höhere Selbstwirksamkeitserwartung hingegen mit besserer Leistung einher. Auch Burns et al. (2021) zeigten, dass Angst den Effekt einer auf naturwissenschaftlichen Unterricht bezogenen Selbstwirksamkeitserwartung von Schüler:innen auf die entsprechende Leistung moderieren kann. Die Selbstwirksamkeitserwartung selbst kann aber auch als Moderator oder Mediator zwischen verschiedenen Variablen und der akademischen Leistung fungieren, wie mehrere Studien zeigten (z. B. Ardura & Galán, 2019; Ardura & Pérez-Bitrián, 2019; S. D. Brown et al., 1989; Kirbulut & Uzuntiryaki-Kondakci, 2019). Ardura und Galán (2019) wiesen zum Beispiel nach, dass eine auf einen Chemie- und Physikkurs bezogene Selbstwirksamkeitserwartung von Schüler:innen als Mediator zwischen verschiedenen Lernansätzen und der akademischen Leistung fungierte. Eine auf Chemie und Physik bezogene Selbstwirksamkeitserwartung vermittelte auch Zusammenhänge zwischen verschiedenen weiteren motivationalen Variablen wie intrinsischer Motivation und der Leistung (Ardura & Pérez-Bitrián, 2019).

Eine weitere Studie, die mit Blick auf die vorliegende Arbeit, in der MER als Charakteristikum naturwissenschaftlicher Kommunikation Hauptthema sind, nicht unerwähnt bleiben soll, ist die Studie von Jian (2017). In der genannten Studie arbeiteten Sechstkläss-

ler:innen mit Text-Bild-Material zu einem Thema aus der Biologie und beantworteten anschließend einen Test mit Fragen zum Inhalt des Materials. Vor der Bearbeitung wurde unter anderem die Selbstwirksamkeitserwartung der Schüler:innen bezüglich des Lesens erfasst. Es zeigte sich, dass Schüler:innen, die hinterher der Gruppe mit hoher Testpunktzahl zugeordnet wurden, eine höhere Selbstwirksamkeitserwartung besaßen als die in der Gruppe mit niedrigerer Testpunktzahl. Eine höhere auf das Lesen bezogene Selbstwirksamkeitserwartung ging also mit einer besseren Leistung beim Arbeiten mit MER-Material in Biologie einher. Allerdings ist zu betonen, dass die Items der Selbstwirksamkeitserwartungsskala lediglich nach Überzeugungen zum Lesen fragten und nicht explizit nach dem Lesen naturwissenschaftlicher oder biologischer Texte. Deshalb ist davon auszugehen, dass die Schüler:innen in die Beantwortung der Items ihre Überzeugungen in Bezug auf das Arbeiten mit Abbildungen, die in naturwissenschaftlichen bzw. biologischen Texten omnipräsent sind und auch im Material der Studie enthalten waren, nicht einbezogen haben. Das heißt, dass die erfasste Selbstwirksamkeitserwartung möglicherweise nur mit dem Arbeiten mit dem textlichen Anteil des Materials zusammenhing.

Die oben aufgeführten Studien haben die Selbstwirksamkeitserwartung häufig mit relativ allgemein gehaltenen Aussagen in Bezug auf einen bestimmten Kurs oder ein Schulfach untersucht (für eine kritische Auseinandersetzung siehe Marsh et al., 2019) oder auch hinsichtlich spezifischerer Anforderungen wie bestimmter Aufgaben aus dem naturwissenschaftlichen oder mathematischen Bereich. Eine Untersuchung einer spezifischen Selbstwirksamkeitserwartung hinsichtlich des Arbeitens mit komplexen, fachspezifischen Abbildungen im Bereich der Biologie und deren Effekt auf eine entsprechende Leistung wurde meines Wissens indes noch nicht berichtet und ist Gegenstand von Studie 2 der vorliegenden Arbeit. Die Bedeutung und das häufige Auftreten derartiger Abbildungen in den Naturwissenschaften und auch spezifisch in einem Biologiestudium oder im Biologieunterricht wurde in den vorstehenden Kapiteln herausgearbeitet. Es ist daher denkbar, dass eine solche Selbstwirksamkeitserwartung in Lehr-Lernkontexten im Bereich der Biologie eine wichtige Rolle spielen könnte.

3 Überblick über die durchgeführten Studien mit ausdifferenzierter Zielsetzung

Die schriftliche Kommunikation in den Naturwissenschaften, gerade auch in der Biologie, ist stark durch den Einsatz von Text-Bild-Kombinationen geprägt. Dies betrifft auch den naturwissenschaftlichen Unterricht, also zum Beispiel den Biologieunterricht in der Schule. Die mit Text kombinierten bildlichen Repräsentationen werden beispielsweise genutzt, um nicht sichtbare oder komplexe Phänomene darzustellen oder dienen zum Beispiel durch die organisierte Darstellung großer Datenmengen und damit einhergehender Sichtbarmachung von Datenmustern der Erkenntnisgewinnung. Die verwendeten bildlichen Darstellungen sind zum Teil fachspezifisch, also weder durch Präsenz im gesellschaftlichen Leben noch durch Präsenz in anderen Schulfächern vertraut, was das Arbeiten mit ihnen erschweren kann. Eine wichtige Eigenschaft von Text-Bild-Material, die auch das Verarbeiten von Informationen aus Text-Bild-Kombinationen bzw. multiplen externen Repräsentationen (MER) beeinflussen kann, ist der Grad an Redundanz, also inwieweit Informationen in Text und Bild identisch sind. Die MER in verschiedenen Medien, mit denen Schüler:innen im Schulunterricht und darüber hinaus sowohl im gesellschaftlichen Leben als auch in einem möglichen Studium in naturwissenschaftlichen Fächern arbeiten (z. B. Schulbücher, Arbeitsblätter, Lehrbücher im Studium, (populär-)wissenschaftliche Artikel) können sich in der Art ihrer bildlichen Repräsentationen und auch im Merkmal Redundanz unterscheiden. Wer eine berufliche Laufbahn bzw. ein Studium in den Naturwissenschaften anstrebt, sollte somit in der Lage sein, Informationen aus fachtypischen MER zu entnehmen und zu integrieren, unabhängig davon, wie die Informationen präsentiert werden. Es sollte deshalb aus wissenschaftspropädeutischer Sicht ein Ziel des naturwissenschaftlichen Fachunterrichts sein, Schüler:innen zu befähigen, mit verschieden gestalteten MER und auch fachspezifischen Abbildungen kompetent zu arbeiten. Dies ist nicht nur für Personen von Vorteil, die eine naturwissenschaftliche Karriere anstreben. Auch für alle anderen können sich entsprechende Fähigkeiten beispielsweise für die gesellschaftliche Teilhabe positiv auswirken. So werden z. B. auch in der Wissenschaftskommunikation zwischen Wissenschaft und Öffentlichkeit häufig Abbildungen eingesetzt und mit Text kombiniert.

In Bezug auf fachspezifische, komplexe Abbildungen in den Naturwissenschaften können verschiedene Faktoren zu Schwierigkeiten bei der Informationsentnahme aus die-

sen beitragen. In dieser Arbeit soll für das Fach Biologie mit der Selbstwirksamkeitserwartung, mit komplexen Abbildungen in der Biologie arbeiten zu können, ein weiterer potentieller Faktor untersucht werden. Bezüglich des Merkmals Redundanz in MER konnten empirische Studien sowohl positive als auch negative Effekte von Redundanz zeigen, haben aber selten redundantes mit nicht-redundantem Text-Bild-Material gleichen Umfangs verglichen und außerdem meist jüngere Schüler:innen als Studienteilnehmende ausgewählt.

In der vorliegenden Arbeit wurde das Arbeiten mit MER zu biologischen Themen in den Blick genommen und dabei auf Personen fokussiert, die die Schullaufbahn mit Erwerb der Hochschulzugangsberechtigung erfolgreich beendet haben. Teilnehmende waren Biologiestudierende im Bachelor. Die Ergebnisse können also zeigen, wie gut Studierende am Anfang ihres Studiums in der Lage sind, mit verschiedenen MER zu arbeiten. Zudem können durch die Ergebnisse Rückschlüsse gezogen werden, inwieweit oben genannte Ziele des schulischen Fachunterrichts erreicht wurden. Die folgenden Aspekte wurden in dem dieser Dissertation zugrunde liegenden Promotionsprojekt untersucht:

1. Welche Testleistung erbringen Biologiestudierende in Abhängigkeit von bezüglich Redundanz in MER unterschiedlich gestaltetem Testmaterial?
2. Wie wirkt sich die Überzeugung hinsichtlich der eigenen Fähigkeit, auch mit komplexen Abbildungen der Biologie arbeiten zu können, auf die Testleistung aus, wenn derartige Abbildungen im Aufgabenmaterial enthalten sind?

Im Folgenden werden die Zielsetzungen der einzelnen in diesem Promotionsprojekt durchgeführten Studien genauer erläutert. Tabelle 3.1 führt Kernpunkte der Studien auf.

Studie 1 (Kapitel 4): Challenges Solving Science Tasks with Text–Picture Combinations Persist beyond Secondary School

Da es aus oben genannten Gründen Ziel schulischer Bildung bzw. des naturwissenschaftlichen Unterrichts sein sollte, Schüler:innen zu befähigen, Informationen aus unterschiedlich gestalteten MER im naturwissenschaftlichen Kontext entnehmen und integrieren zu können, wurde in dieser Studie untersucht, wie gut Personen nach bestandem Abitur tatsächlich in der Lage sind, mit bezüglich Redundanz variiertem Text-Bild-Material in Biologie zu arbeiten. Die Stichprobe war insofern eine Positivselektion, als sie aus Studie-

renden bestand, die sich für ein Studium im Fach Biologie entschieden haben. Die Ergebnisse können (1) Hinweise geben, ob die normative Vorgabe erfüllt wird, dass Schüler:innen nach Beendigung der Schulzeit kompetent mit verschiedenen gestalteten MER arbeiten können und somit auf diesen Aspekt, der auch für eine berufliche Tätigkeit in der Biologie zentral ist, vorbereitet sind, (2) möglichen Förderbedarf für das Arbeiten mit verschiedenen redundanten MER aufdecken, was Bedeutung für den Biologieunterricht und biologiedidaktische Forschung hat, und (3) Erkenntnisse liefern, die möglicherweise zur Sensibilisierung von Lehrkräften oder generell Dozierenden hinsichtlich potentieller Effekte von Redundanz in Lehr- und Testmaterial beitragen können, da davon auszugehen ist, dass die Konstruktion von MER bezüglich Redundanz häufig unbewusst erfolgt.

Zufällig auf drei Experimentalgruppen aufgeteilt bearbeiteten $N = 240$ Biologiestudierende im Bachelor zwei Biologieaufgaben mit offenem Antwortformat. Aus Gründen ökologischer Validität wurden als Aufgaben zwei Abituraufgaben des Fachs Biologie ausgewählt: Eine aus dem Themengebiet Molekulargenetik, die auch fachspezifische, komplexe bildliche Repräsentationen enthielt, die zweite aus den Themengebieten Ökologie und Evolution, die bildliche Repräsentationen enthielt, die auch aus anderen Fächern sowie durch Präsenz im gesellschaftlichen Leben bekannt sind. Das Aufgabenmaterial wurde für drei Gruppen variiert: Eine Gruppe arbeitete mit nicht-redundantem Material, in dem lösungsrelevante Informationen auf Texte und Abbildungen aufgeteilt waren. Die zweite Gruppe arbeitete mit Text-Bild-Kombinationen, bei denen lösungsrelevante Informationen im Text wiederholt wurden, d. h. das Material war teilweise redundant. Um zu schauen, ob die bildlich dargebotenen Informationen für Studierende der zweiten Gruppe einen Mehrwert gegenüber dem reinen Text hatten, arbeitete eine dritte Gruppe mit ausschließlich textlichem Material. In einer Masterarbeit (Klann, 2018), die im Rahmen dieses Promotionsprojekts betreut wurde, wurde das Material der beiden Text-Bild-Bedingungen in einer Eye-Tracking- und Thinking-Aloud-Studie mit sechs Biologiestudierenden eingesetzt, um Erkenntnisse über den Prozess der Aufgabenbearbeitung zu gewinnen.

Studie 2 (Kapitel 5): Fachspezifische, komplexe Abbildungen verarbeiten – (auch) eine Frage der Überzeugung

In den Naturwissenschaften sind fachspezifische Abbildungen ein häufiges Kommunikationsmittel. Auch im öffentlichen Diskurs werden sie mitunter zur Kommunikation von Forschungsergebnissen eingesetzt. So werden in der naturwissenschaftlichen Kommunikation beispielsweise häufig Diagramme mit einer großen Anzahl von Datenpunkten oder Abbildungen mit einer Vielzahl spezifischer grafischer Konventionen genutzt, was sie im Verhältnis zu Abbildungstypen, die durch Präsenz im gesellschaftlichen Leben vertraut sind, komplex erscheinen lässt. Studien konnten Schwierigkeiten bei der Informationsentnahme aus fachspezifischen, komplexen Abbildungen zeigen, zum Teil wurden die Abbildungen sogar übersprungen und stattdessen lediglich Text fokussiert. Wenn Rezipient:innen derartige Abbildungen nicht „lesen“ können, können diese ihr gewinnbringendes Potential nicht entfalten. Warum diese Abbildungen Schwierigkeiten bereiten können, kann unterschiedliche Ursachen haben. Neben etwa mangelndem Wissen über diese Abbildungsarten stellt sich die Frage, ob auch die persönliche Einschätzung der eigenen Fähigkeit, mit entsprechenden Abbildungen arbeiten zu können, einen Einfluss auf den Erfolg beim Arbeiten mit diesen hat. In dieser Studie wurde deshalb die Selbstwirksamkeitserwartung in Bezug auf das Arbeiten mit komplexen Abbildungen in der Biologie in den Blick genommen. Eine Selbstwirksamkeitserwartung bezüglich einer Aufgabe oder Tätigkeit kann sich auf Anstrengung, Widerstandsfähigkeit und Durchhaltevermögen auswirken sowie auf die Wahrscheinlichkeit, eine Aufgabe oder Tätigkeit anzugehen, und letztendlich auch auf die Leistung. Biologiestudierende sollten während ihrer Schulzeit und ihres bisherigen Studiums bezüglich komplexer Abbildungen in der Biologie eine Selbstwirksamkeitserwartung aufgebaut haben, die nun beeinflussen könnte, wie sie an die Bearbeitung von Material herangehen, das solche Abbildungen enthält.

Zufällig auf drei Experimentalgruppen aufgeteilt bearbeiteten $N = 205$ Biologiestudierende im Bachelor zwei Biologieaufgaben mit offenem Antwortformat (s. Studie 1). Eine Aufgabe zum Themenbereich Molekulargenetik enthielt fachspezifische, komplexe Abbildungen. Die zweite Aufgabe, die die Themenbereiche Ökologie und Evolution adressierte, enthielt bildliche Repräsentationen, die auch in anderen Fächern bzw. Fachdisziplinen und im gesellschaftlichen Leben präsent sind. Die Gruppe mit nicht-redundantem Material, in dem lösungsrelevante Informationen auf Texte und Abbildungen aufgeteilt waren, war somit in der Molekulargenetik-Aufgabe darauf angewiesen, Informationen aus komplexen, fachspezifischen Abbildungen entnehmen zu können. Die Gruppe, die mit Text-

Bild-Material arbeitete, in dem lösungsrelevante Informationen im Text wiederholt wurden, war nicht auf Informationen aus Abbildungen angewiesen, genau wie die dritte Gruppe, die nicht mit komplexen Abbildungen konfrontiert wurde, sondern rein textliches Material erhielt. Die Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie wurde erhoben und ein Effekt auf die Testleistung untersucht. Die Ergebnisse können Hinweise geben, ob die untersuchte Selbstwirksamkeitserwartung als eigenständiges Konstrukt angesehen werden kann, und können für die Ableitung von Implikationen für das Unterrichten mit Abbildungen im Fach Biologie dienen.

Tabelle 3.1

Übersicht über die durchgeführten Studien des Promotionsprojekts

	Studie 1	Studie 2
Publikation	Magnus, L., Schütte, K., & Schwanewedel, J. (2020). Challenges solving science tasks with text–picture combinations persist beyond secondary school. <i>Journal of Research on Educational Effectiveness</i> , 13(4), 759–783. https://doi.org/10.1080/19345747.2020.1750744	Magnus, L., Schütte, K., & Schwanewedel, J. (2021). Fachspezifische, komplexe Abbildungen verarbeiten – (auch) eine Frage der Überzeugung. <i>Psychologie in Erziehung und Unterricht</i> , 68. https://doi.org/10.2378/peu2021.art29d
Untersuchungsgegenstand	Arbeiten mit unterschiedlichen multiplen externen Repräsentationen, Redundanz	Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie
Biologische Themenbereiche	Molekulargenetik, Ökologie/Evolution	Molekulargenetik, Ökologie/Evolution
Forschungsfragen/ Hypothesen	Forschungsfrage 1: Wie wirkt sich Redundanz in MER (Text-Bild-Kombinationen mit redundanten im Vergleich zu komplementären Informationen) auf die Testleistung Bachelor-Biologiestudierender in einem Biologietest aus? Forschungsfrage 2: Führen bildliche Repräsentationen, die von Text ergänzt werden, der lösungsrelevante Informationseinheiten aus bildlichen Repräsentationen wiederholt, dazu, dass Bachelor-Biologiestudierende in einem Biologietest besser abschneiden als mit rein textlichem Material?	Hypothese (gekürzt): Die Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie hat positiven Effekt auf Testleistung Studierender, auch unter Kontrolle von fachspezifischen und allgemeinen kognitiven Leistungsmaßen. Effekt wird durch die Experimentalgruppenzugehörigkeit (d. h. Art des Aufgabenmaterials) moderiert. Wenn Studierende zum Lösen der Aufgabe Informationen aus derartigen Abbildungen entnehmen müssen, sollten Selbstwirksamkeitserwartung und Testleistung positiv zusammenhängen. Müssen Studierende zum Lösen derselben Aufgabe keine Informationen aus Abbildungen entnehmen, wird der Zusammenhang nicht erwartet. In dem Maße, wie Studierende, die Text-Bild-Material bekommen, dessen Text alle lösungsrelevanten Informationen enthält, dennoch die Abbildungen zum Lösen hinzuziehen, wird wiederum ein positiver Zusammenhang von Selbstwirksamkeitserwartung und Testleistung erwartet.
Methode	<ul style="list-style-type: none"> • 3 Experimentalgruppen • N = 240 Biologiestudierende im Bachelorstudium • Two-way ANOVA, ANCOVA, Post-hoc-Tests mit Bonferroni-Korrektur 	<ul style="list-style-type: none"> • 3 Experimentalgruppen • N = 205 Biologiestudierende im Bachelorstudium • Multiple Regression, Simple-Slope-Analyse bei Moderationseffekt, Benjamini-Hochberg-Prozedur

4 Studie 1: Challenges solving science tasks with text–picture combinations persist beyond secondary school¹¹

Abstract

Combinations of text and different types of pictures are commonplace in biology as in science in general. The single representations (i.e., text, picture) constituting a text–picture combination may contain redundant or complementary information. The ability to identify and integrate information in different kinds of text–picture combinations is indispensable for engaging in science and is normatively expected to be acquired in school. In this experimental study, which was not preregistered in an independent institutional registry, $N = 240$ undergraduate students worked on 2 constructed-response biology tasks originating from authentic final exams to obtain the higher education entrance qualification. The material carried equivalent information between conditions but consisted of redundant or non-redundant text–picture combinations or only texts. Analysis of variance revealed negative effects of the depictive representations on students' performance, but these effects only occurred in 1 of the 2 tasks: Students in the text-only group reached the highest mean score, those who worked with the non-redundant text–picture combinations the lowest mean score. Implications for science education are discussed.

Keywords: multiple external representations, depictive representations, redundancy, science education, science communication, assessment

¹¹ This is an Accepted Manuscript of an article published by Taylor & Francis in the Journal of Research on Educational Effectiveness on 24 Jun 2020, available at <https://www.tandfonline.com/doi/full/10.1080/19345747.2020.1750744>:

Magnus, L., Schütte, K., & Schwanewedel, J. (2020). Challenges solving science tasks with text–picture combinations persist beyond secondary school. *Journal of Research on Educational Effectiveness*, 13(4), 759–783. <https://doi.org/10.1080/19345747.2020.1750744>. This Open Access article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

4.1 Introduction

Combinations of texts and pictures that together depict a particular subject matter are ubiquitous in everyday life. Such combinations are also widely used in scientific contexts (e.g., Bowen & Roth, 2002; Lemke, 1998), for example a microscopic photo or a line graph and accompanying text. The ability to independently identify and integrate information from text–picture combinations enables an individual to participate in societal discourse and to make reasonable decisions and is thereby an important aim of schooling (e.g., Krajcik & Sutherland, 2010; Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2005). More specifically, students intending to enter a science career who enroll in science majors in tertiary education need to identify and integrate relevant information in scientific text–picture combinations—regardless of how the information is presented—as part of their jobs and their studies. Identifying and integrating information in text–picture combinations is therefore an important receptive part of communication in science. Despite the prevalence of these combinations in science, many students—even students who have graduated from secondary school—struggle with them, especially with pictures in text–picture combinations (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010). Neglecting pictures is especially problematic when the information conveyed by the picture is not repeated in the text.

When presented in combination text and picture may contain complementary or equivalent information, the combinations thereby differ in their redundancy. In the context of learning, educational psychologists have established seemingly contradictory effects of redundancy: On the one hand, Mayer (2001) shows that learning from text and picture is superior to learning from text alone. On the other hand, redundancy of text and picture may increase cognitive load that interferes with learning due to limited working memory capacity (Sweller, 2005). Similar mechanisms are expected to operate when a task at hand needs to be comprehended. Overall, research has yet to establish when redundancy helps and when it hinders. Specifically, it is not yet clear how redundant and non-redundant text–picture combinations in constructed-response items affect students' ability to solve these items. Testing this with university students who have completed school sheds light on how well school prepares students to process text–picture combinations that differ in their redundancy. They are normatively expected to be able to process all kinds of text–picture combinations and deficient science communication skills likely impede students' academic success in tertiary education.

We therefore compared students' test performance in constructed-response science items when working with non-redundant or redundant text–picture combinations or text-only material. This comparison provides information on the extent to which students use or possibly disregard pictorial and textual information as a function of the text–picture combinations' redundancy. In addition, it demonstrates to what extent redundant pictures added to text impact test performance. Our experiment thus gauges the effectiveness of current teaching with regards to students' receptive science communication skills and its results might inform new policy.

4.2 Theoretical Framework

4.2.1 Types of External Representations and their Importance in Biology Contexts

Representations are objects that stand for something else (Schnotz, 2002). External representations (ERs) are divided into descriptive representations and depictions (Schnotz, 2002). Descriptive representations consist of symbols characterized by an arbitrary structure and conventions determine which object a symbol represents; typical examples are texts and mathematical equations (Schnotz, 2002). Depictive representations, however, consist of iconic signs. Whereas realistic pictures, such as photographs and drawings, are similar to the represented object, logical pictures, for example graphs and diagrams, are more abstract (Schnotz, 2002). Combinations of two or more single representations are referred to as multiple external representations (MERs).

In the discipline of biology, it is common to combine different types of ERs (Tsui & Treagust, 2013), for example texts and pictures; that is, MERs are highly prevalent in biology. The use of depictive representations is highly important to communicate within the discipline of biology, for example when representing biological phenomena on the macro or micro level (Eilam & Gilbert, 2014), such as a plant and its different cell types. Pictures can visualize biological phenomena that cannot be seen by the naked eye (e.g., at the submicroscopic level). Logical pictures allow scientists to economically summarize and present large amounts of data (Roth, Bowen, & McGinn, 1999). In sum, depictive representations added to text are indispensable to discover and explain findings in the discipline of biology and thus omnipresent in scientific biology publications and course material of biology studies at universities (Arsenault, Smith, & Beauchamp, 2006; Bowen & Roth,

1998; Roth et al., 1999). Therefore, students—especially those students aspiring to a science career—should have learnt how to process all kinds of text–picture combinations by the end of secondary school.

4.2.2 Processing MERs

When trying to understand an MER, individuals construct an internal mental model in their cognitive system. According to the integrated model of text and picture comprehension (Schnotz, 2014; Schnotz & Bannert, 2003), a sensory register conveys the information of external representations (i.e., texts and pictures) to the working memory. Here, text is first processed in a descriptive subsystem and then in a depictive one; the opposite is assumed to apply to pictures. Individuals also integrate prior knowledge stored in their long-term memory when constructing an internal mental model (Schnotz, 2014).

According to Mayer’s multimedia principle, an MER constituted by text and picture promotes better comprehension than text-only learning material (Mayer, 2001). This is in line with Paivio’s dual coding theory which also assumes two different cognitive subsystems for verbal information such as texts on the one hand, and nonverbal information such as depictive representations on the other (Paivio, 1986)—provided that students access all types of representation, MERs constituted by text and picture would thus result in both subsystems being engaged in their processing. As an important benefit of MERs, Ainsworth (2006) stated that in cases where inserting all information in one single representation would be too complicated, complementary information could be conveyed by more than one single representation. Empirically, undergraduate students working with text–picture combinations on three tasks outperformed students provided only with textual material ($0.57 \leq d \leq 0.80$; Eilam & Poyas, 2008).

However, not every depictive representation added to the text helps identify and process information. For instance, it is assumed that task-inappropriate types of depictive representations may interfere with constructing a task-appropriate internal mental model (Schnotz & Bannert, 2003). Moreover, it appears students do not fully use pictures’ potential: For instance, Cromley et al. (2010) showed that undergraduate students working with texts and pictures dealing with biology topics often just skimmed the pictures or skipped over them. But students who sufficiently focused on pictures employed a higher proportion of high-level cognitive activities (Cromley et al., 2010). Not paying sufficient attention to pictures will be particularly problematic when the depictive information is not repeated in

the text; no adequate mental model can be constructed due to the lack of depictive information.

4.2.3 Redundancy in MERs

As already mentioned, the representations constituting an MER may offer complementary or equivalent information; therefore, MERs differ in terms of redundancy and students encounter somewhat differing demands when processing MERs. Ainsworth (1999) distinguishes two subclasses of complementary MERs: those consisting of single representations with totally different information and MERs consisting of single representations with partially shared information. However, completely redundant MERs also exist. Teachers, test developers, or researchers presumably do not as a rule consciously apply design principles in terms of redundancy when constructing MERs. To advise which principles to adhere to, the effects of redundancy need to be better understood; respective principles would also potentially enhance the practical utility of MERs in learning contexts.

Sweller's redundancy principle in multimedia learning draws attention to the fact that redundant material unnecessarily enhances working memory load and could thereby compromise information processing and transfer to long-term memory (Sweller, 2005). For novel information to be stored in long-term memory it first needs to be processed in the working memory; but working memory capacity is limited. According to Mayer's multimedia principle adding pictures to text in instructional material has positive effects compared to presenting text alone (Mayer, 2001). He summarized nine studies that demonstrated this positive effect regarding students' test performance after studying instructional material on the functional principle of, for example, pumps, brakes, or lightning, with a moderate median effect size of 0.67 for the retention of steps in the presented process and a large median effect size of 1.50 regarding problem-solving transfer. Mayer's cognitive theory of multimedia learning also includes considerations on cognitive load. For example, redundant texts added to narrations and pictures should be avoided to reduce cognitive load caused by text–picture material (Mayer, 2001).

Several experimental studies revealed positive effects of redundancy in MERs. Additional redundant pictures in the stem of multiple-choice science items reduced item difficulty as compared to text-only items when the item stem presented information essential for solving items ($d = 0.53$; Lindner, Ihme, Saß, & Köller, 2018). Saß, Wittwer, Senkbeil,

and Köller (2012) also observed positive effects on students' test performance when combining redundant pictures in the stem, in the answer options, or in both with texts in multiple-choice items ($\eta^2 = .03$ and $.04$ corresponding to d s of approximately 0.35 and 0.41). A study on university students regarding comprehension of a pulley system's kinematics showed that students who worked with redundant text–picture material outperformed students who worked with only text or pictures ($d = 1.69$ and 0.98 , respectively; Hegarty & Just, 1993). However, it should be noted that only 47 students participated in that study.

On the other hand, Bobis, Sweller, and Cooper (1993) observed a negative effect of redundancy in instructional material on paper-folding activities. Performance of students provided only with diagrammatic instructions was superior to that of students who received additional redundant texts; the authors explained this difference through an increased cognitive load in the redundancy group. The performance of students who only received text was not significantly different from those in the redundancy condition. Similarly, Pociask and Morrison (2008) observed a negative effect of redundant text–picture material compared to non-redundant text–picture material presented during the learning phase when 41 university students worked on instructional units regarding orthopedic physical therapy with an effect size of approximately $d = 0.81$.

These studies have several characteristics in common, but some of them also differ in their requirements, which might explain the different findings. The studies compared the effect of redundant text–picture combinations (MERs) with single types of representations (either texts or pictures). Consequently, the amount of test material differed between groups, allowing for other explanations of mean test score differences than the MERs' redundancy. In their low-powered study, Pociask and Morrison (2008) compared redundant and non-redundant MERs, but due to eliminating redundant features the amount of test material also differed between the groups. Saß et al. (2012) and Lindner et al. (2018) placed pictures into multiple-choice items and generated redundancy based on the text. This kind of redundant text–picture combinations is specific to assessment and does not reflect MERs in the scientific discourse; it therefore lacks ecological validity in this science-related context. The presented pictures were less extensive than, for example, those of Pociask and Morrison (2008). Many previous studies used multiple-choice items or practical tasks rather than constructed-response items, potentially placing different demands on working memory. The material in previous research combined a certain type of depictive representation with text, which does not reflect the combination of diverse depictive representations in the authentic scientific discourse. Moreover, in most previous studies, participants were

primary or lower secondary school students (but see Hegarty & Just, 1993; Pociask & Morrison, 2008). Yet, all studies suggest that students did allocate some of their attention to pictures when they were presented with text–picture combinations.

4.2.4 Purpose of This Study

Since MERs are an integral part of communicating in science, the ability to identify and integrate information in different kinds of MERs is indispensable for engaging in science. An aim of science education in school is for all students to acquire this ability. Students who want to enter a career in science are especially required to process MERs, regardless of how the information is presented. In terms of MERs' redundancy, this means that students should be able to identify and integrate information presented in redundant or non-redundant MERs after finishing upper secondary school.

In this experimental study, we therefore analyzed undergraduate students' test performance on constructed-response biology tasks consisting of redundant or non-redundant MERs with equal amounts of material across conditions.¹² We tested undergraduate students because they have already finished upper secondary school and the results give an insight into how well science education in school enabled them to work with different MERs in a biological context. Unlike in the condition with redundant MERs, students working with non-redundant MERs depended on identifying and integrating depictive information to solve the tasks. To gain further insights on how students presented with redundant MERs draw upon the information conveyed by the depictive representations as opposed to relying on the text, we added a text-only group (i.e., a group containing no depictive representations). This third experimental condition allowed us to examine whether redundant depictive representations actually help the students identify task-relevant information in comparison to text alone. Comparisons between experimental conditions allow for inferences about students' utilization of the material and thus further our understanding of how students process multiple external representations. Such knowledge is necessary to deduce how best to foster students' ability to independently identify and integrate information from MERs. We used test material based on final exam tasks to obtain the higher education entrance qualification, the German Abitur, to relate to authentic biology education. In this study, we investigated the following research questions (RQs):

¹² This study was not preregistered with an analysis plan in an independent institutional registry. However, inquiries about study materials may be directed to the first author. Furthermore, the raw data supporting the conclusion of this manuscript will be made available by the authors to any qualified researcher.

RQ1: How does redundancy in MERs (text–picture combinations with redundant as compared with complementary information) affect undergraduate biology students' performance in a biology test?

RQ2: Do depictive representations accompanied by text reiterating task-relevant depictive information units lead to undergraduate biology students performing better in a biology test as compared with text-only material?¹³

4.3 Method

4.3.1 Participants

In light of previous research, we performed an a priori power analysis for detecting an effect of medium size ($f = 0.25$) by means of a one-way analysis of variance (ANOVA)¹⁴ with a power level of 0.95 (Faul, Erdfelder, Lang, & Buchner, 2007). Our sample consisted of 245 university bachelor students majoring in biology in Germany, falling only seven students short of the required sample size of $N = 252$. The dependent variable's score of five students was missing and these students were excluded from the sample. Thus, the remaining sample size is $N = 240$ ($M = 21.8$ years, $SD = 3.4$); 70.0% of the students were female.

4.3.2 Procedure

All data collection was conducted by the first author. At the beginning of the session, students were informed about the scope of the study and that it was part of a dissertation project. They were informed that all data collection would be completely anonymous and that they would face no adverse consequences by declining or withdrawing their participation from the study at any time. Further, students were informed that participation in the study was neither a course requirement nor an opportunity for extra credit and that test performance would have no effect on their grades. Potential participants were made aware that they could leave prior to study conduction or at a later time, which a few students did.

¹³ In authentic scientific communication, redundancy results from information of the depictive representation being repeated in the text. Pictures are therefore considered the primary material.

¹⁴ The targeted sample size could not be obtained within a single classroom. We thank an anonymous reviewer for suggesting to account for this data structure by including classroom as a second factor in our analytic approach, thereby potentially improving its statistical power.

The willingness of the remaining students to participate was implied through survey completion. Dispensing with consent in written form obviated recording participants' names. In a paper-and-pencil test, all students independently worked on two biology tasks in a constructed-response format (described in more detail in the following subsection). The total time allowed for both tasks was 45 minutes. The administrator terminated this part of the session after the designated test duration. To avoid effects of task order, we presented the two biology tasks in a balanced design. For each task, students noted the time when they began and completed it. Within classes, the students were randomly assigned to one of three experimental groups—class-level characteristics shared by students attending the same class were thus distributed equally across experimental groups: In the material of the no-redundancy group, all information relevant for solving the task was distributed between texts and depictive material ($n = 79$), the redundancy group received material in which all relevant information of depictive material was repeated in the text ($n = 79$), and material of the text-only group ($n = 82$) consisted solely of texts, which was nearly the same as the text of the redundancy group (for details see subsection *Material*). The tasks were phrased identically across groups and made no mention of the different experimental conditions. This also means that participants in the two conditions containing depictive material as well as text were not explicitly instructed that the pictures might contain task-relevant information. The test imposed no requirements on students beyond those that are commonplace for biology majors. In a second step, students answered a questionnaire consisting of covariates and questions regarding demographic data. Students could take part in a draw to win diverse vouchers with a value between 5 € and 50 €. Moreover, we rewarded participation with sweets. At the end of the study, students were fully debriefed.

4.3.3 Material

The tasks were based on Abitur exam tasks from 2008 (Task A) and 2009 (Task B) from the German federal state North Rhine-Westphalia (Brixius, Jannan, & Kunze, 2012).¹⁵ The Abitur is the higher education entrance qualification in Germany, the school leaving certificate obtained at the end of secondary school. The reasons for using Abitur exam tasks were twofold: Their authenticity ensures that the tasks do not require abilities beyond those

¹⁵ Although the Abitur examinations in Germany are not centralized at the national level, defined uniform examination requirements mandate comparability between the requirements of Abitur exam tasks in biology in different German federal states (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2004).

which students are normatively expected to possess at the end of secondary school and solving the tasks requires students to analyze and discuss complex biological topics while working with MERs. We slightly adapted the task material according to conditions; the task material in the no-redundancy group was almost the original task material from the Abitur exam tasks. While systematically varying the task material between the three experimental groups, we kept the amount of textual material, content, and the question to be answered in an open-ended format constant.

In Task A, students were asked to explain the molecular genetic cause of a mitochondrial disease and its disease pattern (Brixius et al., 2012, p. 2008-13, 2008-15–2008-17). The students received textual material and—depending on the group they were assigned to—also depictive representations to solve the task. Task A contained three texts and the depictive representations of Task A comprised two biology-specific schematic drawings and a graph. For example, one of the schematic drawings showed the mitochondrial DNA of humans. We present a demonstration example, which closely resembles this particular schematic drawing and accompanying text of the actual task (material) of Task A to illustrate how we created redundancy (Figures 4.3.1–4.3.3); the depictive representations were identical between the two conditions that provided text–picture combinations.¹⁶ To be able to solve Task A, the specific information unit which had to be taken from this particular figure is the base position for the coding of a particular tRNA. This is therefore a depictive information unit. In the no-redundancy group, this information unit was exclusively provided by the figure (see blue marking in the picture of Figure 4.3.1). In the redundancy group, the respective base position was also given in the text (see blue markings in the picture and the text of Figure 4.3.2). In the text-only group, this information unit was exclusively provided by the text; no figures were included in the material (see blue marking in Figure 4.3.3). The text section associated with this particular figure also contained task-relevant information that was only in the text in all three groups (see red markings in Figures 4.3.1–4.3.3). We refer to them as descriptive information units. The second biology-specific schematic of Task A showed a section of the genome of a healthy person and a person suffering from the genetic disease depicted as sequences of letters representing nucleobases of the mitochondrial DNA. The third depictive representation of Task A was a

¹⁶ Copyright issues prevent us from reproducing our actual task materials. Therefore, we developed a demonstration example which is merely intended to show how we created redundancy. In contrast to the actual task, which addresses a hereditary disease, the implied disease of the demonstration example is fictitious and not meant to be biologically correct. Inquiries about intervention materials should be directed to the first author.

scatter plot with 10 data points and a compensation curve. Because in the redundancy group the relevant depictive information was repeated in the text, we had to slightly extend the length of the text in the no-redundancy group to keep the text length of these two groups constant. Text length's equivalence was achieved by content-neutral extension. The text-only group got all information in textual form; we therefore refer to information units that are conveyed by depictive representations in the other two conditions as originally depictive information.

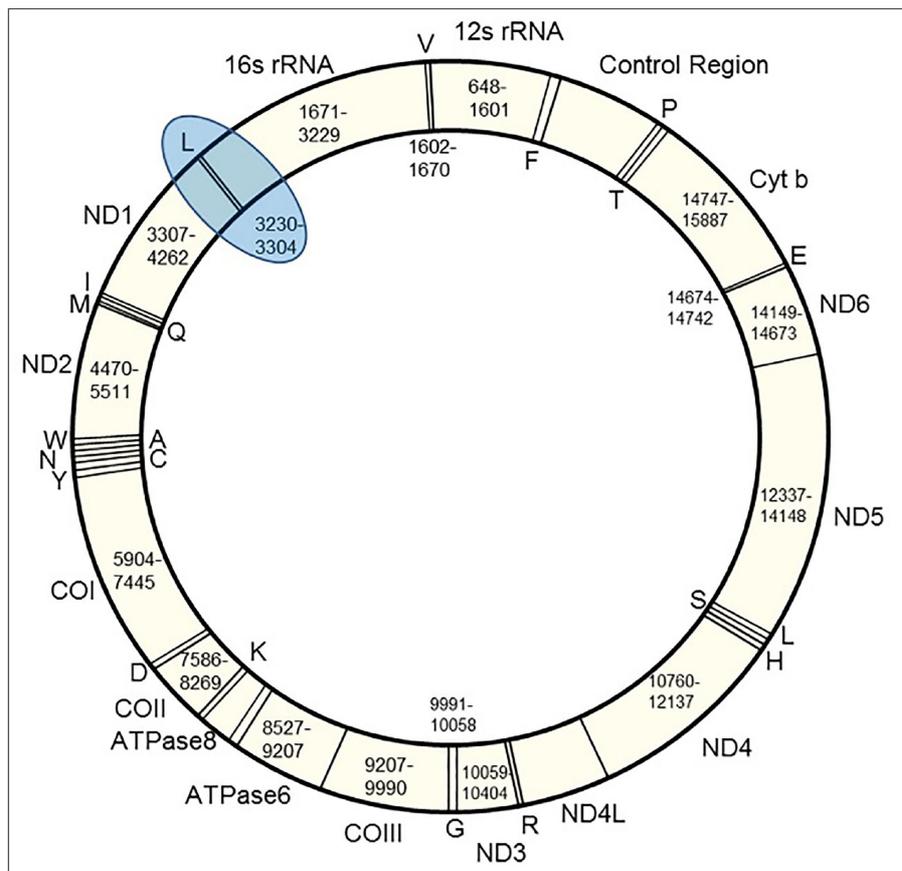


Figure 1. Schematic diagram of human mitochondrial DNA.

The here presented ring-shaped mitochondrial DNA encodes genes for several tRNAs, for example the tRNA^{valine} (which is abbreviated as “V”) and the tRNA^{leucine} (which is abbreviated as “L”). In addition, other areas of the mitochondrial DNA encode the subunits of ribosomes and certain proteins, for example ATPase8 and ATPase6. These are both subunits of the ATP synthase, which is involved in supplying an energy carrier.

Figure 4.3.1. Exemplary test material closely resembling parts of the actual material for the no-redundancy group in Task A. Textual information that was required to solve the task is highlighted in red, relevant depictive information is highlighted in blue. No highlights were included in the material of the actual task provided to the students.

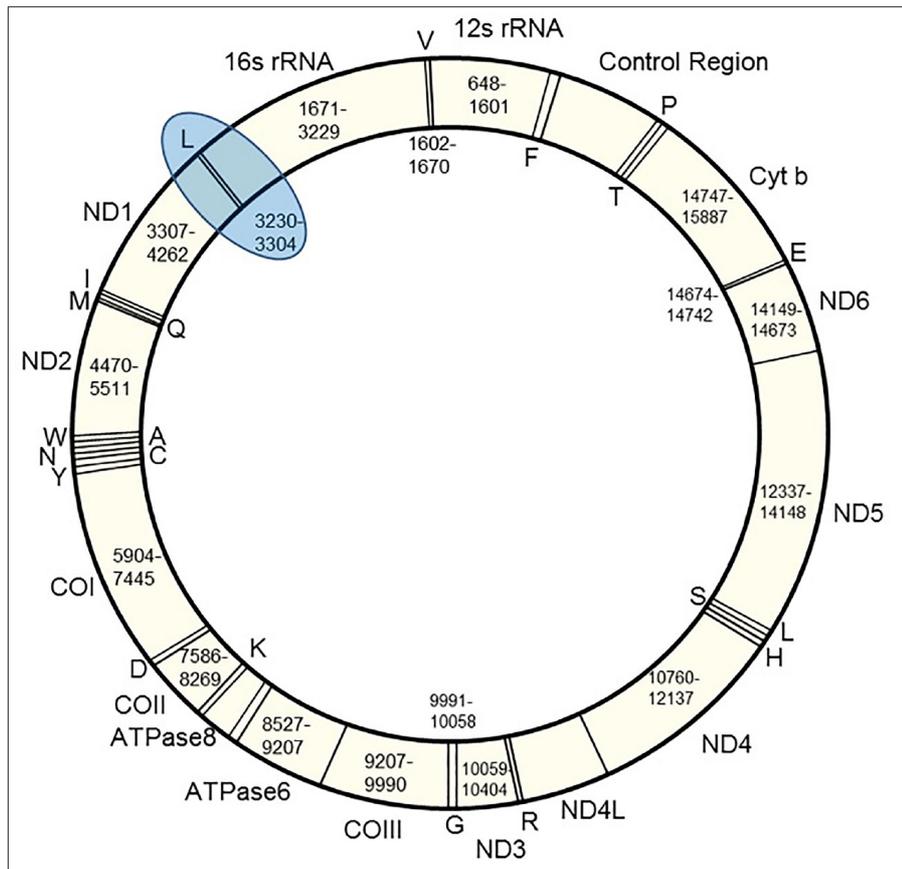


Figure 1. Schematic diagram of human mitochondrial DNA.

The presented ring-shaped mitochondrial DNA encodes genes for tRNAs, for example tRNA^{valine} (abbreviated as “V”) and tRNA^{leucine} (abbreviated as “L”). The tRNA transporting the amino acid L is encoded at position 3230-3304. Other areas of mitochondrial DNA encode subunits of ribosomes and certain proteins, for example ATPase8 and ATPase6, both subunits of ATP synthase, which is involved in supplying an energy carrier.

Figure 4.3.2. Exemplary test material closely resembling parts of the actual material for the redundancy group in Task A. Textual information that was required to solve the task is highlighted in red, relevant (originally) depictive information is highlighted in blue. No highlights were included in the material of the actual task provided to the students.

In the ring-shaped mitochondrial DNA, genes for tRNAs are encoded, for example tRNA^{valine} (abbreviated as “V”) and tRNA^{leucine} (abbreviated as “L”). The tRNA transporting the amino acid L is encoded at position 3230-3304. Other areas of mitochondrial DNA encode subunits of ribosomes and certain proteins, for example ATPase8 and ATPase6, both subunits of ATP synthase, which is involved in supplying an energy carrier.

Figure 4.3.3. Exemplary test material closely resembling parts of the actual material for the text-only group in Task A. Textual information that was required to solve the task is highlighted in red, relevant originally depictive information is highlighted in blue. No highlights were included in the material of the actual task provided to the students.

In contrast to Task A, Task B concerns evolution and ecology rather than genetics. In Task B, students had to explain the specific adaptations of fire salamander larvae to a certain woodland’s environmental factors (Brixius et al., 2012, p. 2009-1–2009-3). Task B’s material contained two texts and the depictive representations of Task B comprised a 2×2 table containing four values, a column chart containing two columns and a scatter plot with 13 data points. The information density of these depictive representations was thus lower than that of figures of Task A. Furthermore, the depictive representations of Task B were less biology-specific than two figures of Task A and it is likely that students knew this type of depictive representations from various disciplines. Overall, the specific depictive representations in Task B might make solving it easier and therefore more likely solved than Task A. Adapting the material to experimental conditions followed the same procedure as for Task A.

4.3.4 Scoring

A category scheme was developed for each task to score students’ responses based on the suggested solutions provided in association with the tasks (Brixius et al., 2012). We identified information units relevant for solving the task both in the figures and in the texts of the original task material. Students could score a maximum of nine points in each task. Depending on the group students were assigned to, they may only have been able to retrieve relevant information units from the text (see Section *Material*). Identification of a task-relevant element was not only scored when the information was named verbatim, but also when students’ answers implied that the information unit must have been extracted.¹⁷ If none of this was the case, students received no point for the information unit in question.

¹⁷ The category schemes can be requested from the first author.

Studie 1

The first author rated all data according to the category schemes. To ensure the objectivity of the judgment, 25% randomly selected student responses for each task were independently rated by a second rater. We then determined Cohen's kappa as a coefficient to measure the degree of raters' agreement (Cohen, 1960). Inter-rater reliability was satisfactory ranging from Cohen's $\kappa = .62$ to $\kappa = .91$ for the nine information units of Task A and Cohen's $\kappa = .64$ to $\kappa = 1.0$ for the nine information units of Task B. This strength of agreement is substantial ($\kappa = .61$ to $.80$) to almost perfect ($\kappa = .81$ to 1.0) following the classification of Landis and Koch (1977). Rating of Task B responses revealed one low kappa value ($\kappa = .42$), although raters' agreement was high (89.1%)—a paradox known to occur when assignments to a category's expressions are very unequal (Feinstein & Cicchetti, 1990). The scoring was based on students' utilization of task-relevant elements of the materials; we refer to those elements as relevant information units. The category scheme for the demonstration example is shown in Supplementary Material Table 4.9.1. In each of the two actual tasks, students were required to use a total of five information units from the figures to solve the task, and four information units from the text. Thus, students could score a maximum of nine points in each task. Additionally, we calculated subscores for each task considering the five depictive and the four descriptive information units separately to investigate whether the different types of representations differentially drive observed effects. Due to our test design, redundancy group students got all depictive information units also in textual form, and students in the text-only group worked only with text. The information units underlying the subscore for depictive information units were thus in fact differentially represented across groups.

4.3.5 Control Measures

4.3.5.1 Cognitive Abilities

Cognitive abilities of students were assessed using the nonverbal scale N2 of the KFT 4-12 + R (Heller & Perleth, 2000), which is a German version of the Cognitive Abilities Tests of Thorndike and Hagen (1971). This scale consists of 25 items, each presenting a pair of meaningfully matching figures, a third figure, and five answer options. Students had to choose the one figure among the answer options which matches the third figure like the presented second figure matches the first. The items were summed up to form a composite score with an acceptable internal consistency, Cronbach's $\alpha = .78$.

4.3.5.2 Working Memory Capacity

Students' working memory capacity was measured using an updating task based on the spatial-figural updating task of Wilhelm, Hildebrandt, and Oberauer (2013). In each trial of this test, between 2 and 5 differently colored rectangles were repeatedly presented within a 3×3 matrix. All colors appearing in a trial were presented together at the beginning of this trial. Then single color rectangles appeared at different positions within the matrix for a number of times unknown to the students. The order of rectangles was generated at random but was kept constant for all tested classes. Students had to remember the last position of each color. Unlike Wilhelm and colleagues who presented the test to each single participant on a computer, we presented it on a large screen to all students of a class. The test consists of 11 trials; an overall score was obtained by averaging the proportion of correctly remembered color positions across trials. Internal consistency was measured using trial scores and showed a poor value of Cronbach's $\alpha = .52$.

4.3.5.3 Additional Covariates

At the end of the session, we asked for students' secondary school leaving certificate grade point average (GPA). In Germany, passing grades range between 1 and 4, whereby 1 is the highest grade. Students also reported their last biology grade in school, their sex, and their age. Furthermore, they indicated whether they had taken an Abitur exam in biology¹⁸ and whether they already knew one or both tasks.

4.3.6 Data Analysis

We calculated ANOVAs and analyses of covariance (ANCOVAs) to answer our research questions. We conducted all analyses with IBM SPSS statistics (version 24, IBM Corporation, 2016). Sampling students from different classes to obtain a sufficiently large sample size means that individual observations were nested in classes; random assignment to groups prevented confounding class and group. In addition to the fixed effect of group, we therefore modeled a fixed effect of class, resulting in a two-way ANOVA. We followed up on significant main effects of the factor experimental group by conducting multiple comparisons: Bonferroni corrected post hoc tests answered our main research questions,

¹⁸ Students in upper secondary school in Germany take different courses having a basic and a higher level of requirements; they take final examinations in only 4–5 subjects, depending on the federal state.

whether text–picture combinations differentially affected students’ task performance when they contained complementary or redundant information (RQ1) and whether depictive representations accompanied by redundant text improved performance as compared with text-only material (RQ2). To be comprehensive, we also report the Bonferroni corrected post hoc tests between the no-redundancy group and the text-only group.

4.4 Results

4.4.1 Baseline Equivalence

Two-way ANOVAs with experimental group and classroom as fixed effects were conducted to ascertain that randomly assigning students to the no-redundancy, the redundancy, or the text-only group had not led to significant differences between the groups with regard to cognitive abilities, working memory capacity, the last biology grade in school, and students’ GPA of the school leaving certificate. Participants in the three groups did not significantly differ with regard to these variables (cf. Table 4.4.1). Working memory was not included in further analyses because of our measure’s low reliability. Moreover, the different experimental conditions did not result in statistically significant differences regarding time-on-task (Table 4.4.1). We further established that the experimental groups did not differ in terms of students’ sex, whether they had taken an Abitur exam in biology and whether they knew at least one task by conducting Pearson’s chi-square tests; neither test was statistically significant, $\chi^2(2) = 0.06, p = .971$; $\chi^2(2) = 0.32, p = .852$; $\chi^2(2) = 0.82, p = .662$, respectively.

Table 4.4.1
Descriptive statistics and ANOVA results for the covariates

Variable	Effect	No-redundancy group		Redundancy group		Text-only group		F	df	p
		M	(SD)	M	(SD)	M	(SD)			
Cognitive abilities	Experimental group	14.60	(4.07)	15.27	(4.71)	14.68	(4.40)	0.44	2, 225	.646
	Classroom							5.32	3, 225	.001
Working memory capacity	Experimental group × Classroom							0.95	6, 225	.462
	Experimental group	.75	(.11)	.73	(.12)	.74	(.11)	0.19	2, 223	.828
Last biology grade ^a	Classroom							1.40	3, 223	.244
	Experimental group × Classroom							0.62	6, 223	.711
Grade point average of the secondary school leaving certificate ^a	Experimental group	1.67	(0.72)	1.76	(0.68)	1.79	(0.64)	1.13	2, 197	.325
	Classroom							19.17	3, 197	<.001
Time-on-task ^b (Task A) (min)	Experimental group × Classroom							1.01	6, 197	.422
	Experimental group	2.00	(0.53)	2.10	(0.58)	2.05	(0.52)	0.86	2, 223	.424
Time-on-task ^b (Task B) (min)	Classroom							26.41	3, 223	<.001
	Experimental group × Classroom							0.69	6, 223	.655
Time-on-task ^b (Task B) (min)	Experimental group	16.18	(6.08)	16.93	(5.31)	16.61	(6.60)	0.10	2, 209	.905
	Classroom							17.01	3, 209	<.001
Time-on-task ^b (Task B) (min)	Experimental group × Classroom							0.77	6, 209	.591
	Experimental group	15.88	(6.08)	16.47	(5.87)	15.08	(5.49)	1.41	2, 215	.247
Time-on-task ^b (Task B) (min)	Classroom							8.84	3, 215	<.001
	Experimental group × Classroom							0.99	6, 215	.433

^aIn the German grading system, 1 is the highest grade and 6 the lowest. The secondary school grade point average of the leaving certificate ranges between 1 and 4. ^bTime-on-task was calculated by subtracting the times recorded by the students. min = minutes.

Studie 1

4.4.2 Effects of Group on Biology Test Performance

4.4.2.1 Effects of Group by Task

We calculated separate two-way ANOVAs for the two biology tasks to investigate whether the group, more specifically, the kind of material students worked with, had an effect on students' performance while accounting for the fact that students had been drawn from different classes. Results for Task A revealed that there was a statistically significant effect of the group, $F(2, 228) = 23.32, p < .001, \eta_p^2 = .170$. Post hoc tests with Bonferroni correction showed that the three mean test scores were significantly different from each other: Students who had to identify relevant information in depictive and textual representations (no-redundancy group; $M = 3.44, SD = 2.49$) reached a significantly lower test score than students who worked exclusively with texts (text-only group; $M = 6.24, SD = 2.44$) and students who got texts and additional pictures containing redundant information (redundancy group; $M = 4.68, SD = 2.86$) with $p < .001, d = 1.13$, and $p = .005, d = 0.46$, respectively; students of the text-only group also outperformed students of the redundancy group, $p < .001, d = 0.59$ (for detailed distribution of the data, see Supplementary Material S1). In contrast, a two-way ANOVA for Task B showed no statistically significant differences between the mean scores of the no-redundancy group ($M = 5.90, SD = 2.13$), the redundancy group ($M = 6.15, SD = 1.63$), and the text-only group ($M = 5.95, SD = 1.60$), $F < 1, p = .669, \eta_p^2 = .004$. The main effect of classroom was statistically significant for both tasks signifying mean level differences between classrooms, $F(3, 228) = 11.51, p < .001, \eta_p^2 = .132$ for Task A, and $F(3, 228) = 8.60, p < .001, \eta_p^2 = .102$ for Task B. The interactions of experimental group and classroom were, however, not statistically significant, $F(6, 228) = 1.35, p = .236, \eta_p^2 = .034$ for Task A and $F(6, 228) = 1.97, p = .071, \eta_p^2 = .049$ for Task B.

4.4.2.2 Effects of Group by Task and Type of Representation

Next, we considered the depictive versus textual representations of each task separately to elucidate whether information conveyed by these representations were differentially driving the effect. We conducted a set of two-way ANOVAs and post hoc tests with Bonferroni correction (Figure 4.4.1). For both depictive and textual information in Task A, the result pattern was comparable to the effect of the experimental group observed for the entire task, but different in effect size, $F(2, 228) = 29.27, p < .001, \eta_p^2 = .204$ and $F(2, 228) = 13.55, p < .001, \eta_p^2 = .106$, respectively (for post hoc test results, see

Supplementary Material Table 4.9.2). The difference between the no-redundancy group and the redundancy group examining textual information units was, however, no longer statistically significant (Figure 4.4.1). In contrast to results for the entire Task B, the main effect of group was statistically significant for the depictive information, $F(2, 228) = 4.20$, $p = .016$, $\eta_p^2 = .036$; yet, no post hoc test with Bonferroni correction reached conventional levels of statistical significance. The mean subtest scores based on the textual information units of the no-redundancy group and the text-only group were, however, statistically significantly different, resulting in a statistically significant main effect, $F(2, 228) = 3.14$, $p = .045$, $\eta_p^2 = .027$. The remaining simple comparisons for textual information units in Task B were not statistically significant (cf. Supplementary Material Table 4.9.2).

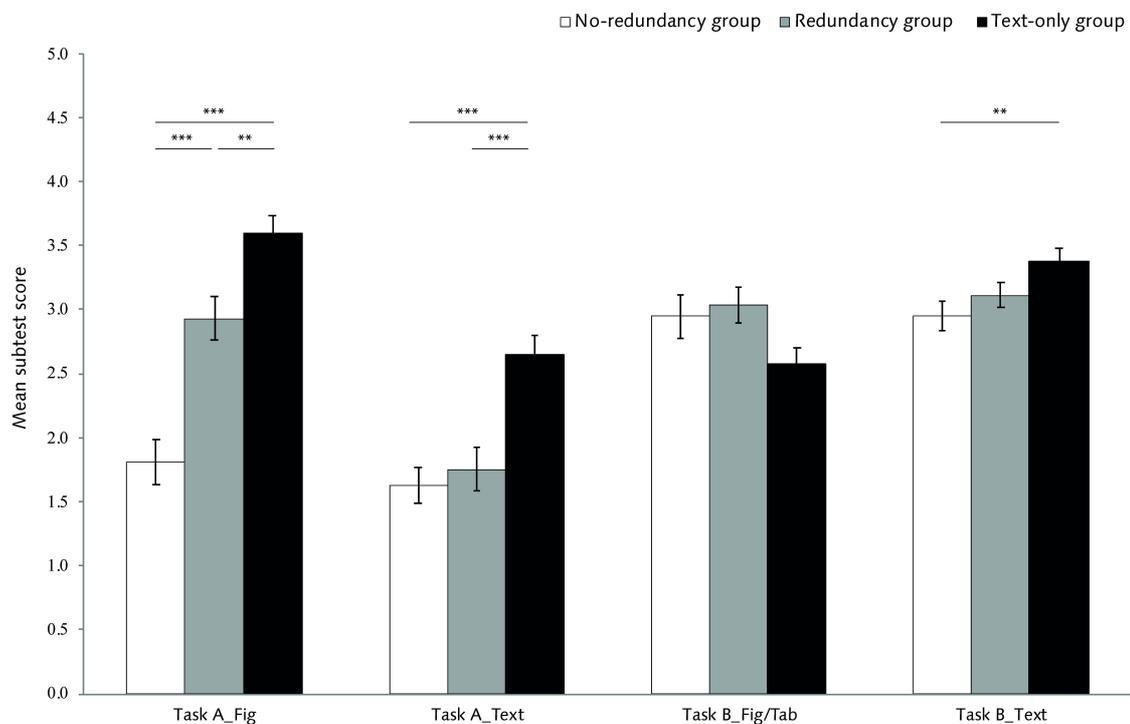


Figure 4.4.1. Mean subtest scores of different information units of Task A and Task B by group collapsed across classroom. The mean subtest scores are presented for relevant information units (originally) represented by figures or texts of Task A and for information units (originally) represented by figures and tables or by texts of Task B. Fig = Figure; Tab = Table. Error bars represent standard errors. All highlighted significance levels for pairwise comparisons were adjusted using Bonferroni correction.

** $p < .01$. *** $p < .001$.

Studie 1

The main effect of classroom was statistically significant for both depictive and textual information in Task A, $F(3, 228) = 11.49, p < .001, \eta_p^2 = .131$ and $F(3, 228) = 7.69, p < .001, \eta_p^2 = .092$, respectively. The interactions of experimental group and classroom were, however, not statistically significant, $F(6, 228) = 1.21, p = .301, \eta_p^2 = .031$ for depictive information in Task A, and $F(6, 228) = 1.39, p = .219, \eta_p^2 = .035$ for textual information in Task A. Likewise for Task B, the main effect of classroom was statistically significant for both depictive and textual information, $F(3, 228) = 7.65, p < .001, \eta_p^2 = .091$ and $F(3, 228) = 4.57, p = .004, \eta_p^2 = .057$, respectively, and the interactions of experimental group and classroom were not statistically significant, $F(6, 228) = 1.92, p = .078, \eta_p^2 = .048$ for depictive information in Task B, and $F(6, 228) = 1.38, p = .222, \eta_p^2 = .035$ for textual information in Task B.

4.4.3 Analyses of Covariance Including Students' Cognitive Abilities

To also account for the effect of individual students' cognitive abilities on their test performance, we ran a set of ANCOVAs with students' cognitive abilities as the covariate. Cognitive abilities had a statistically significant positive effect on students' mean test scores for the entire Task A, $F(1, 224) = 14.59, p < .001, \eta_p^2 = .061$. This effect also existed for the two subscores separately considering depictive or textual information units of Task A, $F(1, 224) = 7.63, p = .006, \eta_p^2 = .033$ and $F(1, 224) = 17.57, p < .001, \eta_p^2 = .073$, respectively. For Task B, however, the covariate cognitive abilities did not significantly affect students' test performance beyond the effects of the two factors experimental group and classroom, neither for the entire Task B nor examining depictive or textual information units separately, $F(1, 224) = 3.68, p = .056, \eta_p^2 = .016$, $F(1, 224) = 2.88, p = .091, \eta_p^2 = .013$, and $F(1, 224) = 1.66, p = .199, \eta_p^2 = .007$, respectively.

Notably, the addition of cognitive abilities did not substantially alter the group's effects. The group effect on students' test performance in Task A was still statistically significant after controlling for the effect of their cognitive abilities, $F(2, 224) = 24.11, p < .001, \eta_p^2 = .177$ for the entire Task A; $F(2, 224) = 28.16, p < .001, \eta_p^2 = .201$ for depictive information in Task A; $F(2, 224) = 15.30, p < .001, \eta_p^2 = .120$ for textual information in Task A. The effect of experimental group remained non-significant for the entire Task B and statistically significant for the depictive information units, $F < 1, p = .726, \eta_p^2 = .003$ and $F(2, 224) = 3.66, p = .027, \eta_p^2 = .032$, respectively. The group effect became non-significant when textual information in Task B was considered, $F(2, 224) = 2.97, p = .053$,

$\eta_p^2 = .026$ (cf. Supplementary Material, Table 4.9.3 for descriptive data and—where appropriate—post hoc test results).

The main effect of classroom also remained statistically significant for the entire Task A and for both depictive and textual information in Task A when students' cognitive abilities were controlled for, $F(3, 224) = 12.38, p < .001, \eta_p^2 = .142$, $F(3, 224) = 12.45, p < .001, \eta_p^2 = .143$, and $F(3, 224) = 8.00, p < .001, \eta_p^2 = .097$, respectively. The interactions of experimental group and classroom were again not statistically significant, $F(6, 224) = 1.16, p = .329, \eta_p^2 = .030$ for the entire Task A, $F(6, 224) = 1.04, p = .399, \eta_p^2 = .027$ for depictive information in Task A, and $F(6, 224) = 1.37, p = .226, \eta_p^2 = .035$ for textual information in Task A. The main effect of classroom was also statistically significant for the entire Task B and for both depictive and textual information in Task B, $F(3, 224) = 7.76, p < .001, \eta_p^2 = .094$, $F(3, 224) = 6.55, p < .001, \eta_p^2 = .081$, and $F(3, 224) = 4.63, p = .004, \eta_p^2 = .058$, respectively. As was the case for Task A, the interactions of experimental group and classroom were not statistically significant, $F(6, 224) = 1.66, p = .132, \eta_p^2 = .043$ for the entire Task B, $F(6, 224) = 1.63, p = .141, \eta_p^2 = .042$ for depictive information in Task B, and $F(6, 224) = 1.27, p = .275, \eta_p^2 = .033$ for textual information in Task B.

4.5 Discussion

In science communication, text–picture combinations are omnipresent. This is true for communication of research findings, but also in the communication of science in the public. Anyone who wants to participate in society in a self-reliant way should therefore be able to identify and integrate information in different types of MERs, but this particularly applies to students who want to pursue a career in science. Students who graduated from secondary school should be equipped to process all kinds of MERs.

4.5.1 Effect of the Task Material

Our study showed that notable differences in undergraduate students' test performance might result from how the task material is presented. In Task A, students who had to identify and integrate information units in both descriptive and depictive representations reached the lowest test scores, far lower than students who could retrieve all information from the text (i.e., the redundancy group [see RQ1] and the text-only group). Students who exclusively worked with textual material performed best. When pictures with some redun-

Studie 1

dant relevant information were presented in addition to nearly the same text, test performance was also markedly lower than in the text-only condition (see RQ2). Unexpectedly, however, the variation of the task material produced such effects on test performance only for Task A, the genetics task; students' performance on Task B, the ecology and evolution task, was not consistently affected by the experimental manipulation. We discuss possible explanations why effects of the task material as observed for Task A failed to show for Task B in the following subsection.

The absence of a positive effect of text–picture material compared to exclusively textual material is contrary to most previous studies, but the studies' material differed in meaningful ways. These differences may explain why the results of our study differ from that of previous studies (e.g., Hegarty & Just, 1993; Lindner et al., 2018; Saß et al., 2012). The way we created redundancy in our task material differed, for example, from studies of Saß et al. (2012) and Lindner et al. (2018). Instead of repeating nearly all information units of the text in a depictive representation, we repeated task-relevant depictive information units in the text. The latter kind of redundancy is common in authentic scientific communication and therefore important for students who want to pursue a career in science. Although there is disagreement over whether scientific journals contain redundant text–picture combinations (Lemke, 1998; Roth et al., 1999), one can summarize that when redundancy occurs only relevant depictive information is repeated in the text and can help the reader correctly interpret the depictive representations (for graphs, see Roth et al., 1999). Furthermore, the depictive representations in our task material contained a greater amount of information than the multiple-choice items generated for the assessment of considerably younger students' science achievement (e.g., Lindner et al., 2018; Saß et al., 2012). This means that students of our study had to exclude irrelevant information in depictive representations and identify the relevant information which has been shown to elevate cognitive load and error rates in graph reading tasks (Strobel, Lindner, Saß, & Köller, 2018). In addition, multiple-choice items with little text make lower demands on mental model construction than the constructed-response items with considerable amounts of material in this study. Hegarty and Just (1993) presented material with which a relatively simple process close to everyday life was to be reconstructed, whereas in our study the domain-specific content requirements were higher, which was also reflected in the depictive representations particularly of Task A.

In line with our results, several studies observed that students sometimes struggle with depictive representations. Research of Roth and colleagues revealed that students—

including college students—have difficulties interpreting graphs (e.g., Bowen & Roth, 1998; Bowen, Roth, & McGinn, 1999; Roth & Bowen, 1999). The fact that students in our study have not sufficiently explored the depictive representations in MERs raises the question of their expectations when dealing with MERs. Because pictures are often decorative in science lessons, pictures may not be first and foremost regarded as a necessary source of information. Cook (2006) assumes that more difficulties are experienced with visual representations the less one knows about the respective topic. Prior content knowledge is necessary to select relevant information and to simultaneously hold more information in working memory. Although students in tertiary education should have knowledge of the content covered by Abitur exams, especially in Task A better knowledge of mitochondrial DNA's structure or genetic disorders may have helped to identify relevant depictive information. However, the observation that students who worked with texts alone outperformed students who worked with almost identical texts and additional redundant pictures in Task A implies that students did not simply disregard the depictive representations—the additional redundant pictures had a deteriorating effect on students' test performance. According to Paivio's dual coding theory (Paivio, 1986), addressing both subsystems should have facilitated mental model construction. This was apparently not the case here: students actually performed worse when presented with redundant pictures in addition to the text (i.e., the redundancy group) rather than merely with verbal information (i.e., the text-only group).

Unlike the assumption of Sweller's redundancy principle, the results of our study did not show a negative effect of redundant text–picture combinations (i.e., repeating task-relevant information from the picture in the text) compared to non-redundant text–picture combinations (where task-relevant information is split between pictures and text). While Sweller refers to learning, Bobis et al. (1993) observed a negative effect of redundant instructional material on paper-folding activities in comparison to diagrams only explaining the results by an increased cognitive load of redundancy group's participants. However, requirements for executing prescribed actions are not directly comparable to requirements for solving constructed-response science items. Whereas our design precluded the potentially confounding influence of differentially extensive material between the no-redundancy and redundancy condition limiting previous research (Bobis et al., 1993; Pociask & Morrison, 2008), cognitive demands might still be higher when information from the verbal and visual subsystem must be integrated (no-redundancy group) rather than when all information can be extracted by a single subsystem (redundancy group). It is conceivable that students working with MERs in Task A focused on the texts and did not pay sufficient

Studie 1

attention to pictures, which would be consistent with the study of Cromley et al. (2010). Consequently, students of the no-redundancy group who needed to identify information in the depictive representations to solve the tasks reached lower test scores than students of the redundancy group who were presented with the same depictive representations, but could have relied completely on the text. Although the text of the text–picture material of both the no-redundancy and the redundancy group occasionally referred to the figure, the material of the no-redundancy group did not contain any content-related referential connections that might have prompted students to focus on relevant parts of the depictive representations as provided by redundant textual information in the redundancy group. Such prompting might contribute to the explanation why the redundancy group outperformed the no-redundancy group. Previous research demonstrated that prompts, such as arrows and bold printing, between associated information in complementary text–picture materials increase science performance, presumably by facilitating text–picture integration (Saß & Schütte, 2016). Eye-tracking research could reveal whether students working with redundant MERs actually move from a particular text passage to the associated redundant part in the depictive representation.

Separating depictive and textual information revealed nearly the same pattern of results as occurred for the entire Task A, while effect sizes differed. Notably, the effect size of group effect examining depictive information was almost twice as large as that examining textual information for Task A, which underlines the relevance of pictures in the presented MERs regarding test performance. With respect to depictive representations, the differences between groups were in line with the above assumptions. However, it was remarkable that for originally textual information units that were provided by text in all groups (in addition to the depictive representations in the redundancy group carrying that same information) the participants working exclusively with textual material also outperformed the participants in both conditions working with MERs and correctly utilized more of those information units in their constructed responses. Original textual information was presented virtually identically in all three materials. A possible explanation might be higher cognitive demands for students of the no-redundancy group and the redundancy group who had to split their attention between external representations and then integrate information of two subsystems. The depictive representations might have served as a distraction in the sense that it directed attention away from the text. In the absence of depictive representations, that is, in the text-only group, students did not have to split their attention between

text and pictures and could thus focus on the text where all relevant information was to be found.

A further explanation might be that negative affective-motivational effects of the complex biology-specific pictures in the material for Task A in the no-redundancy group and the redundancy group contributed to a lack of students' attention to the depictive representations. One might speculate that students of these groups were discouraged when they looked at the pictures and spent less effort on comprehending task material than students of the text-only group; yet they spent comparable amounts of time working on the tasks.

4.5.2 Comparison of Results of Task A and B

Differences between the results of the two tasks might be explained by the depictive representations and the topic of each task. While Task A, the task on genetics, included two biology-specific depictive representations, Task B, the evolution and ecology task, did not. Task B contained two graphs and one data table and these types of representations are not per se biology-specific, but relatively common in everyday life. Consequently, students are likely to encounter them in out-of-school contexts as well as in other school subjects, thus having more learning opportunities for these types of representations. Therefore, it is conceivable that Task B seemed less intimidating and was easier for them to comprehend than Task A. Difficulties that occur when working with domain-specific depictive representations were also shown in several previous studies (e.g., Halverson, Pires, & Abell, 2011; Novick, Stull, & Catley, 2012). For example, Novick et al. (2012) showed that undergraduate biology students have difficulties interpreting information in phylogenetic trees. In addition, the depictive representations of Task A contain a high density of information and only a small proportion of these information units was task-relevant and thus repeated in the text of the redundancy group. In contrast, depictive representations of Task B comprised fewer information units. Therefore, a higher proportion of depictive information units was task-relevant and was repeated in the text of the redundancy group. The biology-specificity and the high density of information may have led to the assumed distracting or discouraging effect of depictive representations of Task A that did not occur in Task B.

In terms of the tasks topics, genetics is often described as difficult, for example because of the need to switch between different levels of biological organization, such as cell and organism level (Knippels, Waarlo, & Boersma, 2005). Although students only had to switch between different subcellular levels in depictive representations of Task A, this

requirement may have been a reason to perceive the task as difficult and to be discouraged by the depictive representations.

As expected, students' cognitive abilities were significantly related to test performance. That an effect of individual student cognitive abilities was observed beyond the effects of experimental group and class for Task A but not for Task B might also indicate that Task A was more difficult. Importantly, the effect sizes of the statistically significant group effects did not notably change after controlling for cognitive abilities. MERs challenged students with high cognitive abilities test scores as much as students with low scores and general cognitive ability did not compensate for the lack of ability to identify and integrate relevant information in MERs. In fact, the effects of students' cognitive abilities were small or of medium size and thus considerably lower than the effects we observed for group.

4.5.3 Strength and Limitations

To the best of our knowledge, our study was the first to compare the effects of redundant and non-redundant biological text–picture combinations and one of only a few studies that recruited students who already obtained their secondary school leaving certificate to investigate the effects of MERs' redundancy. The kind of redundancy we created is similar to the redundancy that—if any—awaits students when they get involved in authentic scientific communication.

Both biology tasks can be considered ecologically valid because they have actually been used in Abitur exams. Unlike other studies, our tasks had a constructed-response format, which places higher cognitive demands on students than multiple-choice items. Multiple-choice items might allow excluding answer options based on some aspects' incompatibility with one's mental model. Every single multiple-choice item is usually far less complex than the Abitur tasks and each answer option may be collated with one's mental model; moreover the likelihoods of the answer options may be collated. Solving a constructed-response item, by contrast, requires test-takers to construct answers based on a correct mental model. At the same time, authenticity of our tasks entailed a high selectivity and the constructed-response format constrained the test length to two tasks. Therefore, task topic is confounded with characteristics of the depictive representations that we discuss as distinguishing the genetics task from the evolution and ecology task so that we cannot exclude the possibility that its specific topic or other local specifics in the selected genetics

task (i.e., Task A) caused the observed deleterious effects of depictive representations in multiple external representations.

With regard to scoring, we cannot ascertain that raters were blind to condition at all times. The response format allows for the possibility that students' responses indicated whether they had been assigned to a text–picture condition (no-redundancy or redundancy group) or the text-only condition. In light of the satisfactory inter-rating results a substantial impact seems rather unlikely as both raters' scoring process would have needed to be equally biased.

The differential results we observed between tasks warrant a thorough investigation of potential negative effects on students' performance as a function of the type of depictive representation. As a first step, repetition of the experiment with similar tasks is required to examine whether the effect of task material replicates with other biology-specific representations high in information density and fails to show in tasks lacking these characteristics. Secondly, tasks covering a broader range of topics should be examined to test whether similar effects do indeed occur. From an applied perspective, it would also be instructive to include more advanced students and systematically vary students' time spent in tertiary education as an additional factor in the research design. Ultimately, one would hope, that the effects of task material would be absent after a certain amount of time as a biology major. Furthermore, due to our working memory test's low reliability, future research must yet shed light on working memory's role in solving constructed-response tasks based on MERs.

Moreover, we could observe the outcome of task solving in our study but not the underlying process. Therefore, eye-tracking research could disclose to which extent students focus on and use depictive and descriptive representations in each task. The present results indicate that depictive representations have received less attention than text and that not all of their relevant elements have been identified. However, textual redundancy seems to help identify relevant parts in depictive representations. Eye-tracking may therefore also provide information as to whether textual redundancy serves as a prompt. It can also shed light on whether students are able to systematically decode depictive information or whether they try to decode it in an unstructured way.

4.5.4 Conclusion and Implications

It is not enough for science education in school to only convey content knowledge. Process-related skills such as communication, which includes interpreting domain-specific depictive representations and different kinds of MERs, must also be conveyed to the students. These skills are fundamental for all students but especially for those who want to pursue a career in science. Given that authentic communication heavily depends on the use of domain-specific figures and MERs that present information in different ways, our results suggest that secondary school education—at least in Germany—might improve on how it prepares students to pursue a career in science. Because the tasks we used were Abitur exam tasks, bachelor students majoring in biology, that is, students who have successfully completed upper secondary school, should be able to solve these tasks. As domain-specific depictive representations negatively affected students' test performance they might benefit from more explicit instructions and more learning opportunities regarding how to independently identify and integrate information from all kinds of MERs. Moreover, students might benefit from more targeted instruction on how to process MERs that are more or less redundant. National guidelines are already in place which require proficient working with representations as part of communication competence (for Germany: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2005). But process-related skills of students have only been proposed to be included in science classes in school alongside subject contents for the past few years. This change in science education is ongoing. Teachers may need time to adapt and possibly more tangible support which is ideally based on rigorous intervention research. The encouraging policy decision to explicitly incorporate process-related skills into required student learning outcomes so far seems to fall short in translating into (effective) teaching.

In light of the significance of multiple external representations in biology and science more generally, we propose that curricula should explicitly include both working with MERs of different degrees of redundancy and working with biology-specific depictive representations. Overall, Western science curricula include working with different representations (e.g., Department for Education, 2015; NGSS Lead States, 2013), but not specifically enough considering the requirements that studying science places on students. This is not only lacking in Germany. Similar demands have already been made by Halverson et al. (2011) with regard to biology-specific depictive representations, while Roth and colleagues criticized that science education does not even prepare students to work with authentic more

general depictive representations (i.e., graphs; e.g., Roth & Bowen, 1999). Furthermore, teachers should not only explicitly teach students how to deal with any kind of MERs and depictive representations, but use more authentic scientific literature in science lessons, for example excerpts of research articles. Being thus empowered students will hopefully also overcome any unease they might experience when confronted with text–picture combinations.

4.6 Acknowledgments

We thank Safa Dadkhah Aseman and Christina Kelting for their help in inter-rating and data entry. Special thanks to all instructors who kindly provided access to their courses.

4.7 Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

4.8 References

- Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education*, 33, 131–152. doi:10.1016/S0360-1315(99)00029-9
- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16, 183–198. doi:10.1016/j.learninstruc.2006.03.001
- Arsenault, D. J., Smith, L. D., & Beauchamp, E. A. (2006). Visual inscriptions in the scientific hierarchy: Mapping the “treasures of science”. *Science Communication*, 27, 376–428. doi:10.1177/1075547005285030
- Bobis, J., Sweller, J., & Cooper, M. (1993). Cognitive load effects in a primary-school geometry task. *Learning and Instruction*, 3, 1–21. doi:10.1016/S0959-4752(09)80002-9
- Bowen, G. M., & Roth, W.-M. (1998). Lecturing graphing: What features of lectures contribute to student difficulties in learning to interpret graphs? *Research in Science Education*, 28, 77–90. doi:10.1007/BF02461643
- Bowen, G. M., & Roth, W.-M. (2002). Why students may not learn to interpret scientific inscriptions. *Research in Science Education*, 32, 303–327. doi:10.1023/A:1020833231966
- Bowen, G. M., Roth, W.-M., & McGinn, M. K. (1999). Interpretations of graphs by university biology students and practicing scientists: Toward a social practice view of scientific representation practices. *Journal of Research in Science Teaching*, 36, 1020–1043. doi:10.1002/(SICI)1098-2736(199911)36:9<1020::AID-TEA4>3.0.CO;2-#
- Brixius, R., Jannan, M., & Kunze, H. (2012). *Abitur; Prüfungsaufgaben mit Lösungen 2013; NRW; Biologie LK* [Abitur; Abitur exam tasks with solutions 2013; NRW; Biology LK] (6. ed.). Hallbergmoos, Germany: Stark.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46. doi:10.1177/001316446002000104
- Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, 90, 1073–1091. doi:10.1002/sce.20164

- Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology, 35*, 59–74. doi:10.1016/j.cedpsych.2009.10.002
- Department for Education (2015). *Statutory guidance – National curriculum in England: Science programmes of study*. Retrieved from <https://www.gov.uk/government/publications/national-curriculum-in-england-science-programmes-of-study/national-curriculum-in-england-science-programmes-of-study> Accessed 3 June 2018.
- Eilam, B., & Gilbert, K. K. (2014). The significance of visual representations in the teaching of science. In B. Eilam & K. K. Gilbert (Eds.), *Science teachers' use of visual representations* (pp. 3–28). Cham, Switzerland: Springer.
- Eilam, B., & Poyas, Y. (2008). Learning with multiple representations: Extending multimedia learning beyond the lab. *Learning and Instruction, 18*, 368–378. doi:10.1016/j.learninstruc.2007.07.003
- 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*, 175–191. doi:10.3758/BF03193146
- Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low Kappa: I. the problems of two paradoxes. *Journal of Clinical Epidemiology, 43*, 543–549. doi:10.1016/0895-4356(90)90158-L
- Halverson, K. L., Pires, C. J., & Abell, S. K. (2011). Exploring the complexity of tree thinking expertise in an undergraduate systematics course. *Science Education, 95*, 794–823. doi:10.1002/sce.20436
- Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32*, 717–742. doi:10.1006/jmla.1993.1036
- Heller, K. A., & Perleth, C. (2000). *KFT 4-12+R - Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision* [KFT 4-12+R - Cognitive ability test for Grade 4 to 12, revision]. Göttingen, Germany: Beltz.
- IBM Corporation (2016). IBM SPSS Statistics for Windows (24) [Software]. Armonk, NY: IBM Corp.

Studie 1

- Knippels, M.-C. P. J., Waarlo, A. J., & Boersma, K. T. (2005). Design criteria for learning and teaching genetics. *Journal of Biological Education*, *39*, 108–112. doi:10.1080/00219266.2005.9655976
- Krajcik, J. S., & Sutherland, L. M. (2010). Supporting students in developing literacy in science. *Science*, *328*, 456–459. doi:10.1126/science.1182593
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. doi:10.2307/2529310
- Lemke, J. (1998). Multiplying meaning – Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of science* (pp. 87–113). London, United Kingdom: Routledge.
- Lindner, M. A., Ihme, J. M., Saß, S., & Köller, O. (2018). How representational pictures enhance students' performance and test-taking pleasure in low-stakes assessment. *European Journal of Psychological Assessment*, *34*, 376–385. doi:10.1027/1015-5759/a000351
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge, United Kingdom: Cambridge University Press.
- NGSS Lead States (2013). *Next generation science standards: for states, by states*. Washington, DC: The National Academies Press.
- Novick, L. R., Stull, A. T., & Catley, K. M. (2012). Reading phylogenetic trees: The effects of tree orientation and text processing on comprehension. *BioScience*, *62*, 757–764. doi:10.1525/bio.2012.62.8.8
- Paivio, A. (1986). Dual coding theory. In A. Paivio (Ed.), *Mental representations: A dual coding approach* (pp. 53–83). New York, NY: Oxford University Press.
- Pociask, F. D., & Morrison, G. R. (2008). Controlling split attention and redundancy in physical therapy instruction. *Educational Technology Research and Development*, *56*, 379–399. doi:10.1007/s11423-007-9062-5
- Roth, W.-M., & Bowen, G. M. (1999). Of cannibals, missionaries, and converts: Graphing competencies from grade 8 to professional science inside (classrooms) and outside (field/laboratory). *Science, Technology, & Human Values*, *24*, 179–212. doi:10.1177/016224399902400201

- Roth, W.-M., Bowen, G. M., & McGinn, M. K. (1999). Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching*, *36*, 977–1019. doi:10.1002/(SICI)1098-2736(199911)36:9<977::AID-TEA3>3.0.CO;2-V
- Saß, S., & Schütte, K. (2016). Helping poor readers demonstrate their science competence: Item characteristics supporting text–picture integration. *Journal of Psychoeducational Assessment*, *34*, 91–96. doi:10.1177/0734282915588389
- Saß, S., Wittwer, J., Senkbeil, M., & Köller, O. (2012). Pictures in test items: Effects on response time and response correctness. *Applied Cognitive Psychology*, *26*, 70–81. doi:10.1002/acp.1798
- Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, *14*, 101–120. doi:10.1023/a:1013136727916
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2 ed., pp. 72–103). Cambridge, United Kingdom: Cambridge University Press.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, *13*, 141–156. doi:10.1016/s0959-4752(02)00017-8
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2004). *Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie* [Uniform examination requirements for the Abitur exam in biology]. Munich, Germany: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004* [National educational standards for the intermediate school leaving certificate in biology. Resolution of 16.12.2004]. Munich, Germany: Luchterhand.
- Strobel, B., Lindner, M. A., Saß, S., & Köller, O. (2018). Task-irrelevant data impair processing of graph reading tasks: An eye tracking study. *Learning and Instruction*, *55*, 139–147. doi:10.1016/j.learninstruc.2017.10.003

Studie 1

- Sweller, J. (2005). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159–167). Cambridge, United Kingdom: Cambridge University Press.
- Thorndike, R. L., & Hagen, E. (1971). *Cognitive abilities test*. Boston, MA: Houghton-Mifflin.
- Tsui, C.-Y., & Treagust, D. F. (2013). Introduction to multiple representations: their importance in biology and biological education. In D. F. Treagust & C.-Y. Tsui (Eds.), *Multiple representations in biological education* (pp. 3–18). Dordrecht, The Netherlands: Springer.
- Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, 4(433), 1–22.
doi:10.3389/fpsyg.2013.00433

4.9 Supplementary Material

4.9.1 Supplementary Material S1: Detailed Representations of the Data by Boxplots

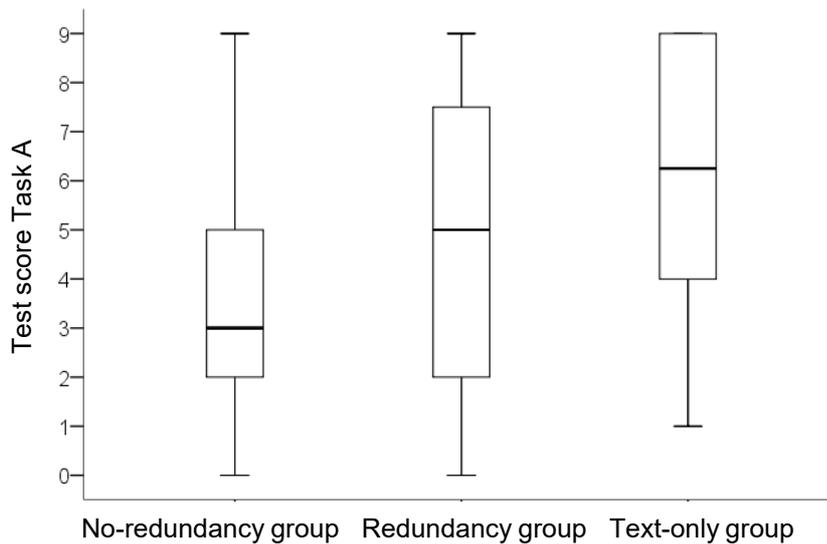


Figure 4.9.1. Boxplots of test performance for the entire Task A separated by group. Boxes show the middle 50% of scores, and whiskers show the top and bottom 25%. The bold horizontal line in each box depicts the median.

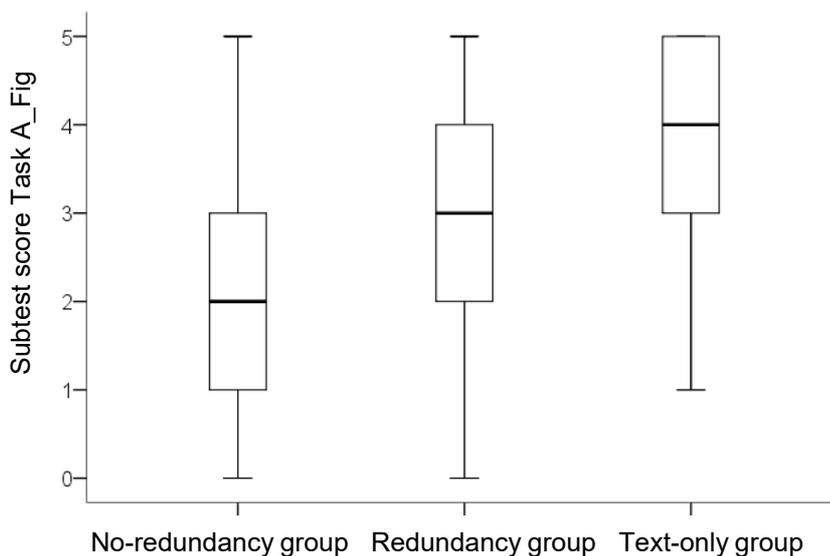


Figure 4.9.2. Boxplots of test performance examining the (originally) depictive information units of Task A separated by group. Boxes show the middle 50% of scores, and whiskers show the top and bottom 25%. The bold horizontal line in each box depicts the median.

Studie 1

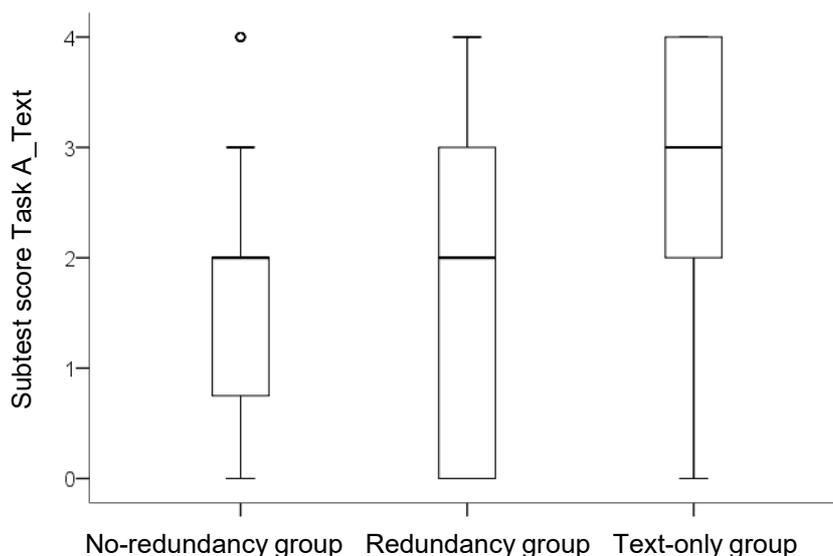


Figure 4.9.3. Boxplots of test performance examining the (originally) descriptive information units of Task A separated by group. Boxes show the middle 50% of scores, whiskers show the top and bottom 25%, and the individual dot represents outliers (scores of four students). The bold horizontal line in each box depicts the median.

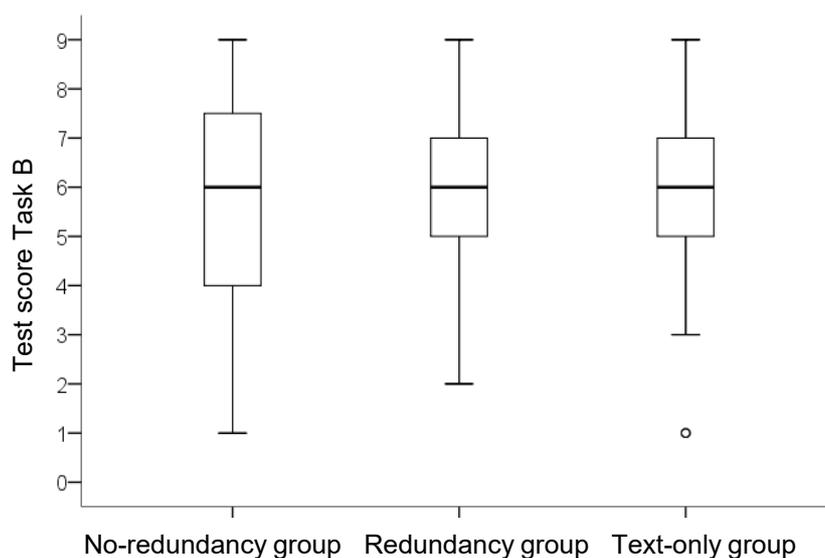


Figure 4.9.4. Boxplots of test performance for the entire Task B separated by group. Boxes show the middle 50% of scores, whiskers show the top and bottom 25%, and the individual dot represents outliers (scores of two students). The bold horizontal line in each box depicts the median.

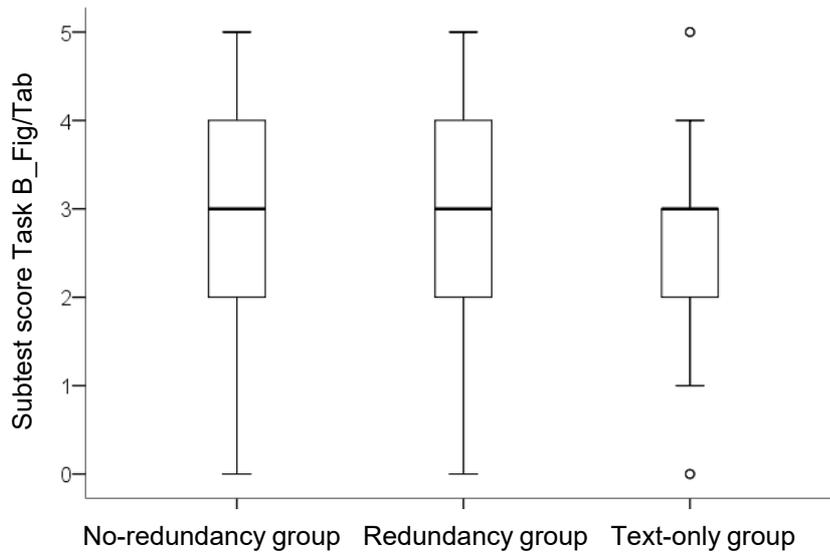


Figure 4.9.5. Boxplots of test performance examining the (originally) depictive information units of Task B separated by group. Boxes show the middle 50% of scores, whiskers show the top and bottom 25%, and the individual dots represent outliers (each dot represents scores of two students). The bold horizontal line in each box depicts the median.

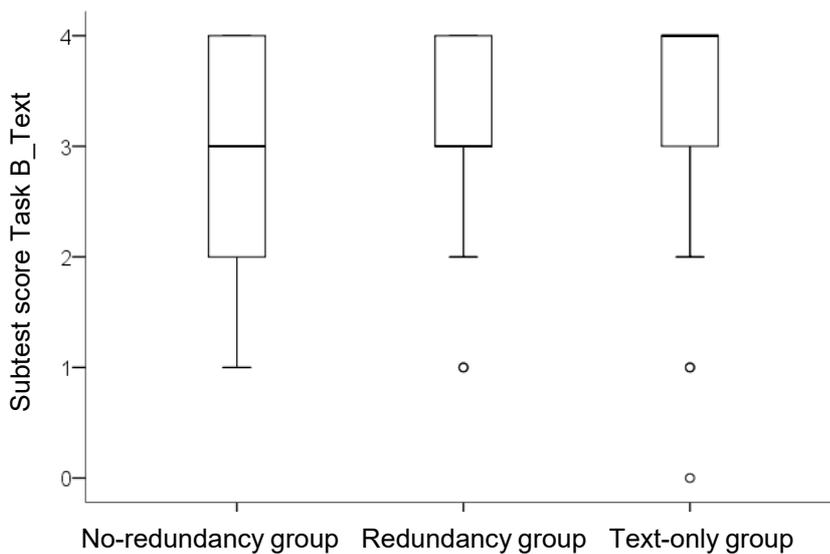


Figure 4.9.6. Boxplots of test performance examining the (originally) descriptive information units of Task B separated by group. Boxes show the middle 50% of scores, whiskers show the top and bottom 25%, and the individual dots represent outliers (each dot at a score = 1 represents results of two students, the dot at score = 0 the result of one student). The bold horizontal line in each box depicts the median.

4.9.2 Supplementary Tables

Table 4.9.1

Excerpt of the category scheme for the demonstration example (cf. Figures 4.3.1–4.3.3)

Representation	Information unit	Responses scored as correct
(originally) Figure 1 ^a (maximum one point)	Position of L (tRNA ^{Leucine})	<ul style="list-style-type: none"> • “L” is encoded at position 3230-3304 • tRNA^{Leucine} is encoded at position 3230-3304 • other responses indicating that the position of “L” has been identified
Text 1 ^a (maximum three points)	The mitochondrial DNA encodes tRNAs	<ul style="list-style-type: none"> • The mtDNA encodes tRNAs • Figure 1 shows areas that are transcribed in tRNA • The mtDNA encodes the tRNA^{Leucine} at position 3230-3304
	tRNA ^{Leucine} is abbreviated here as “L”	<ul style="list-style-type: none"> • “L” is the abbreviation for the tRNA that transports Leucine • The mtDNA encodes the tRNA^{Leucine} at position 3230-3304
	ATPase6/ATP synthase is involved in the supply of an energy carrier	<ul style="list-style-type: none"> • ATPase6/ATP synthase is important for energy supply • ATPase6 is a subunit of the ATP synthase, which is involved in the supply of an energy carrier

^aDepending on the group students were assigned to, they might only have been able to retrieve relevant information units from the text.

Table 4.9.2

Pairwise comparisons between experimental groups for Task A and B

Measurement	Task A			Task B		
	Total information units	Depictive information units	Textual information units	Total information units	Depictive information units	Textual information units
Post hoc test with Bonferroni correction						
No-redundancy group vs. redundancy group	$p = .005$, $d = 0.46$	$p < .001$, $d = 0.73$	$p = 1.00$, $d = 0.09$		$p = 1.00$, $d = 0.06$	$p = .751$, $d = 0.17$
No-redundancy group vs. text-only group	$p < .001$, $d = 1.13$	$p < .001$, $d = 1.25$	$p < .001$, $d = 0.78$		$p = .167$, $d = 0.28$	$p = .008$, $d = 0.45$
Redundancy group vs. text-only group	$p < .001$, $d = 0.59$	$p = .007$, $d = 0.46$	$p < .001$, $d = 0.62$		$p = .055$, $d = 0.39$	$p = .190$, $d = 0.31$

Note. Post hoc test results are reported for two-way ANOVAs in which the effect of experimental group was statistically significant. Sample size: $N = 240$, $n_{\text{Group 1}} = 79$, $n_{\text{Group 2}} = 79$, $n_{\text{Group 3}} = 82$.

Table 4.9.3

Descriptive data and pairwise comparisons between experimental groups for Task A and B with cognitive abilities as covariate

Measurement	Task A			Task B		
	Total information units	Depictive information units	Textual information units	Total information units	Depictive information units	Textual information units
Estimated marginal means (sub) test score (SE)						
No-redundancy group	3.73 (0.28)	2.01 (0.16)	1.72 (0.15)	6.11 (0.20)	3.07 (0.15)	3.04 (0.11)
Redundancy group	4.79 (0.28)	3.01 (0.16)	1.79 (0.15)	6.22 (0.20)	3.10 (0.14)	3.11 (0.10)
Text-only group	6.39 (0.27)	3.66 (0.15)	2.74 (0.14)	6.00 (0.19)	2.62 (0.14)	3.38 (0.10)
Post hoc test (with Bonferroni correction)						
No-redundancy group vs. redundancy group	$p = .021$, $d = 0.43$	$p < .001$, $d = 0.71$	$p = 1.00$, $d = 0.05$		$p = 1.00$, $d = 0.03$	
No-redundancy group vs. text-only group	$p < .001$, $d = 1.10$	$p < .001$, $d = 1.19$	$p < .001$, $d = 0.78$		$p = .079$, $d = 0.36$	
Redundancy group vs. text-only group	$p < .001$, $d = 0.66$	$p = .010$, $d = 0.47$	$p < .001$, $d = 0.73$		$p = .050$, $d = 0.38$	

Note. Post hoc test results are reported for two-way ANCOVAs in which the experimental effect of group was statistically significant. Sample size: $N = 237$, $n_{\text{Group 1}} = 78$, $n_{\text{Group 2}} = 79$, $n_{\text{Group 3}} = 80$.

5 Studie 2: Fachspezifische komplexe Abbildungen verarbeiten – (auch) eine Frage der Überzeugung¹⁹

Processing subject-specific complex pictures – (also) a question of one's belief

Zusammenfassung

Abbildungen in den Naturwissenschaften verwenden spezielle grafische Konventionen und können große Datenmengen visualisieren. In der Kommunikation in den Naturwissenschaften sind sie – meist kombiniert mit Texten – omnipräsent. Kompetentes Arbeiten mit ihnen ist daher unverzichtbar, wird aber während der Schullaufbahn nur unzureichend erlernt. In dieser Studie wurde der Effekt einer Selbstwirksamkeitserwartung bezüglich des Arbeitens mit fachspezifischen, komplexen Abbildungen in der Biologie auf Testleistung untersucht. Dazu bearbeiteten $N = 205$ Bachelorstudierende der Biologie 2 auf Abituraufgaben basierende Biologieaufgaben, wobei das Testmaterial zwischen Gruppen variiert wurde. Multiple Regressionsanalysen zeigten bei der Biologieaufgabe mit fachspezifischen, komplexen Abbildungen einen statistisch signifikanten positiven Effekt der Selbstwirksamkeitserwartung auf die Testleistung, der durch die Anforderungen des Materials moderiert wurde. Es werden Implikationen für das Lehren von Biologie diskutiert.

Schlüsselbegriffe: Selbstwirksamkeitserwartung, komplexe Abbildungen, Kommunikation in Naturwissenschaften, Testleistung, Sekundarbildung

¹⁹ Dies ist die akzeptierte Manuskriptversion des folgenden Artikels, der – verfügbar unter <https://www.reinhardt-journals.de/index.php/peu/article/view/153866/0> – im Journal *Psychologie in Erziehung und Unterricht* im Ernst Reinhardt Verlag erschienen ist: Magnus, L., Schütte, K., & Schwanewedel, J. (2021). Fachspezifische, komplexe Abbildungen verarbeiten – (auch) eine Frage der Überzeugung. *Psychologie in Erziehung und Unterricht*, 68. <https://doi.org/10.2378/peu2021.art29d>

Abstract

Pictures in science use special graphical conventions, visualize a large amount of data points or variables. They are omnipresent in science communication – usually combined with texts. Therefore, working competently with these pictures is indispensable, but is insufficiently learned during the school career. This study examined the effect of a self-efficacy belief regarding working with subject-specific, complex pictures on test performance. For this, $N = 205$ undergraduate students of biology worked on 2 biology tasks based on German Abitur exams, whereby the test material varied between groups. Multiple regression analyses showed a statistically significant positive effect of the self-efficacy belief on test performance for the task with subject-specific, complex pictures, which was moderated by the requirements of the test material. Implications for teaching biology are discussed.

Keywords: Self-efficacy, complex representations, science communication, test performance, secondary education

5.1 Einleitung

In den Naturwissenschaften sind Abbildungen allgegenwärtig (Lemke, 1998; Roth & Pozzer-Ardenghi, 2013). Sie werden beispielsweise genutzt, um schwer fassbare, komplexe Phänomene abzubilden (Schönborn & Anderson, 2006) oder Daten zu visualisieren und dabei Erkenntnisse zu gewinnen und zu kommunizieren (Heßler, 2006). Auch während der COVID-19-Pandemie wurde die Vielfältigkeit der Abbildungen im naturwissenschaftlichen Bereich sehr deutlich, als regelmäßig von Wissenschaftler:innen erstellte Abbildungen zu neuesten Forschungsergebnissen (z. B. Viruskonzentrationen in Proben von Patientengruppen) in die Öffentlichkeit getragen wurden. Kommentare unter solchen Abbildungen in den sozialen Medien dokumentieren, dass mitunter große Schwierigkeiten bei deren Interpretation auftraten und sogar zum Teil Annahmen resultierten, die zu den eigentlichen Forschungsergebnissen konträr waren. Dies deutet darauf hin, dass Abbildungen ihr gewinnbringendes Potenzial nicht ausschöpfen können, wenn die Rezipierenden nicht wissen, wie sie die Abbildungen „lesen“ müssen. Anders als im Alltag allgegenwärtige Abbildungen (z. B. Diagramme mit wenigen Datenpunkten in Zeitungen) verwenden fachspezifische Abbildungen in den Naturwissenschaften häufig eine Vielzahl spezieller grafischer Konventionen oder

visualisieren eine große Anzahl von Datenpunkten oder Variablen, was ihnen eine verhältnismäßig hohe Komplexität verleiht (Lowe, 2000; Trafton, Marshall, Mintz & Trickett, 2002). Verschiedene empirische Studien implizieren, dass gerade diese fachspezifischen, komplexen Abbildungen Probleme bei der Interpretation oder dem Erfassen der dargestellten Informationen hervorrufen können (z. B. Magnus, Schütte & Schwanewedel, 2020; Novick, Stull & Catley, 2012), dies, obwohl nationale Bildungsstandards für die Naturwissenschaften in vielen Ländern (z. B. NGSS Lead States, 2013; Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2005, 2020) das kompetente Arbeiten mit verschiedenen Abbildungsarten von Schüler:innen explizit fordern. Zudem gibt es Hinweise darauf, dass Abbildungen mitunter nur überflogen oder übersprungen werden, wenn sie in Kombination mit Texten präsentiert werden (Cromley, Snyder-Hogan & Luciw-Dubas, 2010).

Ein auch von Cromley et al. (2010) in Betracht gezogener Erklärungsansatz dafür, dass Abbildungen zum Teil nicht hinreichend beachtet werden, wodurch wertvolle Informationen verloren gehen, könnte die persönliche Überzeugung von den eigenen Fähigkeiten sein, sich Informationen aus komplexen Abbildungen erschließen zu können. Diese spezifische Selbstwirksamkeitserwartung könnte mit einer bestimmten auf das Arbeiten mit komplexen Abbildungen bezogenen Motivation einhergehen. Wird sie nicht zum Beispiel durch positive Erfahrungen mit solchen Abbildungen gestärkt, könnte dies negative Auswirkungen darauf haben, wie Personen an komplexe Abbildungen herangehen und wie sie die Informationsverarbeitung bewältigen. In dieser Studie haben wir untersucht, ob sich ein Effekt der Selbstwirksamkeitserwartung Biologiestudierender bezüglich des Arbeitens mit komplexen Abbildungen auf ihre Testleistung in Aufgaben nachweisen lässt, die unter anderem komplexe Abbildungen enthielten.

5.1.1 Die Rolle und Effekte von Abbildungen in den Naturwissenschaften

Die Naturwissenschaften bedienen sich einer Vielzahl von Abbildungen, um erhobene Daten aufzubereiten und Ergebnisse zu vermitteln (Arsenault, Smith & Beauchamp, 2006; Heßler, 2006). Zudem lassen sich Phänomene auf makro-, mikro- und submikroskopischer Organisationsebene und auch abstrakte Konzepte durch Abbildungen wie beispielsweise Schemazeichnungen oder Fotografien sichtbar machen (Schönborn & Anderson, 2006). Abbildungen werden in naturwissenschaftlichen

Studie 2

Fachpublikationen und in Lehrbüchern zumeist mit Text und teilweise mit weiteren Abbildungen kombiniert (z. B. Lemke, 1998). Der Grad an Redundanz, also inwieweit sich Informationen in den einzelnen Informationsmedien (Texte, Abbildungen) wiederholen, variiert. Wenn Informationen aus Abbildungen nicht im Text wiederholt werden, müssen Rezipierende zwingend die wesentlichen Elemente in Text und Bild identifizieren und anschließend zueinander in Beziehung setzen, ein Prozess, der als Kohärenzbildung bezeichnet wird (Seufert, 2003).

Zwar konnten empirische Studien positive Effekte von Abbildungen in Lern- und Testmaterialien auf die zu erbringende Leistung nachweisen (z. B. Eilam & Poyas, 2008; Lindner, Ihme, Saß & Köller, 2018), andere Studien zeigten jedoch, dass gerade fachspezifische, komplexe Abbildungen Probleme bei der Aufgabenbearbeitung bereiten können (z. B. Magnus et al., 2020; Novick et al., 2012). Im Gegensatz zu alltäglichen Abbildungen (z. B. Beschilderungen an Flughäfen, einfache Graphen mit wenigen Variablen und Datenpunkten), zeichnen sich fachspezifische Abbildungen in den Naturwissenschaften durch folgende Eigenschaften aus: Sie sind häufig abstrakt, verwenden Farbcodierungen und eine Vielzahl spezieller grafischer Konventionen, bilden zum Teil eine große Anzahl von Datenpunkten oder Variablen ab (z. B. spezielle Wetterkarten in der Meteorologie) und sind somit komplexer als Abbildungen aus dem Alltag (Lowe, 2000; Trafton et al., 2002). Die Ergebnisse bisheriger Studien legen nahe, dass Schüler:innen nach Beendigung der Schullaufbahn nicht automatisch in der Lage sind, mit jeder Art von Abbildung in den Naturwissenschaften umzugehen (Magnus et al., 2020; Novick et al., 2012). Cromley et al. (2010) beobachteten, dass Biologiestudierende bei der Bearbeitung von Biologieaufgaben mit Text-Bild-Kombinationen die Bilder häufig lediglich überflogen oder sogar übersprungen haben. Wenn sie aber mit den Abbildungen arbeiteten, erreichten sie einen höheren Anteil kognitiver Aktivitäten auf erhöhtem Level im Vergleich zum Arbeiten mit reinem Text. Es stellt sich also die Frage, warum das gewinnbringende Potenzial von Abbildungen häufig nicht ausgeschöpft wird, diese mitunter nicht einmal beachtet werden. Über alle Schulfächer betrachtet werden vornehmlich Texte als Informationsmedium genutzt und der Umgang mit diesen wird – mindestens im Deutschunterricht – explizit geübt (Rosebrock & Nix, 2010). Dagegen wird das Lesen von Abbildungen häufig nicht explizit eingeübt, obwohl es gerade für komplexe, fachspezifische Abbildungen förderlich wäre (Lowe, 2000; Schönborn & Anderson, 2006). Dies könnte sich potentiell über

die Schulzeit hinaus auf die Selbstwirksamkeitserwartung der Schüler:innen in Bezug auf das Arbeiten mit solchen Abbildungen auswirken.

5.1.2 Die Beziehung zwischen Selbstwirksamkeitserwartung und akademischer Leistung

Die Selbstwirksamkeitserwartung, verankert in Banduras sozial-kognitiver Lerntheorie, bezeichnet die Überzeugung, durch eigenes Handeln gewünschte Ergebnisse erzielen zu können (Bandura, 2001). Sie ist somit aufgaben- und domänenspezifisch (Pajares, 1996), aber zwischen Aufgaben und zwischen Domänen positiv assoziiert, umso mehr, je ähnlicher diese sich jeweils sind (Bong, 1997). Es wurden verschiedene Quellen der Selbstwirksamkeitserwartung identifiziert: Erfolgserfahrungen, stellvertretende Erfahrung, verbale Überzeugungsarbeit und emotionale Erregung (Bandura, 1977). Das Erfahren von Erfolgen, die Lernende auf ihre eigene Fähigkeit und Anstrengung zurückführen können, wird dabei als einflussreichste Quelle betrachtet (z. B. Schwarzer & Jerusalem, 2002). Wird eine positive Selbstwirksamkeitserwartung in einem Gebiet aufgebaut, ist sie relativ stabil (Schwarzer & Jerusalem, 2002). Die spezifische Selbstwirksamkeitserwartung kann sich auf die Anstrengung, das Durchhaltevermögen und die Widerstandsfähigkeit einer Person bezüglich einer Tätigkeit in genau diesem Gebiet auswirken und somit letztendlich deren Leistung beeinflussen (Pajares, 1996). Zahlreiche empirische Studien belegen den Zusammenhang von Selbstwirksamkeitserwartung und Leistungen im naturwissenschaftlichen Bereich (z. B. Ardura & Pérez-Bitrián, 2019; Schütte, Frenzel, Asseburg & Pekrun, 2007). Dabei wurde jeweils eine spezifische Selbstwirksamkeitserwartung von Schüler:innen oder Studierenden erfragt und deren Effekt auf die Leistung, die in Form von Noten oder Leistungstests erfasst wurde, untersucht.

5.1.3 Fragestellung und Hypothese

Auf das Arbeiten mit komplexen Abbildungen in den Naturwissenschaften und spezifisch in der Biologie übertragen, ist erwartbar, dass die während der Schullaufbahn und des bisherigen Studiums gesammelten Erfahrungen mit solchen Abbildungen dazu führen, dass Biologiestudierende eine spezifische diese Anforderung betreffende Selbstwirksamkeitserwartung ausgebildet haben. Eine solche Selbstwirksamkeitserwartung würde sich darauf auswirken, wie Personen an die Bearbeitung von

Studie 2

Aufgaben herangehen, die unter anderem fachspezifische, komplexe Abbildungen enthalten, und welche Leistung sie letztendlich bei der Bearbeitung dieser Aufgaben erzielen.

Unter Kontrolle experimentell variiertes Anforderungen an die Arbeit mit komplexen Abbildungen auf die Testleistung (deren Effekte Gegenstand von Magnus et al., 2020, sind), wurde der Einfluss der Selbstwirksamkeitserwartung auf die Testleistung untersucht. Konkret wurde folgende Hypothese geprüft:

Die Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie hat einen positiven Effekt auf die Testleistung der Studierenden, dies auch unter Kontrolle von kognitiven Leistungsmaßen (sowohl fachspezifisch als auch allgemein), der durch die Art des zur Verfügung gestellten Aufgabenmaterials, also die Experimentalgruppenzugehörigkeit, moderiert wird. Bei Studierenden, die zum Lösen der Aufgaben Informationen aus Abbildungen entnehmen mussten (im Folgenden als Gruppe *keine Redundanz* bezeichnet), sollte ein positiver Zusammenhang von Selbstwirksamkeitserwartung und Testleistung vorliegen. Bei Studierenden, die zum Lösen derselben Aufgaben keine Informationen aus Abbildungen entnehmen mussten, wird dieser Zusammenhang grundsätzlich nicht erwartet (im Folgenden als Gruppe *Redundanz* bzw. Gruppe *nur Text* bezeichnet). In dem Maße, wie Studierende aber auch dann Abbildungen zum Lösen der Aufgaben hinzuziehen, wenn der Text alle dafür notwendigen Informationen enthält (Gruppe *Redundanz*), wird wiederum ein positiver Zusammenhang zwischen Selbstwirksamkeitserwartung und der Testleistung erwartet. Es ist dabei unerheblich, ob sich die Studierenden, die über die verschiedenen Versionen des Materials nicht informiert worden waren, der Art des Materials als redundant oder nicht redundant bewusst geworden sind.

5.2 Methode

5.2.1 Stichprobe und Studiendurchführung

An der Studie, die 2016/2017 durchgeführt wurde, nahmen $N = 240$ Biologiestudierende im Bachelorstudium teil. Studierende wurden aus der Stichprobe ausgeschlossen, wenn sie die Items zu Selbstwirksamkeitserwartung oder kognitiven Fähigkeiten nicht vollständig bearbeitet oder ihre letzte Biologienote nicht angegeben hatten. Es ergab sich eine finale Stichprobe von $N = 205$ ($M = 21.7$ Jahre, $SD = 3.3$) Studierenden, von denen 71,2 % weiblich waren.

Die Studierenden wurden zufällig einer von drei Experimentalgruppen zugeordnet und aufgefordert, zwei Biologieaufgaben mit offenem Antwortformat in insgesamt 45 Minuten zu bearbeiten. Die Erstautorin führte alle Datenerhebungen durch und informierte die Teilnehmenden darüber, dass die Teilnahme freiwillig und anonym erfolgte und die Testleistung keinerlei Einfluss auf ihre Kursnote haben würde. Um Effekte der Aufgabenreihenfolge zu vermeiden, wurden die Biologieaufgaben in einem balancierten Design dargeboten, d. h. in jeder Gruppe gab es zwei Testheftversionen, in denen entweder zuerst Aufgabe A und dann B dargestellt war oder zuerst Aufgabe B und dann A. Das Aufgabenmaterial variierte zwischen den Gruppen: Die Gruppe *keine Redundanz* ($n = 68$) erhielt Material mit Abbildungen und Text und war zum Lösen der Aufgaben darauf angewiesen, Informationen sowohl aus Texten als auch aus den Abbildungen zu entnehmen. Im Material der Gruppe *Redundanz* ($n = 71$) wurden hingegen lösungsrelevante Informationen aus den Abbildungen zusätzlich im Text wiederholt, sodass die Aufgaben selbst dann gelöst werden konnten, wenn die Abbildungen nicht hinzugezogen wurden. Die Gruppe *nur Text* ($n = 66$) arbeitete mit Material, das ausschließlich aus Text bestand. Die Aufgabenstellungen waren in allen Gruppen identisch und die Teilnehmenden wurden nicht auf die unterschiedlichen Experimentalgruppen hingewiesen. Nach Bearbeitung der Biologieaufgaben hatten die Studierenden wiederum bis zu 45 Minuten Zeit, einen Fragebogen mit unter anderem den Skalen zur Erfassung ihrer Selbstwirksamkeitserwartung und kognitiven Fähigkeiten zu bearbeiten und ihre letzte Biologienote sowie einige demografische Daten anzugeben. Am Ende der Studie wurden die Teilnehmenden umfassend über die Ergebnisse unterrichtet. Die Teilnahme wurde mit Süßigkeiten honoriert und zusätzlich bestand die Möglichkeit, an einer Verlosung von Gutscheinen im Wert von 5 € bis 50 € teilzunehmen.

5.2.2 Messinstrumente

5.2.2.1 Selbstwirksamkeitserwartung

Zur Erfassung der Selbstwirksamkeitserwartung in Bezug auf das Arbeiten mit komplexen Abbildungen in der Biologie wurde eine Skala mit 4 Items entwickelt. Bei der Formulierung diente die gut etablierte Skala zur Erfassung der schulbezogenen Selbstwirksamkeitserwartung von Jerusalem und Satow (1999) als Orientierung. Die Items lauten wie folgt: „Selbst wenn ich eine komplexe biologische Abbildung nicht

Studie 2

„sofort verstehe, versuche ich weiterhin, sie mir zu erschließen.“, „Ich kann mir biologische Inhalte auch dann erschließen, wenn sie in für mich ungewohnten, neuen Abbildungen dargestellt sind.“, „Wenn Informationen in Biologie in komplexen Abbildungen dargestellt sind, macht es mir Spaß, die Informationen herauszuarbeiten.“, „Auch wenn eine biologische Abbildung auf den ersten Blick komplex erscheint, fühle ich mich in der Lage, mit ihr zu arbeiten“. Die Studierenden schätzten sich auf einer 4-stufigen Likert-Skala ein (1 = „Stimme ganz zu“, 4 = „Stimme gar nicht zu“). Im Nachgang wurden die Items recodiert, sodass hohe Werte eine hohe Selbstwirksamkeitserwartung bedeuten. Die interne Konsistenz der Skala war akzeptabel (Cronbachs $\alpha = .76$). Für weitere Analysen wurden Mittelwerte gebildet und z -transformiert.

5.2.2.2 Kognitive Fähigkeiten

Zur Erfassung der kognitiven Fähigkeiten der Studierenden wurde die nonverbale Skala N2 des Kognitiven Fähigkeitstests (KFT) 4–12 + R eingesetzt (Heller & Perleth, 2000). Diese Skala besteht aus 25 Items, die jeweils zwei sinnvoll zusammenpassende Figuren oder Zeichnungen sowie eine dritte Figur oder Zeichnung zeigen. Aus weiteren fünf dargestellten Figuren oder Zeichnungen muss dann die eine ausgewählt werden, die auf die Art und Weise zur dritten Figur oder Zeichnung passt, wie die zweite zur ersten. Es wurde jeweils die Gesamtpunktzahl berechnet und einer z -Transformation unterzogen.

5.2.2.3 Biologienote

Die von den Studierenden in unterschiedlichen Notensystemen angegebenen letzten Schulnoten im Fach Biologie wurden zur Vereinheitlichung dem Notensystem mit Werten von 1 (sehr gut) bis 6 (ungenügend) ohne Dezimalstellen zugeordnet, wobei nur Noten beobachtet wurden, die dem Bereich von 1 bis 4 entsprachen. Anschließend wurden die Werte z -transformiert.

5.2.3 Material und Aufgabenbewertung

Das Aufgabenmaterial basierte auf Biologie-Abituraufgaben aus Nordrhein-Westfalen aus den Jahren 2008 (Aufgabe A) und 2009 (Aufgabe B) (Brixius, Jannan & Kunze, 2012, S. 2008-13, 2008-15–2008-17 bzw. S. 2009-1–2009-3). Aufgabe A war eine Genetikaufgabe und thematisierte Ursache und Symptome einer Muskelerkrankung, die genaue Aufgabenstellung lautete: „Erläutern Sie unter Bezug auf die im 108

Material dargestellten Erkenntnisse die molekulargenetische Ursache des MERRF-Krankheitsbildes. Erläutern Sie darüber hinaus, wie es zu dem für das MERRF-Syndrom typischen Symptom der Muskelschwäche kommt.“²⁰ In Abhängigkeit von der Experimentalgruppe standen den Studierenden für Aufgabe A zwei biologiespezifische schematische Abbildungen aus dem Bereich der Molekulargenetik (eine Karte der mitochondrialen DNA sowie der Ausschnitt der Basenabfolge der mitochondrialen DNA einer gesunden und einer erkrankten Person), ein Graph mit zehn beschrifteten Datenpunkten und einer angepassten Kurve (zu Proteinen in den Mitochondrien Erkrankter) und textliches Material (3 Textblöcke) zur Verfügung (Gruppen *keine Redundanz* und *Redundanz*) oder rein textliches Material (Gruppe *nur Text*). Beispielhaft soll die Variation des Testmaterials der drei Experimentalgruppen anhand der Karte der mitochondrialen DNA erläutert werden. Dies ist eine schematische Abbildung eines ringförmigen DNA-Doppelstrangs, in den mit den zugehörigen Basenpaarpositionen Regionen eingezeichnet sind, die für bestimmte Gene codieren. Zum Lösen der Aufgabe musste in dieser Abbildung abgelesen werden, welcher Bereich dieser DNA für das Gen einer bestimmten tRNA codiert. Diese Information konnte im Material der Gruppe *keine Redundanz* ausschließlich aus der Abbildung entnommen werden. Im Material der Gruppe *Redundanz* wurde der spezifische Bereich der DNA zusätzlich im Text genannt, sodass diese Information aus der Abbildung oder dem Text entnommen werden konnte. Da die Gruppe *nur Text* keine Abbildungen erhielt, wurde die Information aus der Abbildung wie bei Gruppe *Redundanz* in den Text eingefügt und konnte anders als im Material der Gruppe *Redundanz* lediglich aus diesem entnommen werden. In dem zu dieser Abbildung gehörenden Textblock waren weitere zum Lösen der Aufgabe relevante Informationen enthalten, die aber in allen Gruppen nur in textlicher Form dargeboten wurden.²¹ Da das eingesetzte Material nahezu dem der Original-Abituraufgaben entsprach, kann angenommen werden, dass der Schwierigkeitsgrad sowohl des bildlichen Materials als auch der Texte für Personen, die das Abitur erfolgreich absolviert haben, angemessen war. Es wurden nur Fachbegriffe verwendet, von denen erwartet werden kann, dass Schüler:innen sie bis zum Abitur (im Schulunterricht) erworben haben. Dies gilt für beide Aufgaben. Da in der Gruppe *Redundanz*

²⁰ Die Aufgabenstellung ist stark an die Aufgabenstellung der Unteraufgaben der entsprechenden Abituraufgabe angelehnt. Der erste Satz ist fast wortgleich übernommen (Brixius et al., 2012, S. 2008-13).

²¹ Das eingesetzte Material ist urheberrechtlich geschützt und darf daher hier nicht reproduziert werden. Eine grafische Demonstration des Materials, die die Variation zwischen den Gruppen zeigt, ist in Magnus et al. (2020) enthalten.

Informationen aus bildlichen Repräsentationen im Text wiederholt wurden, wurde, um die Textlänge und damit die Leselast zwischen den Gruppen konstant zu halten, der Text der Gruppe *keine Redundanz* durch moderate Umformulierungen unter weitgehendem Erhalt der Textschwierigkeit verlängert.²² Der Text des Aufgabenmaterials der Gruppe *nur Text* entsprach jeweils weitgehend dem der Gruppe *Redundanz*, wobei kleine Abweichungen dadurch begründet sind, dass im Material der Gruppe *Redundanz* (wie auch der Gruppe *keine Redundanz*) Verweise auf die Abbildungen enthalten waren, die im Material für Gruppe *nur Text* durch Umformulierungen entfernt wurden.

Die Aufgabenstellung von Aufgabe B lautete: „[E]rklären Sie die spezifischen Anpasstheiten der in stehenden Gewässern lebenden Feuersalamanderlarven im Kotttenforst“ (Brixius et al., 2012, S. 2009-1). Die bildlichen Darstellungen in Aufgabe B umfassten eine Tabelle mit vier Werten (bzgl. des Gewichts der Larven in verschiedenen Gewässertypen), ein Balkendiagramm mit zwei Balken (zum Energiegehalt von Nahrungsorganismen) und einen Graphen mit 13 Datenpunkten (in dem Metamorphosegewicht gegen Dauer der Larvenphase aufgetragen wurde). Sie wiesen eine geringere Informationsdichte auf als die Abbildungen von Aufgabe A und waren insgesamt weniger biologiespezifisch, da alle genutzten Abbildungsarten auch im Alltag und in anderen Disziplinen präsent sind. Der Text von Aufgabe B wurde in zwei Textblöcken präsentiert.

Anders als in den beiden anderen Gruppen mussten Studierende der Gruppe *keine Redundanz* zwingend relevante Informationen in Abbildungen identifizieren und mit den textlich dargebotenen Informationen in Beziehung setzen, um die Aufgaben beantworten zu können, da diese Informationen nicht im Text standen. Im Aufgabenmaterial der Gruppen, denen Text- und Bildmaterial zur Verfügung stand, konnten jeweils fünf für das Lösen der Aufgabe relevante Informationseinheiten aus den bildlichen Darstellungen entnommen werden, vier dem textlichen Material (in der Gruppe

²² Der Lesbarkeitsindex LIX zeigt keine bedeutsamen Unterschiede der Textschwierigkeit zwischen den Experimentalgruppen an. Für Aufgabe B wurden für den LIX nach Lenhard und Lenhard (2014–2017) für Gruppe *keine Redundanz*, *Redundanz* und *nur Text* die Werte 57.6, 57.7 und 57.7 ermittelt; damit wird die Komplexität als hoch eingestuft. Für Aufgabe A ergaben sich in gleicher Reihenfolge die LIX-Werte 57.0, 54.7 und 53.2. Den beiden letztgenannten Werten wird von Lenhard und Lenhard (2014–2017) noch mittlere Komplexität zugeschrieben, einem LIX-Wert von 57 wiederum hohe Komplexität. Diese geringfügig unterschiedlichen LIX-Werte zwischen den Experimentalgruppen gehen darauf zurück, dass in die LIX-Formel die Satzanzahl eingeht und diese in Aufgabe A nicht konstant gehalten wurde: In der Gruppe *keine Redundanz* ist ein langer Satz enthalten, während an gleicher Stelle im Text der anderen beiden Gruppen bei identischer Zeichenzahl zwei kurze Sätze enthalten sind. Dieser Unterschied erscheint insofern nicht als Beeinträchtigung der Validität, als der Satz in der Gruppe *keine Redundanz* keine lösungsrelevante Information enthält und zudem kritisch zu hinterfragen ist, ob ein längerer Satz mit weniger Information einen Text tatsächlich in jedem Fall schwieriger werden lässt als zwei kürzere Sätze mit mehr Information.

Redundanz zuzüglich der fünf im Text wiederholten ursprünglich bildlichen Informationseinheiten). Pro Aufgabe konnten die Studierenden somit Testleistungen im Wertebereich von 0 bis 9 erzielen; die Punktzahl zeigt an, wie viele der relevanten Informationen aus dem Material korrekt in die offene Antwort eingebunden wurden. Die Bewertung erfolgte für jede Aufgabe anhand eines dafür entwickelten Kategoriensystems. Die Objektivität der Bewertung wurde durch ein Interrating mit je Aufgabe 25 Prozent zufällig ausgewählten Antworttexten der Ursprungsstichprobe sichergestellt (vgl. Magnus et al., 2020). Für das Interrating berechnete Cohens-Kappa-Werte lagen über beide Aufgaben mit Werten zwischen $\kappa = .62$ und $\kappa = 1.0$ in einem nach Landis und Koch (1977) als substantiell bis fast perfekt eingeordneten Bereich der Übereinstimmung. Für Aufgabe B trat für eine der neun Informationseinheiten ein niedrigerer Kappa-Wert ($\kappa = .42$) trotz hoher Übereinstimmung (89,1 %) auf, was darauf zurückgeht, dass sich die Häufigkeiten der Ausprägungen der Kategorie erheblich unterscheiden. Der Kappa-Wert ist in solchen Fällen nicht sinnvoll interpretierbar.

5.2.4 Regressionsanalysen

Die Hypothesen wurden anhand multipler Regressionsanalysen unter Nutzung der Software IBM SPSS statistics (Version 24) geprüft. Da die Variable Experimentalgruppe kategorial ist, wurden zwei Dummy-Variablen erstellt; die Gruppe *nur Text* fungierte als Referenzgruppe. Zur Überprüfung der Moderationshypothese wurden zwei Interaktionsterme gebildet, indem jede Dummy-Variable mit dem z -standardisierten Prädiktor Selbstwirksamkeitserwartung multipliziert wurde; die fünf Terme wurden gleichzeitig in das Modell eingefügt (Aiken & West, 1991). In zwei weiteren Modellen wurden die letzte in der Schule erhaltene Biologienote bzw. die kognitiven Fähigkeiten als zusätzliche Prädiktorvariablen für die Testleistung einbezogen. Die Benjamini-Hochberg-Prozedur wurde mit einem q -Wert von .05 auf die Gesamtheit der Tests angewendet, um den multiplen Tests Rechnung zu tragen und so den erwarteten Anteil falschpositiver Aussagen zu kontrollieren. Die adjustierten p -Werte werden für die Fälle berichtet, in welchen die Fehlerkontrolle zu einer anderen Entscheidung führt als der zugrunde liegende Test.

5.3 Ergebnisse

In Tabelle 5.3.1 sind die Mittelwerte und Standardabweichungen aller in den Regressionsmodellen berücksichtigten Variablen für die drei Experimentalgruppen angegeben. Mittels einfaktorieller Varianzanalyse wurde überprüft, ob sich die Selbstwirksamkeitserwartung zwischen den Gruppen unterschied. Es wurde kein statistisch signifikanter Unterschied beobachtet, $F(2, 202) = 1.89, p = .153$.

Tabelle 5.3.1

Mittelwerte und Standardabweichungen nach Experimentalgruppe

Variable	Gruppe <i>keine Redundanz</i>		Gruppe <i>Redundanz</i>		Gruppe <i>nur Text</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Testleistung Aufgabe A gesamt	3.58	2.49	4.76	2.90	6.31	2.41
Testleistung Aufgabe A Score_Abb	1.88	1.54	2.97	1.56	3.60	1.34
Testleistung Aufgabe A Score_Text	1.70	1.23	1.79	1.53	2.71	1.35
Testleistung Aufgabe B gesamt	5.91	2.13	6.23	1.64	5.77	1.62
Selbstwirksamkeitserwartung	3.02	0.51	3.13	0.57	3.20	0.48
Letzte Biologienote	1.68	0.72	1.77	0.68	1.77	0.63
Kognitive Fähigkeiten	14.65	4.16	15.15	4.62	14.76	4.51

Anmerkung. Für Aufgabe A erfolgte zusätzlich zur Betrachtung der Testleistung der gesamten Aufgabe eine separate Betrachtung der Testleistung für Informationseinheiten, die aus Abbildungen entnommen werden mussten bzw. konnten (in Gruppe *keine Redundanz* und *Redundanz*, in Gruppe *nur Text* wurden sie hingegen rein textlich dargeboten; Score_Abb), und Informationseinheiten, die in allen drei Gruppen ausschließlich im Text gegeben waren (Score_Text). Für Aufgabe B erfolgte diese separate Betrachtung nicht.

5.3.1 Multiple Regression für Aufgabe A

Die Ergebnisse der multiplen Regression mit Experimentalgruppe als Moderator (Gruppe *nur Text* als Referenzgruppe) und Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie (SWE) als Prädiktor zeigten, dass dieses Modell einen statistisch signifikanten Erklärungsbeitrag zur Varianz der Testleistung in Aufgabe A leistete (korrigiertes $R^2 = .192$; $R^2 = .212$, $F(5, 199) = 10.69, p < .001$). Der für Dummy 1 geschätzte Regressionskoeffizient repräsentiert

den Unterschied in der mittleren Testleistung zwischen Gruppe *keine Redundanz* und Gruppe *nur Text* (Referenzgruppe) zwischen Studierenden bei gleichermaßen dem Mittelwert entsprechend ausgeprägter SWE. Dieser wies einen statistisch signifikanten Effekt von $B = -2.437, p < .001$, auf die Testleistung aus, d. h. die Testleistung war um ca. 2.4 (von 9) Punkte geringer, wenn Informationen aus Text und Bild entnommen werden mussten und nicht rein textlich dargeboten wurden. Auch der den Unterschied zwischen Gruppe *Redundanz* und Gruppe *nur Text* repräsentierende geschätzte Regressionskoeffizient für Dummy 2 zeigte einen statistisch signifikanten negativen Effekt auf die Testleistung, $B = -1.378, p = .002$ – auch die Gruppe *Redundanz* erzielte eine geringere Punktzahl als die Gruppe *nur Text* (die Effekte des Testmaterials auf die Testleistung sind Gegenstand von Magnus et al., 2020). Darüber hinaus wurde für die SWE ein statistisch signifikanter positiver Effekt auf die Testleistung beobachtet, $B = 1.087, p = .002$, welcher den Effekt (simple slope) in der Referenzgruppe *nur Text* repräsentiert: Eine höhere SWE ging mit einer besseren Testleistung einher. Diese Beziehung zwischen SWE und Testleistung in Aufgabe A fand sich nicht unabhängig von der Experimentalgruppenzugehörigkeit, was durch den statistisch signifikanten Effekt des Interaktionsterms Dummy 2 \times SWE ersichtlich wurde, $B = -1.169, p = .009$. Für Studierende, die mit redundantem Material arbeiteten, wurde ein statistisch signifikant vom Effekt für die Referenzgruppe *nur Text* abweichender Zusammenhang von SWE und Testleistung beobachtet ($1.087 - 1.169 = -0.082$). Der Interaktionsterm Dummy 1 \times SWE war nicht statistisch signifikant, $B = -0.386, p = .415$; das heißt, für die Gruppe *keine Redundanz* wurde ein Effekt der SWE auf die Testleistung beobachtet, welcher nicht vom Effekt in der Gruppe *nur Text* verschieden ist.

Um den beobachteten Moderationseffekt genauer zu untersuchen, wurden analog zum Ausgangsmodell mit Referenzgruppe *nur Text* die bedingten einfachen Regressionsgewichte (simple slopes) für den Effekt von SWE auf die Testleistung in Aufgabe A für die beiden Experimentalgruppen auf statistische Signifikanz überprüft, in welchen zusätzlich zum Text fachspezifische komplexe Abbildungen dargeboten wurden (Abbildung 5.3.1). Dazu wurden neue Dummy-Variablen und Interaktionsterme erstellt, wobei einmal Gruppe *keine Redundanz* und einmal Gruppe *Redundanz* als Referenzgruppe fungierte. Für Gruppe *keine Redundanz* ergab sich ein signifikanter Effekt von $B = 0.701, p = .030$, für Gruppe *Redundanz* erwies sich wie oben beschrieben kein bedeutsamer Zusammenhang von SWE und Testleistung mit $B = -0.082, p = .769$. Für die Gruppen *keine Redundanz* und *nur Text* (zu Gruppe *nur*

Studie 2

Text als Referenzgruppe s. oben: $B = 1.087$, $p = .002$) besteht somit anders als für Gruppe *Redundanz* eine statistisch signifikante positive Beziehung zwischen SWE und Testleistung (Abbildung 5.3.1).

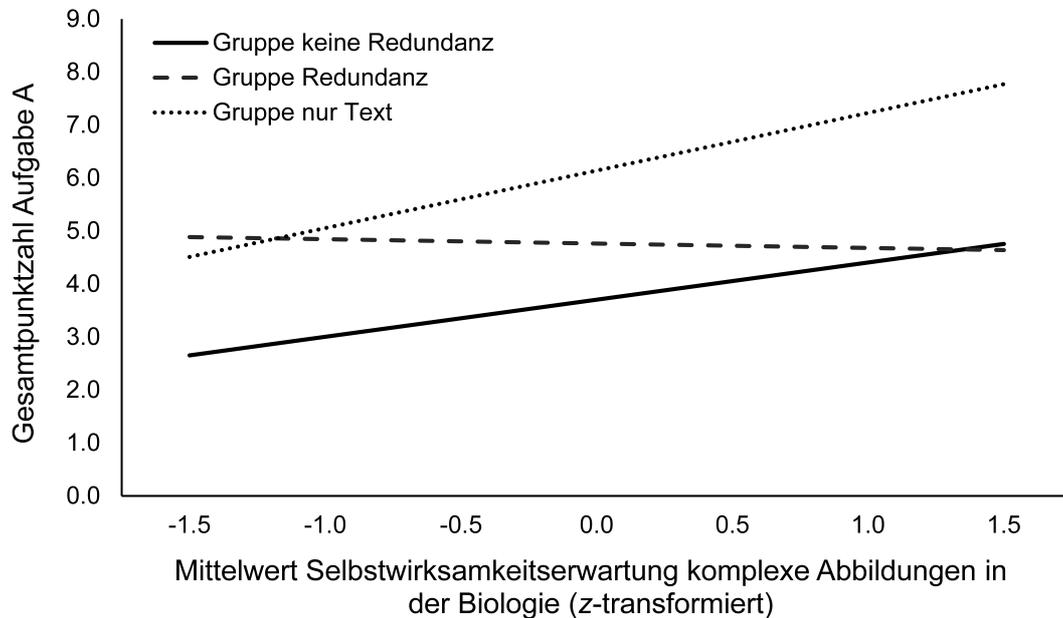


Abbildung 5.3.1. Testleistung in Aufgabe A als Funktion von Selbstwirksamkeitserwartung und Experimentalgruppe (simple slopes).

5.3.1.1 Berücksichtigung des Repräsentationstyps

Im Folgenden wurde die Testleistung in Aufgabe A zerlegt in die Informationseinheiten, die in der Gruppe *keine Redundanz* und *Redundanz* aus Abbildungen entnommen werden mussten bzw. konnten (diese wurden in der Gruppe *nur Text* rein textlich dargeboten; $Score_Abb$), und die Informationseinheiten, die in allen drei Gruppen nur im Text gegeben waren ($Score_Text$). Die oben dargestellten Regressionsanalysen wurden mit diesen neuen Kriteriumsvariablen wiederholt. So konnte unter anderem untersucht werden, ob sich der Effekt der SWE auf die Testleistung in Gruppe *keine Redundanz* in erster Linie auf die Testleistung bezüglich der Informationsentnahme aus Abbildungen bezog. Wie für die gesamte Aufgabe A zeigte sich, dass dieses Modell einen statistisch signifikanten Anteil der Varianz in der Testleistung aufklärt, korrigiertes $R^2 = .217$, $R^2 = .236$, $F(5, 199) = 12.31$, $p < .001$ für $Score_Abb$ und korrigiertes $R^2 = .127$, $R^2 = .148$, $F(5, 199) = 6.94$, $p < .001$ für $Score_Text$. Der Unterschied in der Testleistung zwischen Gruppe *keine Redundanz* und Gruppe

nur Text repräsentierende geschätzte Regressionskoeffizient für Dummy 1 zeigte wie für den Gesamtscore von Aufgabe A einen signifikanten Effekt auf *Score_Abb* und *Score_Text*, $B = -1.555, p < .001$ bzw. $B = -0.881, p < .001$. Auch der den Unterschied zwischen Gruppe *Redundanz* und Gruppe *nur Text* repräsentierende geschätzte Regressionskoeffizient für Dummy 2 wies einen statistisch signifikanten Effekt auf die Testleistung auf, $B = -0.543, p = .031$ für *Score_Abb* und $B = -0.836, p < .001$ für *Score_Text*. Der Regressionskoeffizient für die SWE betrug $B = 0.532, p = .008$ (für *Score_Abb*) und $B = 0.555, p = .003$ (für *Score_Text*) – der statistisch signifikante Einfluss der SWE auf die Subscores war in der Gruppe *nur Text* nahezu gleich groß. Eine höhere SWE ging mit einer besseren Testleistung einher. Wie für den Gesamtscore der Aufgabe zeigte sich ein statistisch signifikanter Effekt des Interaktionsterms *Dummy 2* \times SWE, $B = -0.572, p = .025$ für *Score_Abb* bzw. $B = -0.597, p = .012$ für *Score_Text*. Auch bei einzelner Betrachtung der Informationseinheiten aus Abbildungen oder Texten wurde die Beziehung zwischen SWE und der Testleistung in Aufgabe A also durch die Experimentalgruppenzugehörigkeit der Studierenden moderiert; in der Gruppe *Redundanz* war die Testleistung unabhängig von der SWE. Der Interaktionsterm *Dummy 1* \times SWE hatte keinen statistisch signifikanten Effekt auf die Testleistung, $B = -0.090, p = .739$ für *Score_Abb* bzw. $B = -0.296, p = .238$ für *Score_Text*.

Die bedingten einfachen Regressionsgewichte zeigten wiederum lediglich für Gruppe *keine Redundanz* einen statistisch signifikanten Effekt von $B = 0.442, p = .016$ auf den *Score_Abb*, für Gruppe *Redundanz* erwies sich dieser Zusammenhang als nicht statistisch signifikant, $B = -0.040, p = .802$. Für die Gruppen *keine Redundanz* und *nur Text* besteht somit eine statistisch signifikante positive Beziehung zwischen der SWE bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie und der Entnahme von Informationen aus Abbildungen (*Score_Abb*), während dieser Zusammenhang für die Gruppe *Redundanz* nicht vorliegt. Für die Entnahme von Informationen aus dem Text gibt es lediglich für die Gruppe *nur Text* einen statistisch signifikanten Zusammenhang zwischen SWE und Testleistung (*Score_Text*), beide bedingten einfachen Regressionsgewichte für den Effekt der SWE auf *Score_Text* für die Gruppen *keine Redundanz* bzw. *Redundanz* waren nicht statistisch signifikant, $B = 0.259, p = .130$ bzw. $B = -0.042, p = .776$.

5.3.1.2 Berücksichtigung objektiver Leistungsmaße

Um Effekte objektiver Leistungsmaße auf die Testleistung zu untersuchen und um zu prüfen, ob die SWE nach Berücksichtigung dieser auf die Testleistung weiterhin einen eigenständigen Beitrag zur Varianzaufklärung leistet, wurden die letzte in der Schule erhaltene Biologienote und allgemeine kognitive Fähigkeiten nacheinander ins Modell eingefügt (Modell 2 und 3, Tabelle 5.3.2). Die Biologienote klärte mit 2,1 % ($F(1, 198) = 5.36, p = .022$) einen statistisch signifikanten zusätzlichen Varianzanteil am Gesamtscore von Aufgabe A auf (korrigiertes $R^2 = .209$; $R^2 = .233, F(6, 198) = 10.00, p < .001$). Für den Score_Abb führte das Einfügen der Biologienote ins Modell nicht zu einer statistisch signifikanten Änderung des $R^2, \Delta R^2 = .010 (F(1, 198) = 2.69, p = .102)$. Für den Score_Text zeigte sich hingegen wie für den Gesamtscore eine statistisch signifikante Änderung des R^2 um $.028 (F(1, 198) = 6.76, p = .010)$. Dieses Modell klärte insgesamt einen statistisch signifikanten Anteil der Testleistung für Score_Text auf, korrigiertes $R^2 = .152, R^2 = .177, F(6, 198) = 7.08, p < .001$. Wurde zusätzlich der Prädiktor allgemeine kognitive Fähigkeiten ins Modell eingefügt (korrigiertes $R^2 = .238; R^2 = .264, F(7, 197) = 10.08, p < .001$), wurden weitere 3,1 % ($F(1, 197) = 8.33, p = .004$) der Varianz der Testleistung (Gesamtscore) in Aufgabe A erklärt. Auch für Score_Abb und Score_Text trugen die kognitiven Fähigkeiten statistisch signifikant zur Varianzaufklärung bei, $\Delta R^2 = .016 (F(1, 197) = 4.32, p = .039^{23})$ für Score_Abb (korrigiertes $R^2 = .236; R^2 = .263, F(7, 197) = 10.03, p < .001$) und $\Delta R^2 = .041 (F(1, 197) = 10.26, p = .002)$ für Score_Text (korrigiertes $R^2 = .190; R^2 = .217, F(7, 197) = 7.82, p < .001$). Tabelle 5.3.2 zeigt den Einfluss sämtlicher Prädiktoren in den Modellen 2 und 3 auf die Testleistung (Gesamtscore, Score_Abb und Score_Text) in Aufgabe A. Im Vergleich zu Modell 1 änderte sich jeweils durch Hinzufügen der Prädiktoren Biologienote und kognitive Fähigkeiten der Einfluss der SWE und des Moderators Experimentalgruppe nur geringfügig. Erwartungsgemäß war der Zusammenhang von Biologienote und Testleistung positiv. Der Betrag des Regressionskoeffizienten war dabei jeweils nahezu halb so groß wie der Wert für die SWE (Tabelle 5.3.2). Für Score_Abb wurde der Effekt der Biologienote auf die Testleistung nicht statistisch signifikant. Wurden zusätzlich allgemeine kognitive Fähigkeiten im Modell berücksichtigt, zeigte der statistisch signifikante Effekt für alle drei Scores an,

²³ Die Änderung des R^2 durch Aufnahme des KFT in das Vorhersagemodell für das Kriterium Score_Abb war nach Korrektur mittels Benjamini-Hochberg-Prozedur nicht mehr statistisch signifikant.

dass bessere kognitive Fähigkeiten zu einer besseren Testleistung führten, während der Effekt der Biologienote nur noch für Score_Text statistisch signifikant war (Tabelle 5.3.2). Aber auch der Effekt der kognitiven Fähigkeiten auf die Testleistung war bei den drei betrachteten (Sub-)Scores jeweils deutlich geringer als der Effekt der SWE (Tabelle 5.3.2).

Tabelle 5.3.2

Regressionsmodelle 2 und 3 zur Vorhersage der Testleistung von Aufgabe A

	Modell 2			Modell 3		
	<i>B (SE)</i>			<i>B (SE)</i>		
	gesamt	Abbildungen	Texte	gesamt	Abbildungen	Texte
Dummy 1 (Gr. <i>nur Text</i> vs. Gr. <i>keine Redundanz</i>)	-2.547*** (.444)	-1.600*** (.255)	-0.947*** (.235)	-2.534*** (.436)	-1.595*** (.252)	-0.940*** (.229)
Dummy 2 (Gr. <i>nur Text</i> vs. Gr. <i>Redundanz</i>)	-1.400** (.434)	-0.551* (.249)	-0.849*** (.229)	-1.455*** (.426)	-0.574* (.247)	-0.881*** (.224)
SWE ^a	0.931** (.349)	0.469* (.200)	0.462* (.185)	0.870* (.344)	0.443* (.199)	0.427* (.181)
Interaktion 1 (Dummy 1 × SWE ^a)	-0.372 (.467)	-0.084 (.268)	-0.288 (.247)	-0.325 (.459)	-0.064 (.266)	-0.261 (.241)
Interaktion 2 (Dummy 2 × SWE ^a)	-1.066* (.442)	-0.530† (.253)	-0.536* (.234)	-1.030* (.434)	-0.515† (.251)	-0.515* (.228)
Biologienote ^a	-0.425* (.184)	-0.173 (.105)	-0.253* (.097)	-0.346 (.183)	-0.140 (.106)	-0.206* (.096)
Kognitive Fähigkeiten ^a				0.509** (.176)	0.212† (.102)	0.297** (.093)

Anmerkungen. Gr. = Experimentalgruppe; SWE = Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie.

^a(z-standardisiert).

* $p < .05$. ** $p < .01$. *** $p < .001$. † Nach Korrektur mittels Benjamini-Hochberg-Prozedur nicht mehr statistisch signifikant.

Studie 2

5.3.2 Multiple Regression für Aufgabe B

Das Regressionsmodell mit Experimentalgruppen und SWE sowie deren Interaktionstermen erklärte, anders als bei Aufgabe A, keinen statistisch signifikanten Varianzanteil in der Testleistung in Aufgabe B (korrigiertes $R^2 = -.010$; $R^2 = .015$, $F(5, 199) = 0.59$, $p = .708$). Weder der Unterschied zwischen Gruppe *keine Redundanz* und Gruppe *nur Text* (Dummy 1), noch der Unterschied zwischen Gruppe *Redundanz* und Gruppe *nur Text* (Dummy 2) hatten einen statistisch signifikanten Effekt auf die Testleistung, $B = 0.175$, $p = .584$ bzw. $B = 0.483$, $p = .125$. Ebenso zeigten sowohl die SWE ($B = 0.201$, $p = .418$) als auch die beiden Interaktionsterme keinen statistisch signifikanten Effekt auf die Testleistung. Für die Interaktion Dummy 1 \times SWE ergab sich $B = -0.174$, $p = .606$ und für die Interaktion Dummy 2 \times SWE war $B = -0.162$ und $p = .612$.

5.4 Diskussion

Die Ergebnisse dieser Studie legen nahe, dass das Lösen von Aufgaben mit fachspezifischen komplexen Abbildungen auch eine Frage der persönlichen Überzeugung ist hinsichtlich der eigenen Fähigkeit, die dadurch bedingten spezifischen Anforderungen zu bewältigen. Die Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie wies für Aufgabe A einen positiven Effekt auf die Testleistung auf. Diese Beziehung wurde durch die Experimentalgruppenzugehörigkeit der Studierenden moderiert. Der positive Effekt der Selbstwirksamkeitserwartung auf die Testleistung konnte wie erwartet bei Studierenden beobachtet werden, die die Abbildungen zum Lösen der Aufgabe heranziehen mussten (Gruppe *keine Redundanz*), nicht aber bei Studierenden, denen dieselben Abbildungen dargeboten wurden, die aber alle lösungsrelevanten Informationen aus dem Text entnehmen konnten (Gruppe *Redundanz*). Überraschend war die Beobachtung eines statistisch signifikanten Effekts für Gruppe *nur Text*. Die Ergebnisse zeigten zudem, dass die Selbsteinschätzung, mit der Personen komplexen Abbildungen begegnen, selbst dann einen Effekt auf die Testleistung hatte, wenn objektive Leistungsmaße wie die Biologienote oder allgemeine kognitive Fähigkeiten im Modell berücksichtigt wurden. Wie erwartet hatten diese beiden Variablen einen statistisch signifikanten Effekt auf die Testleistung (wenngleich diese bei der gegebenen Stichprobengröße nach Korrektur für multiple

Tests nicht mehr durchgängig statistisch signifikant waren), der Effekt der Selbstwirksamkeitserwartung blieb dabei aber nahezu unverändert groß. Als objektive Maße der kognitiven Fähigkeiten teilen sich KFT und Biologienote Varianzanteile, weshalb der Effekt der Biologienote nicht mehr statistisch signifikant war, als der KFT ins Modell eingefügt wurde.

Die Ergebnisse von Aufgabe B unterstützen die Annahme, dass sich die hier betrachtete Selbstwirksamkeitserwartung spezifisch auf das Arbeiten mit fachspezifischen, komplexen Abbildungen auswirkt. Es gab in Aufgabe B keinen statistisch signifikanten Effekt der Selbstwirksamkeitserwartung auf die Testleistung, was vermutlich daran lag, dass die Abbildungen der Aufgabe weniger fachspezifisch waren und wahrscheinlich als weniger komplex wahrgenommen wurden. Im Folgenden beschränken wir uns auf die Interpretation der Ergebnisse von Aufgabe A.

5.4.1 Moderatoreffekt

Für verschiedene Leistungen im naturwissenschaftlichen Bereich wurde eine entsprechende Selbstwirksamkeitserwartung als Einflussfaktor identifiziert (z. B. Ar dura & Pérez-Bitrián, 2019; Schütte et al., 2007). In dieser Studie konnten wir zeigen, dass sich eine Selbstwirksamkeitserwartung bezüglich des Arbeitens mit fachspezifischen, komplexen Abbildungen in der Biologie auf die Testleistung in Biologie auswirkt, wenn derartige Abbildungen zum Lösen der Aufgabe herangezogen werden müssen. Wie beschrieben werden Strategien zur Texterschließung in der Schule im Deutschunterricht geübt, während Strategien zum Erschließen von Informationen in Abbildungen möglicherweise nicht explizit gelehrt werden (Schönborn & Anderson, 2006; Rosebrock & Nix, 2010; Ziepprecht, 2016). Diese sollten aber gerade im naturwissenschaftlichen Fachunterricht gezielt gefördert werden, da naturwissenschaftliche Erkenntnisse über diverse fachspezifische, komplexe Abbildungen kommuniziert werden, die nicht selbsterklärend sind, sondern spezieller Lesefähigkeiten bedürfen (Lowe, 2000). Wie Ziepprecht (2016) für Schüler:innen der Sekundarstufe I zeigte auch unsere Studie mit Studierenden, dass es ihnen leichter fiel, Informationen aus Texten als aus mit Text kombinierten fachspezifischen Abbildungen zu entnehmen (Hauptgegenstand der Studie Magnus et al., 2020). Wie erwähnt sind Erfolgserfahrungen die wichtigste Quelle der Selbstwirksamkeitserwartung (Schwarzer & Jerusalem, 2002). Bleiben explizite Erfolgserfahrungen beim Erschließen von Informationen aus

Studie 2

fachspezifischen, komplexen Abbildungen in Ermangelung einer entsprechenden Förderung im Fachunterricht aus, wird die Selbstwirksamkeitserwartung in diesem Bereich vermutlich nicht ausreichend gestärkt. Eine mögliche Erklärung für unsere Befunde wäre also, dass Studierende, die in der Vergangenheit Erfolge beim Arbeiten mit komplexen Abbildungen erlebt haben, das notwendige Vertrauen in sich selbst und die nötige Motivation hatten, Aufgaben mit fachspezifischen, komplexen Abbildungen zu bearbeiten, und deshalb besser abschnitten als Studierende mit geringeren Erfolgserfahrungen. Dass sich der positive Effekt der Selbstwirksamkeitserwartung auf die Testleistung in Gruppe *keine Redundanz* ausschließlich auf die Testleistung bezüglich der aus Abbildungen und nicht der aus Text zu entnehmenden Informationen bezog, stützt unsere Annahme, dass die Selbstwirksamkeitserwartung bezüglich fachspezifischer, komplexer Abbildungen als ein eigenständiges Konstrukt behandelt werden sollte.

Die Ergebnisse von Gruppe *Redundanz* legen – wie bereits in Magnus et al. (2020) diskutiert – nahe, dass die Studierenden vor allem mit dem Text gearbeitet haben, da diese Gruppe gegenüber der Gruppe *nur Text* keine bessere Testleistung erzielte. Die Abbildungen wurden also nicht zwingend hinzugezogen, was auch nicht notwendig war, da in dieser Experimentalbedingung alle lösungsrelevanten Informationen im Text gegeben waren. Deshalb ist es schlüssig, dass bei Gruppe *Redundanz* keine Beziehung zwischen Selbstwirksamkeitserwartung und Testleistung vorlag.

Anders als erwartet wirkte sich bei Gruppe *nur Text*, obwohl die Studierenden nicht mit Abbildungen arbeiteten, die Selbstwirksamkeitserwartung positiv auf die Testleistung aus. Eine Erklärung könnte sein, dass für diese Studierenden anders als für Studierende der Gruppen, die Abbildungen im Testmaterial vorfanden, kein Anker für die Vorstellung von fachspezifischen, komplexen Abbildungen gesetzt worden war. Sie waren für die Bearbeitung der SWE-Skala allein auf ihren persönlichen Erfahrungshintergrund angewiesen. Möglicherweise liegen deren Urteilen allgemeinere Urteile über ihre Leistungen im Fach Biologie zugrunde.

5.4.2 Limitationen

Die Ergebnisse stützen die Annahme, dass die untersuchte Selbstwirksamkeitserwartung ein eigenständiges Konstrukt ist, das von einer Selbstwirksamkeitserwartung bezüglich allgemeiner Fähigkeiten in der Biologie (und bezogen auf andere fach-

spezifische Anforderungen) abgrenzbar ist. Die empirische Überprüfung dieser Annahme anhand entsprechender Selbstwirksamkeitsmaße steht jedoch noch aus. Des Weiteren ist unklar, welche Vorstellungen von fachspezifischen, komplexen Abbildungen in der Biologie Studierende der Gruppe *nur Text* hatten. Die Selbstwirksamkeitserwartung wurde gemeinsam mit Kontrollvariablen und demographischen Daten nach Bearbeitung der Biologieaufgaben abgefragt; die Gruppen, die mit fachspezifischen, komplexen Abbildungen gearbeitet haben, konnten also eine konkrete Vorstellung von solchen Abbildungen haben. Abgesehen davon ist aber nicht davon auszugehen, dass die Aufgabenbearbeitung die Einschätzung der Selbstwirksamkeitserwartung beeinflusst hat. Zum einen gilt die Selbstwirksamkeitserwartung, wenn sie einmal gebildet ist, als relativ stabil (Schwarzer & Jerusalem, 2002), zum anderen hatten die Studierenden zum Zeitpunkt der Bearbeitung der Selbstwirksamkeitserwartungs-Items keine Rückmeldung zu ihrem Abschneiden in den Testaufgaben bekommen.

5.4.3 Implikationen

Die Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie zeigte einen Effekt auf die Testleistung von Studierenden, wenn diese fachspezifische, komplexe Abbildungen zum Lösen der Aufgabe heranziehen mussten. Die Ergebnisse zeigten also, dass unter Kontrolle objektiver Leistungsmaße selbst Personen, die sich für ein Studium der Biologie entschieden haben, nicht zwangsläufig mit komplexen Abbildungen in der Biologie arbeiten konnten. Darauf, die entsprechende Selbstwirksamkeitserwartung bei Schüler:innen möglichst frühzeitig zu stärken und somit einen Zugang zum Lesen solcher Abbildungen zu erleichtern, sollten Lehrkräfte schon während der Schulzeit im naturwissenschaftlichen Unterricht hinarbeiten. Dazu sollten die Quellen der Selbstwirksamkeit in den Blick genommen werden und unter ihnen besonders die Erfolgserfahrungen. Zum Vermitteln von Erfolgen gehört das Unterstützen von Bewältigungsstrategien (Schwarzer & Jerusalem, 2002). Mehrere Studien haben sich bereits mit speziellen Anforderungen von fachspezifischen, komplexen Abbildungen an Lernende beschäftigt und Vorschläge ausgearbeitet, wie zur Bewältigung dieser Anforderungen beigetragen werden kann, um Lernende beim Lesen und Verstehen von Abbildungen zu unterstützen (Cromley et al., 2013; Schönborn & Anderson, 2006). Schönborn und Anderson (2006) erstellten zum Beispiel zehn Richtlinien für das Lehren und Lernen mit Abbildungen in der Biochemie. So empfehlen sie unter anderem, die spezielle Symbolik in

Studie 2

diesen Abbildungen explizit zu unterrichten und Lernende mit verschiedenen Repräsentationstypen zu demselben Phänomen zu konfrontieren, damit sie sich mit den einzelnen Repräsentationstypen kritisch auseinandersetzen und sich Gemeinsamkeiten und Unterschiede bewusst machen (Schönborn & Anderson, 2006). Auch das Konstruieren von Abbildungen durch Lernende kann zu einem routinierteren Arbeiten mit diesen führen. Scherb und Nitz (2020) zeigten in einer Interviewstudie mit (angehenden) Biologielehrkräften allerdings, dass diese im Unterricht den Fokus eher auf das Konstruieren realistischer anstatt abstrakter Abbildungen legen. Dazu passt, dass von Kotzebue, Gerstl und Nerdel (2015) beobachtet wurde, dass selbst Studierende Schwierigkeiten bei der Konstruktion von Diagrammen in biologischen Kontexten haben. In der Studie wurden Aspekte benannt, die Lehrkräfte beachten sollten, wenn sie Lernende im Arbeiten mit Diagrammen unterrichten. Zudem entwickelten Nitz, Prechtel und Nerdel (2014) ein Instrument, mit dem unter anderem die Wahrnehmung von Schüler:innen bezüglich der Verwendung bildlicher Repräsentationen im Biologieunterricht erfasst werden kann und das Lehrenden zur Reflexion dienen könnte. Es ist an Lehrenden, solche Richtlinien und Empfehlungen im Unterricht zu berücksichtigen und an Forschenden, weitere Richtlinien für Abbildungen in anderen Themengebieten zu entwickeln, sofern die Richtlinien nicht übertragbar sind. Im Sinne einer Weiterentwicklung der von diSessa und Sherin (2000) beschriebenen metarepräsentationalen Kompetenz sollte zudem das Wissen über Abbildungen in den Blick genommen werden, also zum Beispiel das Wissen über Funktionen und Grenzen bestimmter Abbildungsarten oder auch das eigenständige Entwickeln neuer Abbildungsarten durch Lernende.

Eine gestärkte Selbstwirksamkeitserwartung kann dazu führen, dass sich Personen eher auf das Arbeiten mit komplexen Abbildungen einlassen und mehr Mühe in die Bearbeitung investieren als Personen mit geringerer Selbstwirksamkeitserwartung. Wichtig ist, bei Lehrenden in den Naturwissenschaften in Schule und Universität ein Verständnis darüber zu erreichen, dass fachspezifische, komplexe Abbildungen für Lernende nicht unbedingt selbsterklärend sind und das Arbeiten mit solchen Abbildungen explizit unterrichtet werden sollte, was die entsprechende Selbstwirksamkeitserwartung von Lernenden stärkt.

5.5 Danksagung

Wir danken Safa Dadkhah Aseman und Christina Kelting für ihre Hilfe bei Interrating und Dateneingabe, allen Dozierenden, die uns freundlicherweise die Datenerhebung in ihren Kursen ermöglicht haben, sowie allen Studierenden, die bereit waren, an der Studie teilzunehmen.

5.6 Literatur

- Aiken, L. S. & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.
- Ardura, D. & Pérez-Bitrián, A. (2019). Motivational pathways towards academic achievement in physics & chemistry: a comparison between students who opt out and those who persist. *Chemistry Education Research and Practice*, 20, 618–632. <http://dx.doi.org/10.1039/c9rp00073a>
- Arsenault, D. J., Smith, L. D. & Beauchamp, E. A. (2006). Visual inscriptions in the scientific hierarchy: Mapping the “treasures of science”. *Science Communication*, 27, 376–428. <http://dx.doi.org/10.1177/1075547005285030>
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84, 191–215. <http://dx.doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (2001). Social cognitive theory: an agentic perspective. *Annual Review of Psychology*, 52, 1–26. <http://dx.doi.org/10.1146/annurev.psych.52.1.1>
- Bong, M. (1997). Generality of academic self-efficacy judgments: Evidence of hierarchical relations. *Journal of Educational Psychology*, 89, 696–709. <http://dx.doi.org/10.1037/0022-0663.89.4.696>
- Brixius, R., Jannan, M. & Kunze, H. (2012). *Abitur; Prüfungsaufgaben mit Lösungen 2013; NRW; Biologie LK (6. Aufl.)*. Hallbergmoos: Stark.
- Cromley, J. G., Perez, T. C., Fitzhugh, S. L., Newcombe, N. S., Wills, T. W. & Tanaka, J. C. (2013). Improving students’ diagram comprehension with classroom instruction. *The Journal of Experimental Education*, 81, 511–537. <http://dx.doi.org/10.1080/00220973.2012.745465>
- Cromley, J. G., Snyder-Hogan, L. E. & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology*, 35, 59–74. <http://dx.doi.org/10.1016/j.cedpsych.2009.10.002>
- diSessa, A. A. & Sherin, B. L. (2000). Meta-representation: an introduction. *The Journal of Mathematical Behavior*, 19, 385–398. [http://dx.doi.org/10.1016/S0732-3123\(01\)00051-7](http://dx.doi.org/10.1016/S0732-3123(01)00051-7)
- Eilam, B. & Poyas, Y. (2008). Learning with multiple representations: Extending multimedia learning beyond the lab. *Learning and Instruction*, 18, 368–378. <http://dx.doi.org/10.1016/j.learninstruc.2007.07.003>

- Heller, K. A. & Perleth, C. (2000). *KFT 4-12+R - Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz Test GmbH, Hogrefe.
- Heßler, M. (2006). Einleitung. Annäherungen an Wissenschaftsbilder. In M. Heßler (Hrsg.), *Konstruierte Sichtbarkeiten. Wissenschafts- und Technikbilder seit der Frühen Neuzeit* (S. 11–37). München: Fink.
- Jerusalem, M. & Satow, L. (1999). Schulbezogene Selbstwirksamkeitserwartung. In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen* (S. 15–16). Berlin: Freie Universität Berlin.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. <http://dx.doi.org/10.2307/2529310>
- Lemke, J. L. (1998). *Teaching all the languages of science: words, symbols, images, and actions*. Paper presented at the International Conference on Ideas for a Scientific Culture, Barcelona, Spanien.
- Lenhard, W., & Lenhard, A. (2014–2017). *Berechnung des Lesbarkeitsindex LIX nach Björnson*. Zugriff am 07.07.2021 unter <http://www.psychometrica.de/lix.html>. Bibergau: Psychometrica. <http://dx.doi.org/10.13140/RG.2.1.1512.3447>
- Lindner, M. A., Ihme, J. M., Saß, S. & Köller, O. (2018). How representational pictures enhance students' performance and test-taking pleasure in low-stakes assessment. *European Journal of Psychological Assessment*, *34*, 376–385. <http://dx.doi.org/10.1027/1015-5759/a000351>
- Lowe, R. (2000). *Visual literacy and learning in science*. Zugriff am 27.07.2020 unter <http://eric.ed.gov/?id=ED463945>
- Magnus, L., Schütte, K. & Schwanewedel, J. (2020). Challenges solving science tasks with text–picture combinations persist beyond secondary school. *Journal of Research on Educational Effectiveness*, *13*, 759–783. <http://dx.doi.org/10.1080/19345747.2020.1750744>
- NGSS Lead States (2013). *Next generation science standards: for states, by states*. Washington, DC: The National Academies Press.
- Nitz, S., Precht, H. & Nerdel, C. (2014). Survey of classroom use of representations: development, field test and multilevel analysis. *Learning Environments Research*, *17*, 401–422. <http://dx.doi.org/10.1007/s10984-014-9166-x>

Studie 2

- Novick, L. R., Stull, A. T. & Catley, K. M. (2012). Reading phylogenetic trees: The effects of tree orientation and text processing on comprehension. *BioScience*, 62, 757–764. <http://dx.doi.org/10.1525/bio.2012.62.8.8>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66, 543–578. <http://dx.doi.org/10.3102/00346543066004543>
- Rosebrock, C. & Nix, D. (2010). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung* (3. Aufl.). Baltmannsweiler: Schneider.
- Roth, W.-M. & Pozzer-Ardenghi, L. (2013). Pictures in biology education. In D. F. Treagust & C.-Y. Tsui (Eds.), *Multiple representations in biological education* (Vol. 7, pp. 39–53). Dordrecht: Springer. http://dx.doi.org/10.1007/978-94-007-4192-8_3
- Scherb, C. A. & Nitz, S. (2020). Attitudes of biology teachers towards learner-generated external visual representations. *Research in Science Education*, 50, 2533–2558. <https://dx.doi.org/10.1007/s11165-018-9792-x>
- Schönborn, K. J. & Anderson, T. R. (2006). The importance of visual literacy in the education of biochemists. *Biochemistry and Molecular Biology Education*, 34, 94–102. <http://dx.doi.org/10.1002/bmb.2006.49403402094>
- Schütte, K., Frenzel, A. C., Asseburg, R. & Pekrun, R. (2007). Schülermerkmale, naturwissenschaftliche Kompetenz und Berufserwartung. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 125–146). Münster: Waxmann.
- Schwarzer, R. & Jerusalem, M. (2002). Das Konzept Selbstwirksamkeit. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (S. 28–53), Zeitschrift für Pädagogik, 44. Beiheft. Weinheim: Beltz.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. München: Luchterhand.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2020). *Bildungsstandards im Fach Biologie für die Allgemeine Hochschulreife. Beschluss der Kultusministerkonferenz vom*

18.06.2020. Zugriff am 19.03.2021 unter http://www.kmk.org/fileadmin/Daten/veroeffentlichungen_beschluesse/2020/2020_06_18-Bildungsstandard-sAHR_Biologie.pdf

- Seufert, T. (2003). Kohärenzbildung beim Wissenserwerb mit multiplen Repräsentationen. In K. Sachs-Hombach (Hrsg.), *Was ist Bildkompetenz? Studien zur Bildwissenschaft* (S. 117–129). Wiesbaden: Deutscher Universitätsverlag. http://dx.doi.org/10.1007/978-3-663-11814-5_8
- Trafton, J. G., Marshall, S., Mintz, F. & Trickett, S. B. (2002). Extracting explicit and implicit information from complex visualizations. In M. Hegarty, B. Meyer & N. H. Narayanan (Eds.), *Diagrammatic representation and inference. Diagrams 2002. Lecture notes in computer science* (Vol. 2317, pp. 206–220). Berlin: Springer. http://dx.doi.org/10.1007/3-540-46037-3_22
- von Kotzebue, L., Gerstl, M. & Nerdel, C. (2015). Common mistakes in the construction of diagrams in biological contexts. *Research in Science Education*, 45, 193–213. <http://dx.doi.org/10.1007/s11165-014-9419-9>
- Ziepprecht, K. (2016). *Strategien und Kompetenzen von Lernenden beim Erschließen von biologischen Informationen aus unterschiedlichen Repräsentationen*. Berlin: Logos.

6 Zusammenfassung der Ergebnisse der Studien

Studie 1: Challenges Solving Science Tasks with Text–Picture Combinations Persist beyond Secondary School

Im Fokus dieser Studie stand das Arbeiten mit redundanten und nicht-redundanten Text-Bild-Kombinationen (d. h. MER) in der Biologie. Schüler:innen sollten prinzipiell durch naturwissenschaftlichen Unterricht befähigt werden, mit jeder Art von fachtypischer Text-Bild-Kombination arbeiten zu können. In dieser Studie bearbeiteten $N = 240$ Biologiestudierende im Bachelorstudium, die die Schullaufbahn mit Erwerb der Hochschulzugangsberechtigung absolviert hatten, zwei auf Abituraufgaben basierende Biologieaufgaben aus den Themengebieten Molekulargenetik bzw. Ökologie und Evolution. Ziel war es, die Effekte von Redundanz in MER auf die Testleistung zu untersuchen sowie herauszufinden, ob die Studierenden von Abbildungen in MER profitierten. Es wurde Material für drei Experimentalgruppen erstellt: Gruppe *keine Redundanz* arbeitete mit MER, in denen lösungsrelevante Informationen auf Abbildungen und Texte aufgeteilt waren. Im Material für Gruppe *Redundanz* wurden lösungsrelevante Informationen aus den Abbildungen im Text wiederholt. Gruppe *nur Text* arbeitete mit rein textlichem Material, das alle lösungsrelevanten Informationen enthielt.

Die Ergebnisse für die Molekulargenetik-Aufgabe zeigen, dass Studierende, die zum Lösen der Aufgabe Informationen aus Abbildungen und Texten entnehmen und integrieren mussten (Gruppe *keine Redundanz*), die geringste Punktzahl erzielten. Wurden lösungsrelevante Informationen im Text wiederholt (Gruppe *Redundanz*), wurde eine statistisch signifikant höhere Punktzahl erzielt. Die höchste Punktzahl erreichten allerdings Studierende der Gruppe *nur Text*, deren Testmaterial keine Abbildungen enthielt. Gruppe *nur Text* zeigte auch dann jeweils die beste Testleistung, wenn Informationen separat betrachtet wurden, die in allen drei Gruppen ausschließlich im Text vorlagen oder in den MER-Bedingungen in den Abbildungen enthalten waren (und in Gruppe *Redundanz* im Text wiederholt wurden, in Gruppe *nur Text* ausschließlich textlich vorlagen). Die Molekulargenetik-Aufgabe umfasste komplexe, biologiespezifische Abbildungen. Dagegen wiesen die depiktionalen Repräsentationen der Ökologie/Evolution-Aufgabe eine geringere Informationsdichte auf und waren nicht per se biologiespezifisch, sondern auch aus anderen Fächern oder durch Präsenz im gesellschaftlichen Leben bekannte Repräsentationstypen. Dass die Art der im Material enthaltenen Abbildungen Einfluss auf die Bearbeitung durch

die Studierenden hatte, wurde dadurch deutlich, dass in der Aufgabe aus den Themengebieten Ökologie und Evolution keine Unterschiede der Testleistung zwischen den Experimentalgruppen beobachtet wurden.

Assoziierte Masterarbeit (Klann, 2018)

In Studie 1 wurden in einer quantitativen experimentellen Studie Effekte von verschieden gestalteten Testmaterialien auf die Testleistung Studierender untersucht. Die Methodik der Studie ließ dabei keine Rückschlüsse auf den Prozess der Aufgabenbearbeitung der Studierenden zu, der besonders in Bezug auf das Erschließen von lösungsrelevanten Informationen aus Abbildungen oder Texten wichtige Erkenntnisse liefern könnte. Deshalb wurde durch ein Masterarbeitsprojekt, das im Rahmen des vorliegenden Promotionsprojekts betreut wurde, eine qualitative Studie angeschlossen, die mit dem Testmaterial der Text-Bild-Gruppen (Gruppe *keine Redundanz* und Gruppe *Redundanz*) den Prozess der Aufgabenbearbeitung mittels Eye-Tracking- und Thinking-Aloud-Methode in den Blick nahm (Klann, 2018). Zusätzlich wurden die Teilnehmenden befragt, wie sie die einzelnen Aufgaben und Abbildungen wahrgenommen haben. Die Studie wurde mit $N = 6$ Biologiestudierenden im Bachelorstudium aufgeteilt auf die zwei Gruppen durchgeführt.

Zwar können aufgrund der geringen Stichprobengröße keine verallgemeinernden Aussagen getroffen werden, es wurden aber folgende Beobachtungen gemacht: Studierende der Gruppe *Redundanz*, denen lösungsrelevante Informationen aus Abbildungen auch im Text geboten wurden, zeigten in der Molekulargenetik-Aufgabe eine höhere Anzahl von Sakkaden (Augenbewegungen zwischen Abbildungen und zugehörigen Texten) als Studierende der Gruppe *keine Redundanz*. Sie erreichten wie in der quantitativen Studie eine höhere Punktzahl als Gruppe *keine Redundanz* bei ähnlicher Bearbeitungszeit. In Abbildung und Text dargebotene redundante Informationen wurden dabei meistens zunächst dem Text entnommen und anschließend vom Text geleitet zum Teil in der entsprechenden Abbildung gefunden. Es kann festgehalten werden, dass Studierende beider Gruppen sowohl die Abbildungen als auch die Texte fokussierten, die jeweilige Verweildauer auf einzelnen Repräsentationen war aber zwischen den Teilnehmenden individuell sehr unterschiedlich. Der Bearbeitungsprozess der Aufgabe aus den Themengebieten Ökologie und Evolution unterschied sich nicht zwischen den Gruppen. Es fiel auf, dass Studierende beider Gruppen meistens stärker die depiktionalen Repräsentationen als die Texte fokussierten.

Studierende beider Gruppen empfanden die Molekulargenetik-Aufgabe als schwieriger als die Ökologie/Evolution-Aufgabe. Ein häufig genannter Grund für die Schwierigkeit war die Komplexität der Abbildungen, während in der Ökologie/Evolution-Aufgabe die relativ gering eingeschätzte Schwierigkeit mitunter sogar durch die Übersichtlichkeit der Abbildungen begründet wurde. Bei der jeweiligen Bewertung der Komplexität mittels eines 5-stufigen Items (1 = sehr einfach; 5 = sehr komplex) wurden die drei depiktionalen Repräsentationen der Molekulargenetik-Aufgabe im Mittel mit 3,2 bewertet (3,7; 2,0; 4,0), die drei der zweiten Aufgabe im Mittel mit 1,5 (1,3; 1,3; 1,8).

Studie 2: Fachspezifische, komplexe Abbildungen verarbeiten – (auch) eine Frage der Überzeugung

Im Fokus dieser Studie standen Effekte einer Selbstwirksamkeitserwartung bezüglich des Arbeitens mit fachspezifischen, komplexen Abbildungen auf das Arbeiten mit derartigen Abbildungen. Es gibt zahlreiche Studien, die einen Zusammenhang zwischen Selbstwirksamkeitserwartung – also der Überzeugung, durch eigenes Handeln bestimmte Ergebnisse erzielen zu können – und Leistungen im naturwissenschaftlichen Bereich aufzeigen. In der vorliegenden Studie sollte die Annahme untersucht werden, dass die Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie einen positiven Effekt auf die Testleistung Biologiestudierender hat, wenn sie derartige Abbildungen zum Lösen einer Aufgabe heranziehen müssen. Die Studierenden schätzten ihre Selbstwirksamkeitserwartung auf einer Likert-Skala ein. Das Testmaterial wurde bezüglich der Anforderungen an das Arbeiten mit Abbildungen zwischen drei Gruppen variiert. Die Stichprobe und das Aufgabenmaterial für die Experimentalgruppen (*keine Redundanz*, *Redundanz*, *nur Text*) wurden bereits in der Zusammenfassung von Studie 1 beschrieben. Durch Ausschluss der Teilnehmenden, die notwendige Skalen nicht vollständig bearbeitet hatten, ergab sich ein Stichprobenumfang von $N = 205$ Studierenden.

Wie erwartet wurde der Effekt der Selbstwirksamkeitserwartung durch die Experimentalgruppenzugehörigkeit moderiert. Für Studierende, die zum Lösen der Molekulargenetik-Aufgabe auf Informationen aus fachspezifischen, komplexen Abbildungen angewiesen waren (Gruppe *keine Redundanz*), wurde unter Kontrolle objektiver Leistungsmaße (kognitive Fähigkeiten, Biologienote) ein positiver Effekt der Selbstwirksamkeitserwartung auf die Testleistung nachgewiesen. Dieser Effekt bezog sich allein auf die aus Abbildungen zu entnehmenden Informationen und nicht auf textlich dargebotene. Wenn Studie-

Zusammenfassung der Ergebnisse der Studien

rende die lösungsrelevanten Informationen der Abbildungen zusätzlich aus dem Text entnehmen konnten (Gruppe *Redundanz*), lag kein Effekt vor. Überraschenderweise wurde für die Gruppe, die ausschließlich mit textlichem Material arbeitete (Gruppe *nur Text*), ein positiver Effekt der Selbstwirksamkeitserwartung auf die Testleistung beobachtet, der sich vermutlich durch das Fehlen eines Ankers für die Vorstellung von fachspezifischen, komplexen Abbildungen begründen lässt. Für die Ökologie/Evolution-Aufgabe, deren depiktionale Repräsentationen weniger komplex und weniger biologiespezifisch waren, zeigte sich kein Effekt der Selbstwirksamkeitserwartung auf die Testleistung. Insgesamt stützen die Ergebnisse die Annahme, dass die untersuchte Selbstwirksamkeitserwartung als eigenständiges Konstrukt von einer Selbstwirksamkeitserwartung bezüglich allgemeiner Fähigkeiten in der Biologie abgrenzbar ist.

7 Diskussion und Ausblick

7.1 Studienübergreifende Diskussion der Ergebnisse

In der vorliegenden Arbeit wurden Erkenntnisse darüber gewonnen, wie gut Personen nach erfolgreicher Beendigung der Schullaufbahn mit Erwerb der Hochschulzugangsberechtigung, die sich für das Studienfach Biologie entschieden haben, in der Lage sind, mit verschiedenartigen Text-Bild-Kombinationen (d. h. MER) in der Biologie zu arbeiten, und wodurch auftretende Schwierigkeiten begründet sein könnten. Die Arbeit leistet somit einen Beitrag zur evidenzbasierten Forschung zu Repräsentationen in den Naturwissenschaften, in denen MER omnipräsent sind und in deren assoziierten Schulfächern das Erlernen eines kompetenten Arbeitens mit MER auch aus wissenschaftspropädeutischer Sicht unerlässlich ist. Die Ergebnisse der einzelnen Studien wurden in den Kapiteln 4 und 5 ausführlich diskutiert. Im Folgenden sollen sie studienübergreifend diskutiert werden. Dabei werden folgende Kernpunkte adressiert: Zum einen werden mögliche Ursachen für Schwierigkeiten beim Arbeiten mit depiktionalen Repräsentationen in den Naturwissenschaften, besonders in der Biologie, diskutiert. Zum anderen werden die Unterschiede zwischen Text- und MER-Bedingungen sowie Auswirkungen von Redundanz in MER diskutiert.

7.1.1 Potentielle Ursachen für beobachtete Schwierigkeiten mit bildlichen Repräsentationen

Durch die Testung von Bachelorstudierenden in dieser Arbeit können Rückschlüsse gezogen werden, wie gut Schüler:innen während der Schulzeit sowie ersten Lerngelegenheiten im Studium erlernen und in die Lage versetzt werden, mit verschiedenartigem Text-Bild-Material im Bereich der Biologie zu arbeiten. Dass hier deutliches Verbesserungspotential besteht, wird durch die empirischen Befunde dieser Arbeit deutlich. Die Ergebnisse der ersten experimentellen Studie zeigten, dass Biologiestudierende im Bachelorstudium Schwierigkeiten bei der Arbeit mit Text-Bild-Material hatten, wenn im Material fachspezifische, komplexe Abbildungen enthalten waren. Dies traf auf die Molekulargenetik-Aufgabe zu, nicht aber auf die Aufgabe zu den Themengebieten Ökologie und Evolution. Die Gruppe *keine Redundanz* erzielte bei der Molekulargenetik-Aufgabe die niedrigste Punktzahl, Gruppe *Redundanz*, die lösungsrelevante bildliche Informationen zusätzlich aus dem Text entnehmen konnte, schnitt besser ab. Beide Text-Bild-Gruppen schnitten aber schlechter ab als die Gruppe *nur Text*. Während die aufgetretenen oder ausgebliebenen

Auswirkungen von Redundanz im Text-Bild-Material oder von der Anwesenheit von bildlichen Repräsentationen im Vergleich zu *nur Text* für beide Aufgaben im nachfolgenden Unterkapitel diskutiert werden, soll hier der Blick zunächst auf die Schwierigkeit bei der Bearbeitung von MER-Material mit biologiespezifischen, komplexen Abbildungen gerichtet werden. Durch die zweite experimentelle Studie wurde mit der Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie ein Faktor identifiziert, der das Arbeiten mit MER-Material mit biologiespezifischen, komplexen Abbildungen beeinflussen kann. Dies und weitere potentielle Ursachen für die gezeigten Schwierigkeiten mit derartigen Abbildungen werden mit Bezug zu den in Kapitel 2 vorgestellten theoretischen Grundlagen sowie weiterer Literatur und Hinweisen, die die assoziierte Masterarbeit (Klann, 2018) liefert, diskutiert.

In Kapitel 2 wurde bereits zum einen die große Bedeutung von verschiedenen (M)ER für den Aufbau konzeptuellen Verständnisses und die Kommunikation von Fachinhalten in Naturwissenschaften wie der Biologie herausgestellt. Zum anderen wurde deutlich, dass ein kompetentes Arbeiten mit verschiedenen Repräsentationen Teil von Scientific Literacy ist, die zu erreichen ein Ziel schulischer Bildung ist (Norris & Phillips, 2003). Beides spiegelt sich auch in den normativen Vorgaben hinsichtlich des kompetenten Arbeitens mit verschiedenen Repräsentationen als Teil der Kommunikationskompetenz in den Bildungsstandards im Fach Biologie wider (für Deutschland: KMK, 2005, 2020). Die Ergebnisse dieser Arbeit zeigen allerdings, dass Schüler:innen während der Schulzeit im Biologieunterricht dennoch nicht ausreichend in die Lage versetzt beziehungsweise darauf vorbereitet werden, mit verschiedenartigen MER in der Biologie zu arbeiten, wenn diese biologiespezifische, komplexe Abbildungen enthalten, und dies, obwohl die Stichprobe insofern eine Positivselektion darstellt, als die Teilnehmenden sich für ein Studium der Biologie entschieden haben. Nachfolgend werden absatzweise mögliche Ursachen der beobachteten Schwierigkeiten bei der Informationsentnahme aus biologiespezifischen, komplexen Abbildungen diskutiert.

Vorwissen und Repräsentationskompetenz. Die Entwicklung des Testmaterials zur Untersuchung der Effekte von Redundanz (Studie 1; Kapitel 4) erfolgte in der Annahme, dass Biologiestudierende in Bezug auf Abituraufgaben das nötige fachliche Vorwissen und Wissen über genutzte Repräsentationsarten besitzen sollten, um diese Aufgaben erfolgreich zu bearbeiten. Sie haben die Schule mit Erwerb der Hochschulzugangsberechtigung abgeschlossen, also auch die Oberstufe durchlaufen, in der auf das Arbeiten mit

Abituraufgaben vorbereitet wird, und sich zudem bewusst für ein Biologiestudium entschieden. Zudem waren die eingesetzten Biologieaufgaben so ausgewählt, dass für das Lösen wenig Vorwissen benötigt wurde. In erster Linie war also ein auf der Variation des Aufgabenmaterials basierender Effekt der Redundanz erwartet worden. Aufgrund der Erkenntnisse zu potentiellen Schwierigkeiten bei Konfrontation mit fachtypischen bildlichen Repräsentationen auch über die Schulzeit hinaus (vgl. Kapitel 2.4) wurden die beiden eingesetzten Abituraufgaben so ausgewählt, dass ein Aufgabenmaterial fachspezifische und komplexe bildliche Repräsentationen enthielt, das andere bildliche Repräsentationen, die in der Biologie, aber auch in anderen Fachdisziplinen oder im gesellschaftlichen Leben allgegenwärtig (d. h. nicht fachspezifisch) sind. Obwohl eine große Mehrheit der Teilnehmenden Biologie in der Oberstufe belegt hatte, bereiteten den Biologiestudierenden komplexe, biologiespezifische Abbildungen in der Molekulargenetik-Aufgabe Schwierigkeiten. Die Aufgaben waren zwar so ausgewählt worden, dass die Aufgabenbearbeitung auch mit wenig Vorwissen hätte gelingen können. Dennoch kann das Vorwissen eine Rolle gespielt haben. Kozma und Russell (1997) haben in ihrer Studie mit Noviz:innen und Expert:innen in der Domäne Chemie Unterschiede in der Repräsentationskompetenz festgestellt (siehe auch Kapitel 2.4). Als Noviz:innen wurden dabei *Undergraduates* einer Universität betrachtet, als Expert:innen vorrangig Chemiker:innen pharmazeutischer Labore und Chemiedoktorand:innen. Davon ausgehend können auch die Bachelorstudierenden, die in der vorliegenden Arbeit die Stichprobe bildeten, als Noviz:innen bezogen auf die Disziplin Biologie eingeschätzt werden. Expert:innen können im Gegensatz zu Noviz:innen auf viel konzeptuelles Wissen zurückgreifen, das hilft, Inhalte in verschiedenen Repräsentationsarten zu verstehen (sie dürften zum Beispiel mit der mitochondrialen DNA und ihrer Darstellung vertraut sein), und haben zudem mehr Erfahrung im Gebrauch von verschiedenen Repräsentationsarten in der Kommunikation mit Fachkolleg:innen (Kozma & Russell, 2005).

Im Bereich der Biologie stellten zum Beispiel Cook et al. (2008) und Brandstetter-Korinth (2017) einen Zusammenhang des Vorwissens mit der Informationsentnahme aus Abbildungen fest. So gelang es Highschool-Schüler:innen mit hohem auf das untersuchte biologische Thema bezogenem Vorwissen besser, für das Verständnis relevante Informationen aus schematischen Abbildungen zum Stofftransport durch eine Membran zu entnehmen und zu interpretieren, als Schüler:innen mit geringem Vorwissen (Cook et al., 2008). Letztere konnten beispielsweise Oberflächenmerkmale (z. B. Form und Farbe von Elementen

ten) nicht mit zugrunde liegenden Fachinhalten in Verbindung bringen. Brandstetter-Korinth (2017) stellte – wenn auch mit vermutlich durch Einsatz der Thinking-Aloud-Methode bedingter geringer Stichprobengröße ($N = 42$) – einen positiven Zusammenhang zwischen themenspezifischem Vorwissen und dem Bildverstehen von Schüler:innen bei Konfrontation mit schematischen Abbildungen des Blutkreislaufs und des Kniesehenreflexes fest. Auch in der vorliegenden Arbeit waren in der Molekulargenetik-Aufgabe im Text-Bild-Material fachspezifische schematische Abbildungen enthalten. Bezogen auf die Ergebnisse dieser Arbeit ist anzunehmen, dass ein größeres Vorwissen zu mitochondrialer DNA oder mitochondrialen Erkrankungen dabei geholfen hätte, die lösungsrelevanten Informationen aus bildlichen Abbildungen, also auch biologiespezifischen Abbildungen, zu entnehmen. So wurde auch teilweise von Teilnehmenden der Masterarbeitsstudie bei der Molekulargenetik-Aufgabe, nachdem diese insgesamt als mittelschwer bis schwer eingestuft wurde, unter anderem das Erfordernis von Fach- bzw. Vorwissen als eine Ursache dieser Einschätzung genannt (Klann, 2018). Dies war bei der als leichter eingestuften Ökologie/Evolutions-Aufgabe nicht der Fall (Klann, 2018). Objektiv betrachtet wurde jedoch in beiden Aufgaben wenig Vorwissen benötigt. Allerdings stellten auch Kottmeyer et al. (2020) in ihrer bereits in Kapitel 2.4 vorgestellten Studie fest, dass Studierende, die größtenteils ein biologiebezogenes Studienfach studierten, eine bessere Testleistung bei der Bearbeitung von Fragen zu Abbildungen aus der Biologie als aus Geowissenschaften erzielten, obwohl diese ohne Vorwissen lösbar waren. Das heißt: Auch wenn Vorwissen nicht nötig war, konnte es sich positiv auswirken (Kottmeyer et al., 2020). Es ist zu hoffen und anzunehmen, dass sich im Verlauf des Studiums neben dem Fachwissen auch die Repräsentationskompetenz der Studierenden steigert. Damit dies gelingen kann, sollte die Repräsentationskompetenz in Lehr-Lernkontexten explizit gefördert werden (Kozma & Russell, 1997; s. auch Implikationen Kapitel 7.3). Dies gilt nicht nur für die Lehre im Studium, sondern auch für den Fachunterricht in der Schule. Denn es sei abermals darauf hingewiesen, dass es sich bei dem in dieser Arbeit eingesetzten Material um Abituraufgaben-Material handelte und Schüler:innen durch den Biologieunterricht, besonders auch den Oberstufenunterricht, befähigt sein sollten, mit diesem, also auch den enthaltenen bildlichen Repräsentationen, zu arbeiten.

Selbstwirksamkeitserwartung. In Studie 2 (Kapitel 5) wurde untersucht, ob sich die Überzeugung von Personen, mit komplexen Abbildungen in der Biologie arbeiten zu können, in ihrer Leistung widerspiegelt, wenn diese Anforderung bewältigt werden muss.

Da es hierfür deutliche Hinweise gab, kann die Selbstwirksamkeitserwartung als ein weiterer Faktor betrachtet werden, der zu den in Studie 1 (Kapitel 4) beobachteten Schwierigkeiten bei der Informationsentnahme aus fachspezifischen, komplexen Abbildungen beigetragen hat. Die genannte Selbstwirksamkeitserwartung der Studierenden hatte bei der Molekulargenetik-Aufgabe in Gruppe *keine Redundanz*, die zur Lösung der Aufgabe Informationen aus komplexen Abbildungen der Biologie entnehmen musste, einen positiven Effekt auf die Informationsentnahme. Dieser Effekt blieb bei Studierenden der Gruppe *Redundanz* vermutlich deshalb aus, weil sie auf die Informationen im Text ausweichen konnten. Die Annahme, dass die Studierenden die Abbildungen der Molekulargenetik-Aufgabe tatsächlich als komplex wahrgenommen haben, wird auch durch Beurteilungen der Abbildungen durch Teilnehmende der Masterarbeitsstudie gestützt (Klann, 2018), wobei der geringe Stichprobenumfang im Blick behalten werden muss. Bereits in Studie 1 wurden negative affektiv-motivationale Effekte der komplexen Abbildungen der Molekulargenetik-Aufgabe auf die Testleistung vermutet (Diskussionsteil in Kapitel 4). Eine bestimmte Selbstwirksamkeitserwartung kann sich wie in Kapitel 2.5.2 beschrieben auf die Anstrengung und Persistenz bei der Aufgabenbearbeitung auswirken sowie auf die Wahrscheinlichkeit, ob diese Aufgabe überhaupt in Angriff genommen wird, und letztendlich auf die Leistung (Pajares, 1996). Der Effekt der Selbstwirksamkeitserwartung in Studie 2 legt nahe, dass sich Studierende mit geringer Selbstwirksamkeitserwartung zum Arbeiten mit komplexen Abbildungen in der Biologie möglicherweise entweder an die Abbildungen der Molekulargenetik-Aufgabe gar nicht erst herangetraut haben oder bei auftretenden Schwierigkeiten schnell aufgegeben und keine größeren Anstrengungen unternommen haben, die Abbildungen zu verstehen. Wie in Kapitel 2.5.4 bereits beschrieben wurde der Zusammenhang unterschiedlicher Selbstwirksamkeitserwartungen und akademischer Leistung im Gebiet der Naturwissenschaften in vielen empirischen Arbeiten festgestellt. Diese empirische Befundlage kann durch die Ergebnisse von Studie 2 um Effekte einer Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie auf die Leistung ergänzt werden (einschränkend s. Limitationen Kapitel 5 und 7.2). Eine Stärkung dieser Selbstwirksamkeitserwartung im Schulunterricht, zum Beispiel über entsprechende Lerngelegenheiten mit Erfolgserfahrungen, könnte Schüler:innen zu besseren Schulleistungen im Fach Biologie verhelfen, insbesondere aber auch besser auf ein mögliches Biologiestudium vorbereiten, in dem häufig mit komplexen Abbildungen und fachspezifischen Abbildungen gearbeitet bzw. kommuniziert wird. In Kapitel 2.5.3 wurden mögliche Maßnahmen für den Unterricht zur Stärkung von Selbstwirksamkeitserwartung mit Blick auf deren vier

Hauptquellen vorgestellt. Potentiell wirksame Maßnahmen für den Biologieunterricht zur Stärkung der in dieser Arbeit untersuchten Selbstwirksamkeitserwartung werden in Kapitel 7.3 konkretisiert.

Abstraktionsgrad. Brandstetter-Korinth (2017) zeigte in ihrer oben zitierten Arbeit nicht nur einen Einfluss des Vorwissens, sondern unter anderem auch des Abstraktionsgrads von Schemazeichnungen auf das Bildverstehen von Schüler:innen. Ein Beispiel für eine realistische Version ihres Aufgabenmaterials ist eine Darstellung des Blutkreislaufes, in der der Umriss eines Menschen enthalten und die relative Lage von Organen und (fein verzweigtem) Gefäßsystem realitätsnah dargestellt ist. In der abstrakten Version fehlt der Umriss des Menschen und die Gefäße sind gerade bzw. rechtwinklig angeordnet, die durch Konventionen festgelegte rote und blaue Farbe für sauerstoffreiches und -armes Blut ist aber in beiden Abbildungsversionen enthalten. Trotz Interaktionseffekten mit weiteren Variablen kann festgehalten werden, dass, wenn bei bestimmtem Aufgabenmaterial unterschiedliche Effekte von realistischen und abstrakten Abbildungen gezeigt wurden, abstrakte Abbildungen unabhängig vom Vorwissen der Schüler:innen das Bildverstehen reduzierten (Brandstetter-Korinth, 2017). Hier lassen sich möglicherweise Parallelen zu den Abbildungen der Molekulargenetik-Aufgabe in dieser Arbeit ziehen. In Kapitel 2.2.2 wurde die Klassifikation von Repräsentationen nach Schnotz (2001) vorgestellt, der Schemazeichnungen zu den realistischen depiktionalen Repräsentationen zählt. Das heißt, beide von Brandstetter-Korinth (2017) eingesetzten Abbildungsversionen sind nach Schnotz (2001) als realistisch einzuordnen. Auf dem Kontinuum von konkret zu abstrakt, entlang dessen depiktionalen Repräsentationen eingeordnet werden können (Poizzer-Ardenghi & Roth, 2010), können aber auch realistische Abbildungen in ihrer Position variieren. Auch Schemazeichnungen können also – wie bei Brandstetter-Korinth (2017) – im Abstraktionsgrad variieren. Bezugnehmend auf Arbeiten mehrerer Autor:innen bezeichnet Brandstetter-Korinth abstrakte Abbildungen als „stilisierte, schematische Graphiken, die sich durch reduzierte, einfache Formen, konventionalisierte Zeichen und fachspezifische Symbole auszeichnen [...]“ (2017, S. 21). Dies trifft auf die Abbildungen der Molekulargenetik-Aufgabe dieser Dissertation zu. In der Karte der mitochondrialen DNA (Fig. 4.3.1 und 4.3.2, Kapitel 4) wird die DNA nicht detailliert – etwa in Form von Strukturformeln aneinandergereihter Nukleotide – dargestellt, sondern vereinfacht als zwei ringförmige Striche. Die tRNAs, für die diese DNA codiert, werden einer Konvention folgend als Buchstaben angegeben. In der zweiten Abbildung wird die Basenabfolge ebenfalls konventionalisiert in Form von Buchstaben abgebildet. Auch wenn diese Art der Darstellung mit weniger Details

möglicherweise zur Vereinfachung gedacht ist, ist denkbar, dass der Abstraktionsgrad dieser biologiespezifischen Abbildungen die Informationsentnahme erschwert hat. Zum Beispiel stellten auch Kragten et al. (2013) fest, dass die Vertrautheit mit Konventionen für Komponenten bildlicher Darstellungen biologischer Prozesse einen Einfluss auf die Aufgabenschwierigkeit haben kann. Zwar sind mit Blick auf oben genanntes Kontinuum auch die Achsendiagramme und die Tabelle der Ökologie/Evolution-Aufgabe als abstrakt einzuordnen, diese Repräsentationstypen dürften allerdings auch durch Lehr-Lernkontexte in anderen Fächern oder die Präsenz dieser Repräsentationstypen im gesellschaftlichen Leben vertraut sein. Anders als bei der Studie von Brandstetter-Korinth (2017) umfasste die Stichprobe des in der vorliegenden Dissertation dargelegten Promotionsprojekts keine Schüler:innen der Sekundarstufe I, deren Repräsentationskompetenz im Verlauf der Schulzeit noch gefördert werden könnte, sondern Biologiestudierende, die durch den Fachunterricht – besonders den Oberstufenunterricht – in der Schule hätten befähigt werden sollen, auch mit abstrakteren biologiespezifischen Abbildungen zurechtzukommen. Dass diese vielen Studierenden Schwierigkeiten bereiteten, lässt darauf schließen, dass das Arbeiten mit derartigen Abbildungen im Schulunterricht gezielter unterrichtet werden sollte (s. Implikationen Kapitel 7.3).

Kognitive Belastung und Komplexität von Abbildungen. Mit Blick auf die limitierte Kapazität des Arbeitsgedächtnisses ist denkbar, dass die Abbildungen der Molekulargenetik-Aufgabe den *Intrinsic Cognitive Load* erhöht haben, was bei gleichzeitigem hohem *Extraneous Cognitive Load*, auf den nachfolgend eingegangen wird, die Informationsverarbeitung erschweren kann (vgl. Kapitel 2.3.1). Wenn Material viele Elemente enthält, die gleichzeitig im Arbeitsgedächtnis verarbeitet werden müssen, die Element-Interaktivität somit hoch ist, ist der *Intrinsic Cognitive Load* hoch und das Material als komplex anzusehen (Sweller, 2005a). Dies trifft auf mindestens zwei der drei Abbildungen der Molekulargenetik-Aufgabe zu. So bewerteten auch in der Masterarbeitsstudie (Klann, 2018) die Teilnehmenden die Darstellung der mitochondrialen DNA und den Graphen im Mittel als komplex, die Darstellung der Basenabfolge wurde als einfach eingestuft, wobei es individuelle Unterschiede zwischen Teilnehmenden gab. Dagegen wurden alle bildlichen Repräsentationen der Ökologie/Evolution-Aufgabe im Mittel als sehr einfach bis einfach empfunden. Gefragt nach der Schwierigkeit der Aufgabe, der Komplexität der Aufgabeninhalte und der kognitiven Anstrengung, die sie bei der Aufgabenbearbeitung aufwenden mussten, gaben die Studierenden für die Molekulargenetik-Aufgabe mittlere bis hohe Werte an, für die Ökologie/Evolution-Aufgabe niedrige Werte. Begründet wurden diese Angaben bei der

Molekulargenetik-Aufgabe häufig mit der Komplexität der Abbildungen (Klann, 2018). Dass nicht nur die Karte der mitochondrialen DNA – eine fachspezifische Abbildung mit hoher Informationsdichte und abstrakten Elementen – als komplex wahrgenommen wurde, sondern auch der Graph, kann zum einen daran gelegen haben, dass dessen Datenpunkte mit Abkürzungen beschriftet waren und sich so eine relativ hohe Informationsdichte ergab. Zum anderen stellte das Datenmuster mit eingezeichneter Kurve einen negativ exponentiellen Zusammenhang dar, den die Teilnehmenden eventuell als schwieriger zu interpretieren empfanden als zum Beispiel den linearen Zusammenhang, der im Graph der Ökologie/Evolution-Aufgabe dargestellt war. Brandstetter-Korinth (2017) erhob in der bereits zitierten Bildverstehens-Studie ebenfalls den bei Auseinandersetzung mit den Abbildungen empfundenen *Cognitive Load*. Dieser war beim Verstehen abstrakter Abbildungen höher als bei konkreteren Abbildungen. Ein höherer *Cognitive Load* ging mit schlechterem Bildverstehen einher (Brandstetter-Korinth, 2017).

Auch wenn generell die bildlichen Repräsentationen in Gänze betrachtet werden mussten, damit erkennbar war, was sie aussagen, und damit lösungsrelevante Informationen entnommen werden konnten, waren in den Abbildungen der Molekulargenetik-Aufgabe viele Details enthalten (z. B. in der Abbildung der mitochondrialen DNA: eingezeichnete Positionen und Abkürzungen von weiteren Genen für tRNAs und Proteine), die für das Lösen der Aufgabe irrelevant waren. Diese könnten, ähnlich wie es für *Seductive Details* angenommen wird, einen von eigentlich relevanten Informationen ablenkenden Effekt gehabt haben (z. B. Garner et al., 1991; s. auch Kapitel 7.1.2) und möglicherweise den *Extraneous Cognitive Load* gesteigert haben. Es ist denkbar, dass diese Details bei der Konstruktion der Abituraufgabe bewusst eingefügt wurden, damit die lösungsrelevanten Informationen nicht direkt erkennbar waren (Weiteres zur Gestaltung der Abituraufgaben in Kapitel 7.1.2).

Die Belastung des Arbeitsgedächtnisses kann durch bereits vorhandene Schemata reduziert werden (vgl. Kapitel 2.3.1.3). Wie bereits angeführt, hätte ein höheres – in Schemata organisiertes – Vorwissen zu mitochondrialer DNA und mitochondrialen Erkrankungen möglicherweise geholfen, mehr Informationen simultan im Arbeitsgedächtnis verarbeiten zu können, und es wäre möglicherweise mehr auf die relevanten Elemente der Abbildungen fokussiert worden, weil sie mit Hilfe des Vorwissens leichter identifiziert worden wären. Die Rolle des Vorwissens auch mit Bezug zur Repräsentationskompetenz wurde zu Beginn dieses Kapitels bereits diskutiert.

Die Beobachtung von Schwierigkeiten bei der Informationsentnahme aus den Abbildungen der Molekulargenetik-Aufgabe passt zu den in Kapitel 2.4 vorgestellten Studien, die ebenfalls Defizite in der Repräsentationskompetenz von Studierenden in naturwissenschaftlichen Kursen festgestellt haben. Damit Schüler:innen – bereits bevor sie möglicherweise ein naturwissenschaftliches Studium beginnen – eine hohe Repräsentationskompetenz und positive Selbstwirksamkeitserwartung entwickeln können, um auch mit fachspezifischen, komplexen Abbildungen erfolgreich arbeiten zu können, sollte neben der Vermittlung von Fachwissen beides im Unterricht von den Lehrkräften durch entsprechende Lerngelegenheiten mit verschiedenen Repräsentationsarten explizit gefördert werden (s. auch Implikationen in Kapitel 7.3). Unter anderem Befunde von Nitz et al. (2014) weisen darauf hin, dass Schüler:innen die Anzahl der Lerngelegenheiten bezüglich mancher repräsentationaler Aspekte im Biologieunterricht als nicht sehr hoch empfinden. Gerade die Bedeutung der Selbstwirksamkeitserwartung von Schüler:innen oder Studierenden in Bezug auf komplexe Abbildungen, die sich auf Motivation und letztendlich – wie in dieser Arbeit gezeigt – auf die Leistung auswirken kann, ist Lehrenden vermutlich häufig nicht bewusst und sollte verstärkt berücksichtigt werden.

Schwierigkeiten bei der Informationsentnahme aus bildlichen Repräsentationen – ganz gleich wodurch sie letztendlich begründet sind – sind besonders dann problematisch, wenn die relevanten bildlichen Informationen nicht im Text wiederholt werden. Im folgenden Unterkapitel wird auf die Rolle der Redundanz und die MER-Integration eingegangen.

7.1.2 Auswirkungen von Redundanz in MER und MER gegenüber reinem Text

Die experimentellen Ergebnisse dieser Arbeit zeigen, dass Redundanz in Text-Bild-Material im Vergleich zu nicht-redundantem Text-Bild-Material nicht zu einer verschlechterten Testleistung geführt hat, wie es im Falle des in Kapitel 2.3.4 vorgestellten Redundanzprinzips nach Sweller (2005b) zu erwarten gewesen wäre. Mögliche Ursachen für Unterschiede zu den Ergebnisse der in Kapitel 2.3.4 und im theoretischen Rahmen von Kapitel 4 vorgestellten weiteren experimentellen Studien zu Redundanz wurden im Diskussionsenteil von Kapitel 4 bereits diskutiert. Besonders ist hervorzuheben, dass anders als in vorangegangenen Studien in dieser Arbeit der Materialumfang zwischen den beiden Text-Bild-Bedingungen, in denen die Redundanz variiert wurde, konstant gehalten wurde. Im Folgenden soll nun zunächst der Blick auf das Redundanzprinzip und den *Cognitive Load* sowie davon ausgehend auf die beiden eingesetzten Aufgaben gerichtet werden.

Wenn dieselbe Information in verschiedenen Formen, zum Beispiel Text und Bild, präsentiert wird, kann durch zusätzliche Belastung des beim Verarbeiten neuer Informationen in der Kapazität limitierten Arbeitsgedächtnisses das Lernen erschwert sein, das heißt, Redundanz zu vermeiden könnte das Lernen erleichtern (Redundanzeffekt; Sweller, 2005b) bzw. allgemein das Arbeiten mit multimedialen Material. Denn wenn redundante Informationen in verschiedenen Repräsentationen vorhanden sind, werden kognitive Ressourcen darauf aufgewendet, sie zu koordinieren und miteinander in Verbindung zu bringen, das heißt, der *Extraneous Cognitive Load* wird erhöht (Sweller, 2005b). Der Redundanzeffekt ist eng mit dem *Expertise Reversal Effect* (Kapitel 2.3.1.3) verknüpft, der auftritt, wenn beispielsweise eine bestimmte Gestaltung von MER vorteilhaft für Noviz:innen ist, aber mit zunehmender Expertise ihre Vorteile verliert und gegebenenfalls sogar nachteilig wird (Sweller, 2005b). Laut Sweller ist dies häufig darauf zurückzuführen, dass etwa zusätzliche Erläuterungen, die Noviz:innen helfen, für Expert:innen redundant sind und den *Extraneous Cognitive Load* somit unnötig erhöht. Dies sollte jede Konfrontation mit multimedialen Material betreffen, dem neue Informationen zu entnehmen sind, also auch das Bearbeiten von Aufgabenmaterial.

Molekulargenetik-Aufgabe. Im vorangegangenen Unterkapitel wurde bereits erläutert, warum die Studierenden dieser Dissertationsstudie als Noviz:innen in der Domäne Biologie angesehen werden können. Die Redundanz stellte sich in der Molekulargenetik-Aufgabe, die fachspezifische Abbildungen enthielt, gegenüber der Vermeidung von Redundanz im Text-Bild-Material als vorteilhaft dar, wobei denkbar ist, dass die redundante Textinformation zu lösungsrelevanten Abschnitten in Abbildungen hingeleitet hat oder die Informationen ausschließlich aus dem Text und nicht aus Abbildungen entnommen wurden. Diese Überlegung wird durch Ergebnisse der Masterarbeitsstudie gestützt. Hier wurde beobachtet, dass Studierende der Gruppe *Redundanz*, wenn sie lösungsrelevante Informationen der zwei besonders komplexen Molekulargenetik-Abbildungen identifiziert haben, diese erst dem Text entnommen und dann teilweise in der jeweiligen Abbildung gefunden haben (Klann, 2018). Mit Blick auf die von Ainsworth (1999, 2006) vorgeschlagenen Funktionen von MER könnte in Gruppe *Redundanz* eine einschränkende Funktion vorgelegen haben, dass also die vertraute Repräsentationsart Text bei der Interpretation von weniger vertrauten, fachspezifischen und komplexen bildlichen Repräsentationen geholfen hat (vgl. Kapitel 2.3.2).

In Gruppe *keine Redundanz*, die auf lösungsrelevante Informationen in Abbildungen angewiesen war, konnten die Informationen aus Text und Bild möglicherweise häufig

nicht integriert werden, da bereits die lokale Kohärenzbildung in den bildlichen Repräsentationen nicht erfolgreich war (s. Kapitel 7.1.1), die der globalen Kohärenzbildung zwischen Repräsentationen einer MER vorausgeht (vgl. Kapitel 2.3). Studierende der Gruppe *Redundanz* mussten anders als die der Gruppe *keine Redundanz* nicht zwingend Informationen aus den beiden Kanälen des Arbeitsgedächtnisses integrieren (vgl. Kapitel 2.3.1) und konnten die Molekulargenetik-Aufgabe auch lösen, ohne die komplexen Abbildungen zu verstehen.

Unabhängig davon, wie viel Mühe in das Verstehen der komplexen Abbildungen investiert wurde und ob letztendlich aus ihnen Informationen entnommen wurden, weist das bessere Abschneiden der Gruppe *nur Text* gegenüber beiden Text-Bild-Gruppen darauf hin, dass die Abbildungen zumindest beachtet wurden, wenn auch mit – im Durchschnitt – nachteiligem Effekt. Dies – spezifisch auch, dass die in allen drei Gruppen rein textlich dargebotenen Informationen der Molekulargenetik-Aufgabe in Gruppe *nur Text* erfolgreicher entnommen wurden als in beiden Text-Bild-Gruppen – könnte unter anderem durch den hohen *Intrinsic Cognitive Load* durch das Text-Bild-Aufgabenmaterial begründet sein (vgl. Kapitel 2.3.1), wobei die wahrgenommene Komplexität möglicherweise auch die Motivation zur Aufgabenbearbeitung negativ beeinflusst hat. Zudem ist denkbar, dass gerade die fachspezifischen, also nicht so vertrauten Abbildungen die Aufmerksamkeit teilweise vom Text weggelenkt haben könnten, lösungsrelevante Informationen dann aber aus in Kapitel 7.1.1 diskutierten Gründen nicht immer entnommen werden konnten.

In Kapitel 7.1.1 wurden bereits in Bezug auf irrelevante Details in den Abbildungen Parallelen zum Effekt von *Seductive Details* gezogen. Für den Vergleich von Gruppe *nur Text* und Gruppe *Redundanz* könnte dies auf die bildlichen Repräsentationen in Gruppe *Redundanz* erweitert werden. *Seductive Details* sind mit dem Ziel, die Motivation der Rezipierenden zu erhöhen, in Material eingefügte irrelevante, aber interessante Informationen (textlich oder bildlich dargestellt), die sich allerdings häufig negativ auf das Lernen beziehungsweise die Testleistung auswirken (Mayer, 2019). Mayer und Fiorella (2014) führen in Bezug auf *Seductive Details* das Kohärenzprinzip an, das besagt, dass interessantes, aber irrelevantes Material entfernt werden sollte, da es das *Extraneous Processing* erhöht (vgl. Kapitel 2.3.1.1). Zwar waren die Abbildungen im Testmaterial dieser Arbeit keine *Seductive Details* im oben genannten Sinne, da sie Informationen enthielten, die für das Lösen der Aufgabe relevant waren, und nicht zur Motivation in das Material eingefügt worden waren. In Gruppe *Redundanz* konnten die Aufgaben aber auch allein mit dem Text gelöst werden. Das heißt, ähnlich wie *Seductive Details* haben die Abbildungen möglicher-

weise Aufmerksamkeit auf sich gezogen und damit auch vom Text weggelenkt, aber – wenn sie nicht verstanden wurden (Kapitel 7.1.1) – keinen Beitrag zur Lösung der Aufgabe geliefert und zudem den *Extraneous Cognitive Load* im Vergleich zur Gruppe *nur Text* unnötig erhöht.

Ökologie/Evolution-Aufgabe. Wie im vorangegangenen Unterkapitel erläutert, war der *Intrinsic Cognitive Load* durch das Aufgabenmaterial in der Ökologie/Evolution-Aufgabe vermutlich deutlich geringer als in der Molekulargenetik-Aufgabe. In der Masterarbeitsstudie wurde teilweise die Übersichtlichkeit der Abbildungen explizit hervorgehoben, die bildlichen Repräsentationen wurden wie oben beschrieben von allen Teilnehmenden im Mittel als sehr einfach bis einfach eingestuft und die Aufgabe insgesamt als leicht wahrgenommen (Klann, 2018). Ist der *Intrinsic Cognitive Load* gering, fällt der durch Redundanz erzeugte *Extraneous Cognitive Load* nicht stark ins Gewicht (vgl. Kapitel 2.3.1.3), was erklären könnte, dass sich die Testleistung zwischen den Gruppen nicht unterschied. Dass bildliche Repräsentationen anders als nach dem Multimedia-Prinzip von Mayer (2001) keinen positiven Effekt bewirken konnten gegenüber reinem Textmaterial, könnte daran gelegen haben, dass die Informationen aus den Abbildungen auch sehr einfach textlich dargestellt werden konnten. Zu beachten ist zudem, dass in der Studie das Ergebnis einer Aufgabenbearbeitung untersucht wurde und nicht das Lernen im Sinne eines längerfristig gespeicherten Wissens. Es ist nicht auszuschließen, dass übereinstimmend mit dem in der *Dual Coding Theory* (Paivio, 1991) angenommenen additiven Effekt beim späteren Erinnern die Studierenden der Text-Bild-Gruppen, die eine hohe Punktzahl erzielt haben, in einem Posttest zu Inhalten des Aufgabenmaterials besser abschneiden würden als die Studierenden der Gruppe *nur Text*, die eine hohe Punktzahl erzielten. Dies ist für beide Aufgaben denkbar.

Dass die relevanten Informationen aus den Abbildungen auch sehr einfach textlich dargestellt werden konnten und generell die Ergebnisse dieser Arbeit weisen darauf hin, dass die Gestaltung von Abituraufgaben im Fach Biologie möglicherweise verbesserungswürdig ist (Material der Gruppe *keine Redundanz* entsprach nahezu den Original-Abituraufgaben). Gerade der Einsatz von bildlichen Repräsentationen – die Art, wie sie mit Text kombiniert werden sowie inhärente schwierigkeiterzeugende Merkmale – sollte beim Design reflektiert werden. Die in dieser Arbeit eingesetzten Abituraufgaben stammten aus den Jahren 2008 und 2009. Schon vor Einführung der neuen Bildungsstandards im Fach Biolo-

gie für die Allgemeine Hochschulreife (KMK, 2020) war in den einheitlichen Prüfungsanforderungen in der Abiturprüfung Biologie (KMK, 2004) gefordert, dass das Prüfungsmaterial auch „Daten aus biologischen Experimenten und Untersuchungen“ (S. 17) enthält und unter anderem die Informationsentnahme aus „vielschichtigen Materialien oder einer wissenschaftlichen Veröffentlichung“ (S. 15) gefordert sein soll. Der Graph in der Molekulargenetik-Aufgabe und alle bildlichen Repräsentationen der Ökologie/Evolution-Aufgabe basierten auf Repräsentationen aus wissenschaftlichen Artikeln, die Ergebnisse biologischer Experimente berichteten. Interessant ist hierbei, dass die bildlichen Repräsentationen der letztgenannten Aufgabe im Vergleich zum Original vereinfacht wurden (Reduktion der Zellenanzahl der Tabelle sowie Entfernen von Standardabweichungen, Löschen von Datenreihen im Graph, Entfernen von Standardabweichungen an den Säulen im Diagramm), der ohnehin schon komplexere Graph der Molekulargenetik-Aufgabe hingegen nicht (abgesehen davon, dass eine in die freie Fläche des Graphs eingefügte Gleichung sowie ein p -Wert durch einen kurzen Text ersetzt wurden). Ob Effekte dieser Designentscheidungen bedacht wurden oder sogar erwünscht waren, ist fraglich. Florian et al. (2015), die Aufgabenmerkmale von Abituraufgaben im Fach Biologie untersucht haben, leiten aus fachdidaktischer Sicht Implikationen für die Gestaltung von Abituraufgaben ab. Dabei merken sie an, dass es ratsam sei, Materialzusammensetzung und Materialmenge kritisch zu prüfen, unter anderem, weil es gerade bei hoher Anzahl und Vielfalt von Repräsentationen (pro Teilanforderung) schwieriger werde, die kognitiven Anforderungen an Informationsverarbeitung und Integration von Informationen zu steuern (Florian et al., 2015). Schwieriger sei es dann auch, Informationsredundanzen gezielt einzufügen oder zu eliminieren (Florian et al., 2015). Hier könnte aufgrund der Ergebnisse dieser Arbeit ergänzt werden, dass nicht nur Repräsentationsmenge und -vielfalt bei der Aufgabenkonstruktion beachtet, sondern auch potentielle Effekte bestimmter Abbildungsmerkmale (z. B. fachspezifisch, komplex; vgl. Schwierigkeiten Kapitel 7.1.1) berücksichtigt werden sollten.

Ziepprecht (2016) hatte in ihrer Arbeit gezeigt, dass es Schüler:innen der Sekundarstufe I leichter fiel, Informationen aus biologischen Texten als aus biologischen Bildern und Bild-Text-Kombinationen zu entnehmen. In dieser Arbeit wurde nun beobachtet, dass selbst Biologiestudierende noch Schwierigkeiten mit der Informationsentnahme aus Text-Bild-Material der Biologie hatten, dass also bis zum Ende der Schulzeit diesbezüglich insgesamt keine ausreichende Förderung stattgefunden hat. Ein wichtiger Punkt ist, mit welchen Erfahrungen und Erwartungen Schüler:innen oder Studierende Aufgabenmaterial mit

Texten und Bildern bearbeiten. Slough et al. (2010) stellten fest, dass im Text von Schulbüchern Verweise auf Abbildungen häufig fehlen und letztere oft einen Dekorationscharakter aufweisen. Möglicherweise erwarten Schüler:innen bzw. Studierende nicht, in Abbildungen wichtige Informationen zu finden, die sie nicht auch im Text finden (zu Fokus auf Text vgl. auch Cromley et al., 2010; Hannus & Hyönä, 1999; Kapitel 2.3.3). Ähnlich wie für fachspezifische, komplexe Abbildungen in Kapitel 7.1.1 lässt sich festhalten, dass das Arbeiten mit unterschiedlich gestalteten MER im Unterricht von den Lehrkräften gefördert werden sollte, damit Schüler:innen erfolgreich Informationen aus verschiedenen Repräsentationen entnehmen und integrieren können, was auch für das selbstständige Arbeiten in einem Studium naturwissenschaftlicher Fächer hohe Relevanz hat. Lehrkräften kommt für das Gelingen also eine wichtige Rolle zu. Auswirkungen von beispielsweise Überzeugungen von Lehrkräften in Bezug auf das Lehren mit Bild-Text-Kombinationen zeigten Schroeder et al. (2011). Sie untersuchten Zusammenhänge von Überzeugungen und Unterrichtsverhalten von unter anderem Biologielehrkräften mit dem von Schüler:innen der Sekundarstufe I angegebenen Engagement in Bezug auf das Arbeiten mit Text-Bild-Kombinationen. Die Ergebnisse zeigten unter anderem einen positiven Zusammenhang der Überzeugung der Lehrkräfte, dass Schüler:innen Strategien zum Lesen von Text-Bild-Kombinationen lernen sollten, mit dem genannten Engagement der Schüler:innen. Der Zusammenhang wurde über das von Schüler:innen wahrgenommene Unterrichtsverhalten der Lehrkräfte vermittelt (Schroeder et al., 2011).

Gerade wenn relevante Informationen nicht in mehreren Repräsentationsarten vorliegen, ist es wichtig, zu erkennen, was die einzelnen Repräsentationen zum dargestellten Sachverhalt beitragen. Wenn dies gelingt, können MER ihr Potential entfalten und beispielsweise dem Verständnis komplexer biologischer Phänomene dienen. Implikationen für die Praxis – sowohl für den Schulunterricht und die Hochschullehre im Fach Biologie als auch für die Lehrkräfteaus- und Weiterbildung –, aber auch für zukünftige Forschung werden in den Kapiteln 7.3 und 7.4 diskutiert.

7.2 Stärken und Limitationen

Diese Arbeit liefert neue empirische Befunde dazu, wie gut Personen nach Erlangen der Hochschulzugangsberechtigung mit unterschiedlich redundantem MER-Material in der Biologie arbeiten können, das heißt auch, wie gut sie den Teil der in den Bildungsstandards

im Fach Biologie verankerten Kommunikationskompetenz, der das Arbeiten mit unterschiedlichen Repräsentationen fordert, nach Schulabschluss tatsächlich beherrschen. Es wird deutlich, dass der Fachunterricht diesbezüglich Verbesserungspotential aufweist und besonders das Arbeiten mit komplexen, biologiespezifischen Abbildungen explizit unterrichtet werden sollte. Außerdem werden erstmals empirische Erkenntnisse zu einer Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie und ihrem Effekt auf Leistung erbracht. Detaillierte Diskussionen von Stärken und Limitationen befinden sich in den Kapiteln 4 und 5. Nachfolgend werden die wichtigsten Punkte herausgegriffen und teilweise ergänzt.

Hervorzuheben ist, dass in dieser Arbeit anders als in vorangegangenen Arbeiten Redundanz von Informationen so erzeugt wurde, dass dennoch der Materialumfang in der Gruppe *keine Redundanz* und *Redundanz* konstant gehalten wurde, um Effekte, die rein durch den Umfang des Materials begründet sind, auszuschließen. Das Testmaterial kann als ökologisch valide betrachtet werden, da es auf im Abitur eingesetzten Biologieaufgaben basierte. Darüber hinaus spiegeln Abituraufgaben insofern ein Charakteristikum authentischer Kommunikation in der Disziplin Biologie wider, als in ihrem Material mehrere bildliche Repräsentationen mit Text kombiniert werden, wie es zum Beispiel auch in biologischen Fachartikeln der Fall ist. Die ökologische Validität des Testmaterials, kombiniert mit dem offenen Antwortformat sowie der Wahl von Studierenden als Stichprobe – also Personen, die auch die Oberstufe bereits durchlaufen und sogar schon Kurse im Studium belegt haben –, unterscheidet diese Studie von vorangegangenen Studien zu den Effekten von Redundanz in MER. Wie im Folgenden deutlich wird, bringt der Einsatz umfangreicher und damit zeitintensiver Abituraufgaben allerdings auch Limitationen mit sich, zum Beispiel mit Blick auf die Anzahl eingesetzter Aufgaben.

Die Generalisierbarkeit der Ergebnisse sollte anhand zukünftiger Studien mit Aufgaben auch aus weiteren Themengebieten und einer Variation der depiktionalen Repräsentationsarten unabhängig vom Themengebiet überprüft werden. Aufgrund des Umfangs, den Abituraufgaben aufweisen, und einer auf 90 Minuten begrenzten Erhebungszeit konnten in dieser Arbeit nur zwei Aufgaben eingesetzt werden, sodass die Art depiktionaler Repräsentationen und das Themengebiet konfundiert waren. In Kapitel 4 wurden mögliche Einflüsse des Themengebiets diskutiert. Der trotz Konfundierung gezogene Schluss, dass die in Studie 1 beobachteten Ergebnisse primär auf die Art der depiktionalen Repräsentationen in der Molekulargenetik-Aufgabe zurückzuführen sind, wird sowohl durch Ergebnisse der Masterarbeitsstudie (Klann, 2018) als auch durch Ergebnisse von Studie 2 gestützt.

Durch die quantitative Studie 1 wurden Erkenntnisse über das Resultat des Arbeitens mit verschiedenen MER gewonnen. Die Masterarbeitsstudie von Klann (2018) konnte mit demselben Aufgabenmaterial mittels Eye-tracking- und Thinking-Aloud-Methode sowie einem angeschlossenen Fragebogen erste qualitative Einblicke in den Prozess der Aufgabenbearbeitung geben. Hier muss allerdings einschränkend auf den geringen Stichprobenumfang verwiesen werden. Eine größer angelegte qualitative Untersuchung des Bearbeitungsprozesses wäre anzustreben, um die gemachten Beobachtungen zu untermauern.

Studie 2 liefert Hinweise, dass die Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie ein eigenständiges Konstrukt sein könnte. Dadurch kann die empirische Befundlage zum Zusammenhang unterschiedlicher Selbstwirksamkeitserwartungen und akademischer Leistung in den Naturwissenschaften sowie zu Ursachen von Schwierigkeiten bei der Informationsentnahme aus komplexen Abbildungen in der Biologie um den Effekt dieser Selbstwirksamkeitserwartung ergänzt werden. Der empirische Nachweis anhand entsprechender weiterer Selbstwirksamkeitsmaße, dass sich die untersuchte Selbstwirksamkeitserwartung tatsächlich von einer (1) auf allgemeine Fähigkeiten in der Biologie oder (2) auf andere fachspezifische Anforderungen bezogenen Selbstwirksamkeitserwartung abgrenzen lässt, worauf die Ergebnisse von Studie 2 deutlich hinweisen, steht allerdings noch aus. Eine für Ersteres in dieser Studie eingesetzte Skala, d. h. eine Skala zur Erhebung der Selbstwirksamkeitserwartung in Bezug auf Biologie, zeigte eine zu geringe interne Konsistenz, um für weitere Analysen nutzbar zu sein. Sie war ausgehend von der Skala zur Erfassung der Selbstwirksamkeitserwartung in Naturwissenschaften der PISA-Erhebung von 2006 (Frey et al., 2009) durch kleine Änderungen an das Fach Biologie angepasst worden. Bei der PISA-Erhebung war die Reliabilität mit einem Cronbachs-Alpha-Wert von .82 für die deutsche Stichprobe, die für den internationalen Vergleich herangezogen worden war, gut (Frey et al., 2009).

Wie erwähnt hat die Messung einer Selbstwirksamkeitserwartung, die stark auf die untersuchte zu bewältigende Aufgabe zugeschnitten ist, eine höhere Vorhersagekraft als die Messung einer allgemeineren Selbstwirksamkeitserwartung (Bandura, 1986; Pajares, 1996; Kapitel 2.5.2). Für die in dieser Arbeit untersuchte Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie war die mit der zu bewältigenden Aufgabe korrespondierende Spezifität der Selbstwirksamkeitserwartung bei der Molekulargenetik-Aufgabe erfüllt. Hier musste Material mit komplexen Abbildungen aus der Biologie bearbeitet werden, was auch in den Selbstwirksamkeits-Items adressiert wurde. Es wäre aber zu überlegen, den Effekt der genannten Selbstwirksamkeitserwartung

auf die Leistung bei der Bearbeitung von Aufgabenmaterial zu untersuchen, das komplexe, biologiespezifische Abbildungen noch mehr in den Vordergrund stellt und weniger Textanteile enthält, was hier aufgrund der bewussten Entscheidung für Abituraufgaben-Material nicht der Fall war. Dann sollte die Selbstwirksamkeitserwartung einen noch größeren Beitrag zur Varianzaufklärung leisten können als in Studie 2 dieser Arbeit.

Wie beschrieben wurde in dieser Arbeit Abituraufgaben-Material eingesetzt und angenommen, dass Biologiestudierende in der Lage sein sollten, mit dieser Art von Material zu arbeiten. Mögliche Ursachen für die dennoch aufgetretenen Schwierigkeiten bei der Informationsentnahme aus fachspezifischen, komplexen Abbildungen in der Molekulargenetik-Aufgabe wurden in Kapitel 7.1.1 diskutiert. Unter anderem könnten – obwohl die Aufgaben so gewählt worden waren, dass wenig Vorwissen für die Lösung benötigt wurde – das Vorwissen und die kognitive Belastung eine Rolle gespielt haben. Auch wenn die Biologienote erfasst wurde und als fachspezifisches Leistungsmaß fungiert, wären zusätzliche Erkenntnisse über das Fachwissen der Teilnehmenden in Bezug auf die in den Aufgaben behandelten Themengebiete wünschenswert gewesen. Da aus zeitökonomischen Gründen der Einsatz eines etablierten umfangreichen Fachwissenstests nicht möglich war, wurde eine Skala mit zehn Items entwickelt, deren interne Konsistenz sich aber als zu gering erwies, um die Ergebnisse für weitere Analysen nutzen zu können. Zum anderen wären Erkenntnisse zur Arbeitsgedächtniskapazität der Teilnehmenden gewinnbringend gewesen. Denn wie die kognitive Verarbeitung neuer Informationen bewältigt wird, hängt auch von Charakteristika wie der individuellen Arbeitsgedächtniskapazität ab (Mayer, 2014; Kapitel 2.3.1). Die kognitive Belastung des Arbeitsgedächtnisses und somit die Arbeitsgedächtniskapazität ist auch in Bezug auf das Arbeiten mit umfangreichem MER-Material – besonders das Arbeiten mit redundantem Material – sowie das Konstruieren umfangreicher Antworten bei offenem Antwortformat relevant, weshalb in dieser Arbeit ein Test zur Arbeitsgedächtniskapazität durchgeführt wurde. Meist sind Tests dieser Art für die Durchführung am Computer designet. Da die Erhebung der Daten im Rahmen des in dieser Dissertation dargelegten Promotionsprojekts in Paper-Pencil-Form und zudem mit vielen Teilnehmenden gleichzeitig erfolgte, wurde ein Test ausgewählt, der sich für eine Übertragung auf eine PowerPoint-Präsentation, die den Teilnehmenden präsentiert wurde, mit Eintragung der Antworten im Testheft eignete. Die geringe interne Konsistenz zeigte jedoch, dass vermutlich diese Formatänderung Probleme mit sich brachte, weshalb die Variable Arbeitsgedächtniskapazität nicht für weitergehende Analysen genutzt werden konnte (s. Kapitel 4).

Weitere Überlegungen, die sich aus den Ergebnissen und Limitationen dieser Arbeit ergeben, werden in Kapitel 7.4 angeführt.

Trotz genannter Einschränkungen liefert diese Arbeit neue empirische Erkenntnisse zum Arbeiten mit verschiedenartigen MER und potentiellen Ursachen von dabei auftretenden Schwierigkeiten, aus denen sich Implikationen für die Praxis – in erster Linie für den Schulunterricht, aber auch die Hochschullehre im Fach Biologie sowie für die Lehrkräfteteaus- und Weiterbildung – und für zukünftige Forschung ableiten lassen.

7.3 Implikationen für die Praxis

Obwohl in den Bildungsstandards im Fach Biologie unter anderem das Erschließen von Informationen auch mittels komplexer Repräsentationen als Teil der Kommunikationskompetenz gefordert ist (KMK, 2020), zeigen die Ergebnisse dieser Arbeit, dass selbst Studierende Schwierigkeiten bei der Bearbeitung von Abituraufgaben-Material für das Fach Biologie hatten, in dem Texte mit fachspezifischen, komplexen Abbildungen kombiniert waren. Diesbezüglich scheint der Biologieunterricht also wie bereits erläutert verbesserungswürdig zu sein. Folgende Punkte sollen im Verlauf dieses Kapitels adressiert werden. Die Sensibilisierung von Lehrenden der Biologie für Effekte der Gestaltung von MER (bzgl. Redundanz und Einsatz komplexer Abbildungen) ist erforderlich, damit diese

- entsprechende Effekte bei selbstentwickeltem Lern- und Testmaterial beachten und so auch gezielt den Anforderungsgrad anpassen können
- auch im Sinne einer wissenschaftspropädeutischen Bildung bewusst Lerngelegenheiten zur Stärkung der Repräsentationskompetenz der Schüler:innen schaffen, die auch das Integrieren von Text- und Bildinformationen erfordern sowie an komplexe Abbildungen der Biologie heranführen
- Strategien lehren, die zum Beispiel bei der MER-Integration hilfreich sind, was besonders bei nicht-redundanten MER wichtig ist
- die Selbstwirksamkeitserwartung der Schüler:innen bezüglich komplexer und fachspezifischer Abbildungen stärken

Um die Repräsentationskompetenz von Schüler:innen zu fördern, sollten Lehrkräfte Schüler:innen im Biologieunterricht ausreichend viele Gelegenheiten zur verschiedenartigen Nutzung unterschiedlicher Repräsentationsarten bieten (z. B. Nitz et al., 2014).

Auch mit Blick auf die Abiturvorbereitung und ein mögliches Biologiestudium der Schüler:innen sollten spätestens im Oberstufenunterricht komplexe und fachspezifische Abbildungen behandelt werden. Schüler:innen sollten gezielt bezüglich Konventionen verschiedener Abbildungen unterrichtet werden (vgl. Miller et al., 2016; Schönborn & Anderson, 2006). Um über vorgegebene MER-Gestaltungen in Schulbüchern hinaus MER-Nutzung im Biologieunterricht bzw. allgemein im naturwissenschaftlichen Unterricht gezielt zu steuern, könnten Lehrkräfte selbstgestaltetes Material (z. B. Arbeitsblätter) mit verschiedenartigen bildlichen Repräsentationen und auch unterschiedlich redundanten Text-Bild-Kombinationen einsetzen. Dabei könnten Lehrkräfte bewusst realistische und logische depiktionale Repräsentationen nutzen und Schüler:innen, angepasst an Vorwissen und Fähigkeiten, auch an komplexere Abbildungen heranführen. Eine Abbildung wie die Karte der mitochondrialen DNA könnte zum Beispiel durch eine Aufgabenstellung begleitet werden, die darauf abzielt, sich zunächst mit den dargestellten Konventionen vertraut zu machen. Zudem könnten im Biologieunterricht die im Diskussionsteil von Kapitel 5 genannten Richtlinien von Schönborn und Anderson (2006) berücksichtigt werden, wonach zum Beispiel Lernenden auch Limitationen bestimmter Abbildungsarten bekannt sein und verschiedene bildliche Repräsentationen zu demselben Phänomen genutzt werden sollten, damit Lernende diese vergleichen und sich kritisch mit ihnen auseinandersetzen sowie ein möglichst realitätsnahes mentales Modell bilden. In Bezug auf die Abbildungen aus der Molekulargenetik-Aufgabe, die Schwierigkeiten bei der Informationsentnahme hervorriefen, ist zum Beispiel denkbar, dass das Vorstellen und Vergleichen verschiedener Darstellungsarten von DNA (zwei parallele Striche, Buchstabenfolge, Strickleiter-Darstellung, Strukturformel etc.) im Unterricht hilfreich sein könnte.

Ein weiterer Ansatz, der möglicherweise zur Verbesserung der Repräsentationskompetenz von Schüler:innen (und damit letztendlich auch von Studierenden) führen und sie mit fachwissenschaftlicher Kommunikation vertraut machen kann, ist der Einsatz von Adapted-Primary-Literature-Artikeln im Unterricht, was in Israel bereits in Lehrpläne im Bereich der Biologie integriert wurde (Yarden et al., 2015). *Adapted Primary Literature* (APL) bedeutet, dass wissenschaftliche Fachpublikationen für den Einsatz im Schulunterricht adaptiert, d. h. so umgeschrieben werden, dass sie an das Vorwissen und die Fähigkeiten der Schüler:innen angepasst sind, aber Struktur und Charakteristika des originalen Forschungsartikels beibehalten werden (Yarden et al., 2015). Ziele dieses Ansatzes (u. a. Vermittlung eines Verständnisses von *Nature of Science*; Yarden et al., 2015) gehen teil-

weise weit über Implikationen, die aus Ergebnissen der vorliegenden Dissertation abgeleitet werden können, hinaus, betreffen aber auch das Verständnis von naturwissenschaftlicher Sprache mit den enthaltenen bildlichen Repräsentationen. Auch wenn der Ansatz vorsieht, dass in der APL-Version für die Kernaussage des Artikels irrelevante Abbildungen entfernt und komplexe Abbildungen gegebenenfalls verändert werden (Yarden et al., 2015), wäre der Einsatz von APL-Artikeln dennoch eine Möglichkeit, Schüler:innen mit der Bandbreite an Abbildungsarten bzw. MER, die Biolog:innen in ihrer Kommunikation nutzen, vertraut zu machen. Zudem ist jeder Lehrkraft natürlich selbst überlassen, zu welchem Grad sie die originalen Abbildungen für die jeweilige Lerngruppe für bewältigbar hält und unverändert in den APL-Artikel integriert.

Schüler:innen sollten aber nicht nur mit Charakteristika und Konventionen bestimmter bildlicher Repräsentationen vertraut gemacht werden, sondern auch in Hinblick auf MER-Integration gefördert werden. So sollten sie Informationen aus verschiedenen Repräsentationen auch dann miteinander in Verbindung setzen können, wenn sich keine (lösungs-)relevanten Informationen überschneiden (wie in Gruppe *keine Redundanz*). Beck (2017) hat im Rahmen der Entwicklung und Überprüfung eines Kompetenzstrukturmodells zur MER-Integration in biologischen Kontexten einen Kompetenztest mit Aufgaben entwickelt, die eine Integration von Text und – häufig biologiespezifischem – Bild erfordern und im Unterricht zur Konzeption von Lern- und Testaufgaben genutzt werden könnten. Bei Gestaltung von Lern- und Testmaterial könnten Verweise im Text auf bildliche Repräsentationen, die in Schulbüchern zum Beispiel häufig fehlen (z. B. Slough et al., 2010), und Aufgabenstellungen, die die Integration von Text und Bild fordern, neben inhaltlichem Mehrwert zum einen Kompetenzen bezüglich MER-Integration von Schüler:innen fördern. Zum anderen könnten sie helfen, Schüler:innen generell bewusst zu machen, dass bildliche Repräsentationen Informationen transportieren können, die über die Textinformation hinausgehen und nicht als bloßes Beiwerk betrachtet werden sollten. Ein explizites Unterrichten von Strategien zur Text-Bild-Integration in naturwissenschaftlichen Fächern kann für Schüler:innen verschiedener Klassenstufen gewinnbringend sein (z. B. Jian, 2020; Mason et al., 2015) und im Verlauf der Schulzeit zum Aufbau einer entsprechenden Kompetenz der Schüler:innen führen. Hinsichtlich Lernstrategien mit Relevanz beim schulischen Lernen äußert Artelt (2006), dass „sie im Sinne eines ‚Werkzeugkoffers‘ Schülerinnen und Schüler während und auch nach dem Ende ihrer schulischen Laufbahn dazu befähigen, ihr eigenes Lernen zu steuern und effektiv zu gestalten“ (S. 337). Erlernte Strategien zur MER-Integration könnten also auch in einem naturwissenschaftlichen Studium von Nutzen sein.

Gezielt Gelegenheiten auch für die Informationsentnahme aus oder Konstruktion von komplexen, fachspezifischen Repräsentationen zu schaffen, könnte nicht nur die Repräsentationskompetenz fördern, sondern auch die entsprechende Selbstwirksamkeitserwartung der Schüler:innen positiv beeinflussen (Erfolgserfahrungen sind einflussreichste Informationsquelle der Selbstwirksamkeitserwartung, s. Kapitel 2.5.3). Biologielehrkräfte sollten daher nicht immer bestrebt sein, Material und gerade bildliche Repräsentationen möglichst einfach zu gestalten oder nach Einfachheit auszuwählen, sondern Gelegenheiten schaffen, in denen Schüler:innen komplexe und auch fachspezifische Abbildungen nutzen müssen und mögliche auftretende Schwierigkeiten bewusst bewältigen. Wie in Kapitel 2.5.3 beschrieben sollten Lehrkräfte die Quellen der Selbstwirksamkeitserwartung im Unterricht beachten. Zum einen könnte durch verbale Überzeugungsarbeit bei der Bearbeitung von Aufgaben mit schwierig erscheinenden Abbildungen das Eintreten von Erfolgen unterstützt werden und Erfolge mit komplexen Abbildungen in der Biologie sollten konkret als solche zurückgemeldet werden. Zum anderen könnte es förderlich sein, wenn sich Schüler:innen in – von der Lehrkraft bewusst zusammengesetzten – Gruppen eigenständig komplexe Abbildungen wie bestimmte Graphen oder auch wenig vertraute Abbildungen erschließen – beispielsweise eine Karte der mitochondrialen DNA oder phylogenetische Bäume. So könnten Schüler:innen möglicherweise von stellvertretenden Erfahrungen durch Mitschüler:innen profitieren. Das Maß an externer Hilfe durch die Lehrkraft sollte überlegt gewählt sein: Zu wenig Hilfe führt gegebenenfalls zu Misserfolgen und könnte die Selbstwirksamkeitserwartung schwächen. Werden Erfolge aber zu stark auf externe Hilfe statt auf eigene Fähigkeiten zurückgeführt, ist anzunehmen, dass die Erfolge wenig zur Stärkung der Selbstwirksamkeitserwartung beitragen (Bandura, 1986, 1997; vgl. Kapitel 2.5.3). Dies könnte zum Beispiel der Fall sein, wenn Lehrkräfte ihren Schüler:innen fachspezifische, komplexe Abbildungen direkt erklären und ihnen dadurch die Möglichkeit nehmen, positive Bewältigungserfahrungen zu machen. Denkbar wäre für Aufgaben mit komplexen Abbildungen oder Aufgaben, die MER-Integration benötigen, auch der Einsatz von Hilfekarten, um der Heterogenität von Lerngruppen gerecht werden zu können und verschieden leistungsstarken Schüler:innen möglichst eigenständige, positive Bewältigungserfahrungen zu ermöglichen. Eine Stärkung der Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie könnte die Motivation erhöhen, sich an derartige Abbildungen heranzuwagen, sowie das Durchhaltevermögen, wenn bei der Informationsentnahme Schwierigkeiten auftreten (vgl. Kapitel 2.5.2). Aufgrund der Häufigkeit und Bedeutung bildlich transportierter Informationen in der Biologie

und im Biologieunterricht könnte eine Stärkung der genannten Selbstwirksamkeitserwartung zu einem höheren konzeptuellen Verständnis, besseren Lernleistungen sowie besseren Testleistungen beitragen. Diese positiven Effekte einer gestärkten Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen in der Biologie betreffen nicht allein den schulischen Kontext. Wird eine stark ausgeprägte Selbstwirksamkeitserwartung entwickelt, kann diese Schüler:innen über die Schullaufbahn hinaus für zukünftige entsprechende Aufgaben rüsten (Bandura, 1986).

Damit oben genannte Aspekte im Unterricht berücksichtigt werden können, müssen Bedeutung, Potentiale und Schwierigkeiten verschiedener Repräsentationsarten oder ihrer Kombination den Biologielehrkräften – und weiter gefasst Lehrkräften naturwissenschaftlicher Fächer allgemein – zunächst einmal bewusst sein. Zudem sollten die Lehrkräfte mit Blick auf die Vorbereitung von Schüler:innen auf das Abitur, ein potentielles Studium in naturwissenschaftlichen Fächern sowie eine gesellschaftliche und politische Teilhabe Wege kennen, wie sie die Repräsentationskompetenz sowie die diesbezügliche Selbstwirksamkeitserwartung der Schüler:innen stärken können, um ihnen auch für den zukünftigen Werdegang den Zugang zu komplexen und fachspezifischen (M)ER zu erleichtern.

Diese Punkte sollten in die Lehrkräfteausbildung in naturwissenschaftlichen Fächern und möglicherweise auch in Fortbildungen für Lehrkräfte, die sich bereits im Schuldienst befinden, integriert werden (vgl. auch Eilam et al., 2014; Lowe, 2000), der Repräsentationsaspekt wird aber in der Lehrkräfteaus- und Weiterbildung bisher nicht in ausreichendem Maße berücksichtigt (Eilam et al., 2014). Beginnend mit der ersten Phase der Lehrkräftebildung, dem Lehramtsstudium, über die zweite Phase, den Vorbereitungsdienst, bis hin zu Fort- und Weiterbildungen von Lehrkräften im Schuldienst sollten Kurse oder Kurselemente den Einsatz von bzw. das Arbeiten mit Repräsentationen fokussieren. Einen Beitrag hierzu für die erste Phase der Lehrkräftebildung leistet ein Masterseminar für Lehramtsstudierende in der Biologiedidaktik an der Universität Hamburg, das teilweise aufgrund der Ergebnisse der Studien dieser Dissertation konzipiert und durchgeführt wurde und einige der oben genannten Punkte adressiert. Im Rahmen des laufenden Projekts FiRe₂ wurde eine feedbackgestützte Interventionsmaßnahme für angehende Lehrkräfte entwickelt, die auf die Förderung von fachlichem und fachdidaktischem Wissen über Repräsentationen in der Chemie abzielt sowie Überzeugungen der angehenden Lehrkräfte zum Einsatz von Repräsentationen adressiert und als Modul in der zweiten Phase der Lehrkräftebildung, dem Vorbereitungsdienst, eingesetzt werden kann (Tonyali et al., 2019, 2021).

Weitere Seminare oder Seminareinheiten, die für einen reflektierten Einsatz von Repräsentationen sensibilisieren, könnten Eingang in die Lehrkräfteaus- und Weiterbildung finden.

Auch wenn aufgrund der Stichprobe von Studierenden, die sich eher zu Beginn ihres Studiums befanden, in dieser Arbeit keine Rückschlüsse gezogen werden können, wie sich die Repräsentationskompetenz während des Studiums entwickelt, kann festgehalten werden, dass die Studierenden zumindest bereits einige Lerngelegenheiten und auch Testsituationen in Kursen zu biologischen Themen in ihrem Studium erlebt haben, in denen sicherlich MER zum Einsatz kamen. Dennoch hatten sie die genannten Schwierigkeiten bei der Aufgabenbewältigung. Es wäre daher wünschenswert, wenn auch Dozierenden der Biologie oder weiterer naturwissenschaftlicher Fächer an Universitäten bewusst wäre, dass besonders komplexe, fachspezifische bildliche Repräsentationen für die Studierenden nicht unbedingt selbsterklärend sind, und sie den Studierenden möglicherweise Hilfestellungen geben. Auch im Studium könnten zum Beispiel positive Bewältigungserfahrungen beim Verstehen komplexer Abbildungen die diesbezügliche Selbstwirksamkeitserwartung weiter stärken und im weiteren Verlauf des Studiums hilfreich sein. Bei Gestaltung von Leistungstest mit bildlichen Repräsentationen sollte zudem beachtet werden, inwiefern die Aufgaben das Verstehen der bildlichen Repräsentationen, Fachwissen oder beides abfragen (vgl. Kottmeyer et al., 2020).

7.4 Implikationen für zukünftige Forschung

Einige Implikationen für zukünftige Forschung ergeben sich aus den in Kapitel 7.2 angeführten Limitationen dieser Arbeit. Neben der Ausweitung auf weitere biologische Themengebiete und davon unabhängiger Variation der depiktionalen Repräsentationsarten, wäre eine umfangreiche qualitative Untersuchung des Prozesses der Aufgabenbearbeitung mit Eye-Tracking- und Thinking-Aloud-Methode erstrebenswert. So könnten in größerem Umfang als in der Masterarbeitsstudie (Klann, 2018) Erkenntnisse darüber gewonnen werden, wie die Teilnehmenden bildliche Repräsentationen und Text zum Lösen der Aufgabe heranziehen, ob Teilnehmende der Gruppe *Redundanz* zum Beispiel tatsächlich textgeleitet lösungsrelevante Informationen in den Abbildungen identifizieren. Eine Datenerhebung am Computer ermöglicht auch die Durchführung eines Tests zur Erfassung der Arbeitsgedächtniskapazität wie beispielsweise von Wilhelm et al. (2013) in der ursprünglich vorgesehenen Form. Neben dieser objektiven Messung der Arbeitsgedächtniskapazität könnte

zudem der subjektiv empfundene *Cognitive Load* nach der Bearbeitung jeder Aufgabe erfasst werden. Dazu könnten Skalen genutzt werden, die *Intrinsic*, *Extraneous* und *Germane Cognitive Load* (Leppink et al., 2013)²⁴ oder *Mental Load* (erfasst als wahrgenommene Komplexität der Aufgabe) und *Mental Effort* (erfasst als investierte kognitive Anstrengung bei Bearbeitung der Aufgabe) separat erfassen (Krell, 2015, 2017). So könnte – wenn Ergebnisse dieser Arbeit in zukünftigen Arbeiten bestätigt würden – geprüft werden, ob sich die in Bezug auf Ergebnisse in Kapitel 7.1 diskutierten Annahmen zum *Cognitive Load* mit der Wahrnehmung der Teilnehmenden decken. Zusätzlich könnten nach Aufgabebearbeitung Interviews mit den Teilnehmenden geführt werden, um gezielt Informationen über die bei der Bearbeitung aufgetretenen Schwierigkeiten erhalten zu können. Zudem sollte der empirische Nachweis erbracht werden, dass sich die hier untersuchte Selbstwirksamkeitserwartung tatsächlich von einer Selbstwirksamkeitserwartung bezüglich allgemeiner Fähigkeiten im Fach Biologie abgrenzen lässt und, wie beschrieben, auch in anderen thematischen Kontexten einen Effekt auf die Testleistung hat. Die Datenerhebung der vorliegenden Arbeit erfolgte mit Bachelorstudierenden, die sich in den ersten Semestern ihres Studiums befanden. Interessant wäre, welche Ergebnisse Masterstudierende der Biologie erbringen würden, die deutlich mehr Lerngelegenheiten mit komplexen, fachspezifischen Abbildungen bzw. MER hatten. Besonders bei Anfertigung ihrer Bachelorarbeit dürften sie mit zahlreichen wissenschaftlichen Publikationen gearbeitet haben. Möglicherweise würde hier kein signifikanter Unterschied zwischen den Experimentalgruppen mehr auftreten. Zumindest sollten im Optimalfall komplexe, fachspezifische Abbildungen nicht mehr zu schlechterem Abschneiden führen. Mit Blick auf die Ergebnisse dieser Arbeit könnten Fördermaßnahmen für Schüler:innen sowie Studierende entwickelt werden, die das kompetente Arbeiten mit fachspezifischen, komplexen Abbildungen in der Biologie und auch die Text-Bild-Integration adressieren.

Wie mehrfach betont kann eine Förderung des kompetenten Arbeitens mit (M)ER im schulischen Fachunterricht wie dem Biologieunterricht nur dann gelingen, wenn Lehrkräfte entsprechende Lerngelegenheiten schaffen und die Schüler:innen beispielsweise

²⁴ Leppink et al. (2013) gehen in ihrer Diskussion auf die Argumentation von Kalyuga (2011) ein, den *Germane Cognitive Load* nicht als additiv zu *Intrinsic* und *Extraneous Cognitive Load* zu begreifen, sondern als dem Umgang mit *Intrinsic Cognitive Load* statt mit *Extraneous Cognitive Load* gewidmete Ressourcen des Arbeitsgedächtnisses (vgl. Kapitel 2.3.1.3), argumentieren auf Basis ihrer Ergebnisse aber für eine Trennung in *Intrinsic*, *Extraneous* und *Germane Cognitive Load*: „The present findings suggest, however, that such a two-factor framework may not be sufficient; the three-factor solution is consistent across lectures“ (S. 1068).

auch bewusst an komplexe und fachspezifische Abbildungen heranführen und so deren Repräsentationskompetenz und Selbstwirksamkeitserwartung in diesem Bereich stärken. Damit dies gelingen kann, müssen (angehende) Biologielehrkräfte selbst ein gewisses Maß an Repräsentationskompetenz und eine ausreichend hohe Selbstwirksamkeitserwartung besitzen. Traut eine Lehrkraft sich selbst zum Beispiel das Arbeiten mit fachspezifischen, komplexen Abbildungen nicht zu, ist nicht anzunehmen, dass sie derartige Abbildungen gezielt im Unterricht einsetzen wird. Somit lassen sich aufgrund der Befunde dieser Arbeit, die auf eine unzureichende Berücksichtigung dieses Themas im Biologieunterricht schließen lassen, auch Implikationen für die Professionsforschung ableiten.

Hierbei werden verschiedene Aspekte der professionellen Kompetenz von Lehrkräften berührt (vgl. Kompetenzaspekte nach Baumert & Kunter, 2006): Zum einen das Professionswissen (vgl. auch für Diagramme in der Biologie von Kotzebue & Nerdel, 2012), zum anderen mit der Selbstwirksamkeitserwartung auch der Kompetenzaspekt motivationale Orientierungen und zudem der Kompetenzaspekt Überzeugungen und Werthaltungen. Es ergeben sich zum Beispiel folgende Ansatzpunkte für zukünftige Forschung: Anknüpfend an bereits bestehende Instrumente zu Repräsentationen in naturwissenschaftlichen Fächern könnten mit Blick auf die in dieser Dissertation fokussierten Aspekte Redundanz sowie fachspezifische, komplexe Abbildungen Instrumente für die Biologie entwickelt werden zum Erfassen (1) der Repräsentationskompetenz von (angehenden) Lehrkräften (und möglicherweise auch von Schüler:innen) mit explizitem Einschluss von fachspezifischen, komplexen Abbildungen sowie unterschiedlich redundanten Text-Bild-Kombinationen, (2) der Bedeutung, die (angehende) Lehrkräfte verschiedenen Repräsentationsarten als Kommunikations- oder Erkenntnismittel oder auch dem Einsatz von MER aus Fachartikeln im Biologieunterricht beimessen und (3) der Selbstwirksamkeitserwartung bezüglich des Unterrichtens mit fachspezifischen, komplexen Abbildungen in der Biologie.

Die unter (1)–(3) aufgeführten Instrumente sowie bereits bestehende Instrumente zu fachdidaktischem Wissen über Repräsentationen und die in der vorliegenden Arbeit genutzte Selbstwirksamkeitserwartungsskala könnten eingesetzt werden, um zu untersuchen, ob diese Merkmale von Lehrkräften mit Repräsentationskompetenz, konzeptuellem Wissen sowie der Selbstwirksamkeitserwartung bezüglich des Arbeitens mit fachspezifischen, komplexen Abbildungen von deren Schüler:innen zusammenhängen. In Bezug auf Lehramtsstudiengänge der Biologie ist denkbar, im Rahmen einer Längsschnittanalyse zu untersuchen, ob sich oben genannte Punkte bei Studierenden im Verlauf ihres Lehramtsstu-

diums verändern. Mit den genannten Untersuchungen könnte an Arbeiten großer Forschungsprojekte angeschlossen werden, die unter anderem untersuchen, wie sich Kompetenzen von (angehenden) Lehrkräften der Naturwissenschaften bezüglich (multipler) externer Repräsentationen auf Leistungen oder Motivation von Schüler:innen auswirken oder sich während des Lehramtsstudiums entwickeln (u. a. BITE, Schnotz et al., 2010; KiL/KeiLa: z. B. Kleickmann et al., 2014, Mahler et al., 2022, Taskin et al., 2017), und die Ergebnisse um den Aspekt der Selbstwirksamkeitserwartung bezüglich des Arbeitens mit komplexen Abbildungen und Redundanz in MER erweitert werden. Je nach Bedarf könnten Fördermaßnahmen für (angehende) Lehrkräfte entwickelt und in Lehrkräfteaus- und Weiterbildung implementiert werden (vgl. Kapitel 7.3), damit das Unterrichten mit verschiedenen Repräsentationsarten gelingt.

Dem Oberstufenunterricht kommt wie bereits angeführt eine wissenschaftspropädeutische Funktion zu (z. B. KMK, 2020, 2021). Für den naturwissenschaftlichen Unterricht bedeutet dies auch, Schüler:innen zu befähigen, mit MER in wissenschaftlichen Fachartikeln zurechtzukommen, was das Verständnis diverser Repräsentationen und ihrer Kombination einschließt. Nachdem in dieser Arbeit Auswirkungen von Redundanz in MER der Biologie untersucht wurden, wäre eine detaillierte systematische Analyse des Merkmals Redundanz in Fachpublikationen der Biologie ein weiterer nächster Schritt. Eine solche Analyse, die Redundanz zwischen Texten und verschiedenen depiktionalen Repräsentationsarten sowie zudem Fachpublikationen verschiedener Themenbereiche der Biologie (z. B. Ökologie, Evolution, Genetik) in den Blick nimmt und damit einhergehend auch potentielle Unterschiede der Ausprägung der Redundanz in Abhängigkeit von depiktionaler Repräsentationsart oder Themengebiet aufdecken kann, könnte weitere Erkenntnisse darüber liefern, mit welcher Art von MER Personen konfrontiert werden, die sich für ein Biologiestudium beziehungsweise einen Berufsweg in dieser Fachdisziplin entscheiden. Diese Erkenntnisse könnten möglicherweise bereits im Biologieunterricht, besonders in der Oberstufe, berücksichtigt werden.

Für diese Analyse der Redundanz in Fachpublikationen aus der Biologie wurde ein Kategoriensystem entwickelt, mit dem ausgehend von den depiktionalen Repräsentationen eines Fachartikels die Redundanz zwischen in ihnen enthaltenen Informationen und Informationen des Haupttexts sowie der Abbildungsunterschrift oder Tabellenüberschrift erfasst werden kann (Magnus et al., 2019). Dabei werden zum Beispiel Einzelelemente, Färbun-

gen, Werte, Minima und Maxima in depiktionalen Repräsentationen und die Art ihrer Erwähnung in Haupttext oder Abbildungsunterschrift bzw. Tabellenüberschrift in den Blick genommen. Auch wenn in der Analyse das formale Kriterium Redundanz von MER und kein reiner Text untersucht werden soll, wurde nach der Methode der Strukturierung der qualitativen Inhaltsanalyse vorgegangen, wonach die Kategorien des Kategoriensystems deduktiv definiert werden (Mayring, 2010). Die deduktive Kategorienbildung erfolgte anhand von Literatur zur Klassifikation von (depiktionalen) Repräsentationen, Elementen bzw. Merkmalen von depiktionalen Repräsentationen oder zu Text-Bild-Beziehungen, wobei die Kategorien zum Teil mittels Analyse von Artikeln aus hochrangigen Zeitschriften aus dem Bereich der Biologie induktiv ausgeschärft wurden. Letzteres bezog sich meist auf die Art der Erwähnung der Elemente oder Merkmale einer depiktionalen Repräsentation (z. B. Färbung im Mikroskopbild, Minima oder Maxima in einem Liniendiagramm) im Text. Das heißt, um die Ausprägungen einzelner Kategorien festzulegen, wurden teilweise Fachartikel der Biologie herangezogen. Aufgrund des explorativen Charakters dieser Untersuchung konnten nicht alle Kategorien deduktiv gebildet werden. Es wurden zusätzlich anhand genannter Fachartikel Kategorien bezüglich Redundanz induktiv gebildet. Nach weiterer Überarbeitung der Kategorien bei Anwendung auf weitere Fachartikel kann das Kategoriensystem der Analyse der Redundanz von Informationen jeder depiktionalen Repräsentation und genannten Textanteilen (Haupttext und Abbildungsunterschrift/Tabellenüberschrift) von Fachartikeln dienen, wobei auf eine hinreichende Intercoder-Reliabilität geprüft werden muss. Das Kategoriensystem könnte zudem nicht nur für die Analyse von Redundanz in Fachpublikationen genutzt werden, sondern auch für die detaillierte systematische Analyse von Redundanz in anderen Textgenres wie populärwissenschaftlichen Artikeln oder Lehrbüchern zur Verfügung stehen.

7.5 Ausblick

Die Ergebnisse dieser Arbeit weisen unter anderem darauf hin, dass im Biologieunterricht das Arbeiten mit komplexen, fachspezifischen Abbildungen und verschiedenartigen MER stärker gefördert werden sollte. Auf Deutschland bezogen lässt sich festhalten, dass die relativ neuen Bildungsstandards im Fach Biologie für die Allgemeine Hochschulreife (KMK, 2020) im Rahmen der Kommunikationskompetenz die Fähigkeit aufführen, auch mittels „komplexer Darstellungsformen“ (S. 16) Informationen erschließen und „fachtypische[] Darstellungen“ (S. 16) zur Kommunikation nutzen zu können. Somit ist

zu hoffen, dass die Berücksichtigung dieser normativen Vorgaben in den nächsten Jahren verstärkt Eingang in die Unterrichtspraxis findet und Biologieunterricht auch in diesem Punkt dem Ziel wissenschaftspropädeutischer Bildung näher kommt. Es wäre zu überlegen, diese Vorgaben noch explizit um das Arbeiten mit verschiedenartigen MER zu ergänzen und die genannten Punkte (d. h. auch das Arbeiten mit komplexen Darstellungsformen) auch in die Bildungsstandards für den Mittleren Schulabschluss zu integrieren. Schließlich ist die Fähigkeit, mittels verschiedenartiger, auch komplexer Repräsentationen kommunizieren zu können, Informationen aus diesen entnehmen und sie interpretieren zu können, ein wichtiger Aspekt naturwissenschaftlicher Grundbildung, die zur aktiven gesellschaftlichen und politischen Teilhabe befähigt (vgl. Kapitel 1 und 2.1). Hier sei beispielsweise an die eingangs thematisierte öffentliche Kommunikation zur COVID-19-Pandemie erinnert, die unter anderem mittels verschiedener bildlicher Repräsentationen stattfindet, die nicht immer einfach zu verstehen sind.

Während schulischer Fachunterricht also stärker darauf hinarbeiten sollte, Schüler:innen in die Lage zu versetzen, mit verschiedensten, auch komplexen, fachspezifischen Repräsentationen der Biologie arbeiten zu können, stellt sich gleichzeitig von der anderen Seite betrachtet für Forschende die Frage, wie Forschungsergebnisse so aufbereitet werden können, dass sie für die Öffentlichkeit verständlich sind. Im Sommer 2021 wurde die Förderung von vier neuen Zentren zur Wissenschaftskommunikationsforschung verkündet (VolkswagenStiftung, 2021). Eines dieser Konsortien, das *Kiel Science Communication Network*, beschäftigt sich explizit mit dem Gelingen und somit der Beforschung von Wissenschaftskommunikation mittels Visualisierungen zu Gesundheitsthemen. Wenn das Anstreben gelingender Wissenschaftskommunikation insgesamt – auch durch derartige Projekte – breitflächig stärker ins Bewusstsein von Forschenden gelangt und berücksichtigt wird, könnten in Zukunft Bestrebungen von Schulunterricht und Bestrebungen von Forschenden beiderseits stärker zum Gelingen eines öffentlichen Diskurses zu naturwissenschaftlichen Themen, der reich an bildlichen Repräsentationen ist, beitragen.

Des Weiteren ist eine Publikation zur Konzeption des in Kapitel 7.3 vorgestellten, teilweise aufgrund der Forschungsergebnisse des in dieser Dissertation dargelegten Promotionsprojekts entwickelten Masterseminars zur Kommunikationskompetenz geplant. Diese soll Dozierenden von Kursen für Lehramtsstudierende naturwissenschaftlicher Fächer einen Ansatz vorstellen, wie dieses wichtige Thema explizit in die Lehramtsausbildung integriert werden könnte. Im Optimalfall sind sich die am Kurs teilnehmenden Studierenden bei zukünftiger Ausübung ihres Berufs als Naturwissenschaftslehrkräfte der Potentiale und

Schwierigkeiten des Arbeitens mit – insbesondere bestimmten bildlichen – Repräsentationen bewusst, setzen (multiple) Repräsentationen reflektiert ein und fördern Schüler:innen auch gezielt hinsichtlich ihrer Selbstwirksamkeitserwartung und Repräsentationskompetenz.

8 Literatur²⁵

- Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education*, 33(2–3), 131–152. [https://doi.org/10.1016/S0360-1315\(99\)00029-9](https://doi.org/10.1016/S0360-1315(99)00029-9)
- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183–198. <https://doi.org/10.1016/j.learninstruc.2006.03.001>
- Ainsworth, S. (2014). The multiple representation principle in multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (2. Aufl., S. 464–486). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.024>
- Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, 333(6046), 1096–1097. <https://doi.org/10.1126/science.1204153>
- Ardura, D., & Galán, A. (2019). The interplay of learning approaches and self-efficacy in secondary school students' academic achievement in science. *International Journal of Science Education*, 41(13), 1723–1743. <https://doi.org/10.1080/09500693.2019.1638981>
- Ardura, D., & Pérez-Bitrián, A. (2019). Motivational pathways towards academic achievement in physics & chemistry: A comparison between students who opt out and those who persist. *Chemistry Education Research and Practice*, 20(3), 618–632. <https://doi.org/10.1039/c9rp00073a>
- Arsenault, D. J., Smith, L. D., & Beauchamp, E. A. (2006). Visual inscriptions in the scientific hierarchy: Mapping the “treasures of science”. *Science Communication*, 27(3), 376–428. <https://doi.org/10.1177/1075547005285030>
- Artelt, C. (2006). Lernstrategien in der Schule. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 337–351). Hogrefe.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Hrsg.), *The Psychology of learning and motivation* (Bd. 2, S. 89–195). Academic Press.
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556–559. <https://doi.org/10.1126/science.1736359>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall, Inc.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman and Company.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>

²⁵ Dieses Literaturverzeichnis umfasst die im übergreifenden Rahmen dieser Dissertation verwendeten Quellen. Die Literaturangaben der eingebundenen Artikel sind in Kapitel 4 bzw. Kapitel 5 in eigenständigen Literaturverzeichnissen enthalten.

- Bandura, A., Adams, N. E., & Beyer, J. (1977). Cognitive processes mediating behavioral change. *Journal of Personality and Social Psychology*, 35(3), 125–139. <https://doi.org/10.1037/0022-3514.35.3.125>
- Bartimote-Aufflick, K., Bridgeman, A., Walker, R., Sharma, M., & Smith, L. (2016). The study, evaluation, and improvement of university student self-efficacy. *Studies in Higher Education*, 41(11), 1918–1942. <https://doi.org/10.1080/03075079.2014.999319>
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Beck, C. H. (2017). *Biologiespezifisches Bildverständnis beim Umgang mit multiplen externen Repräsentationen: Entwicklung und Validierung eines Kompetenzmodells zur Integration von Diagrammen und Schemata* [Dissertation, Technische Universität München]. mediaTUM. <https://mediatum.ub.tum.de/1368009>
- Benight, C. C., & Bandura, A. (2004). Social cognitive theory of posttraumatic recovery: The role of perceived self-efficacy. *Behaviour Research and Therapy*, 42(10), 1129–1148. <https://doi.org/10.1016/j.brat.2003.08.008>
- Bjorklund, K. (2021, Januar 11). ‘We’re in a worse place today than we were before he came in’: Former U.S. Secretary of State Rex Tillerson on the mess Donald Trump is leaving behind. *Foreign Policy*, <https://foreignpolicy.com/2021/01/11/rex-tillerson-interview-trump/>
- Bobis, J., Sweller, J., & Cooper, M. (1993). Cognitive load effects in a primary-school geometry task. *Learning and Instruction*, 3(1), 1–21. [https://doi.org/10.1016/S0959-4752\(09\)80002-9](https://doi.org/10.1016/S0959-4752(09)80002-9)
- Bong, M. (1997). Generality of academic self-efficacy judgments: Evidence of hierarchical relations. *Journal of Educational Psychology*, 89(4), 696–709. <https://doi.org/10.1037/0022-0663.89.4.696>
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, 15(1), 1–40. <https://doi.org/10.1023/A:1021302408382>
- Bowen, G. M., & Roth, W.-M. (2002). Why students may not learn to interpret scientific inscriptions. *Research in Science Education*, 32(3), 303–327. <https://doi.org/10.1023/A:1020833231966>
- Brandstetter-Korinth, M. (2017). *Abbildungen im Biologieunterricht*. Logos.
- Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, 43(5), 485–499. <https://doi.org/10.1002/tea.20131>
- Brown, I., & Inouye, D. K. (1978). Learned helplessness through modeling: The role of perceived similarity in competence. *Journal of Personality and Social Psychology*, 36(8), 900–908. <https://doi.org/10.1037/0022-3514.36.8.900>
- Brown, S. D., Lent, R. W., & Larkin, K. C. (1989). Self-efficacy as a moderator of scholastic aptitude – academic performance relationships. *Journal of Vocational Behavior*, 35(1), 64–75. [https://doi.org/10.1016/0001-8791\(89\)90048-1](https://doi.org/10.1016/0001-8791(89)90048-1)

- Burns, E. C., Martin, A. J., Kennett, R. K., Pearson, J., & Munro-Smith, V. (2021). Optimizing science self-efficacy: A multilevel examination of the moderating effects of anxiety on the relationship between self-efficacy and achievement in science. *Contemporary Educational Psychology*, *64*, Article 101937. <https://doi.org/10.1016/j.cedpsych.2020.101937>
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2
- Cook, M., Carter, G., & Wiebe, E. N. (2008). The interpretation of cellular transport graphics by students with low and high prior knowledge. *International Journal of Science Education*, *30*(2), 239–261. <https://doi.org/10.1080/09500690601187168>
- Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, *90*(6), 1073–1091. <https://doi.org/10.1002/sci.20164>
- Corradi, D., Elen, J., & Clarebout, G. (2012). Understanding and enhancing the use of multiple external representations in chemistry education. *Journal of Science Education and Technology*, *21*(6), 780–795. <https://doi.org/10.1007/s10956-012-9366-z>
- Corradi, D., Trinenuh, D. T., Clarebout, G., & Elen, J. (2014). How multiple external representations can help or constrain learning in science. *Journal of Cognitive Education and Psychology*, *13*(3), 411–423. <https://doi.org/10.1891/1945-8959.13.3.411>
- Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology*, *35*(1), 59–74. <https://doi.org/10.1016/j.cedpsych.2009.10.002>
- Department for Education. (2015). *Statutory guidance – National curriculum in England: Science programmes of study. Updated 6 May 2015*, <https://www.gov.uk/government/publications/national-curriculum-in-england-science-programmes-of-study/national-curriculum-in-england-science-programmes-of-study>
- DiBenedetto, M. K., & Bembenuy, H. (2013). Within the pipeline: Self-regulated learning, self-efficacy, and socialization among college students in science courses. *Learning and Individual Differences*, *23*, 218–224. <https://doi.org/10.1016/j.lindif.2012.09.015>
- diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, *22*(3), 293–331. https://doi.org/10.1207/s1532690xci2203_2
- diSessa, A. A., & Sherin, B. L. (2000). Meta-representation: an introduction. *The Journal of Mathematical Behavior*, *19*(4), 385–398. [https://doi.org/10.1016/S0732-3123\(01\)00051-7](https://doi.org/10.1016/S0732-3123(01)00051-7)
- Eilam, B., & Gilbert, J. K. (2014). The significance of visual representations in the teaching of science. In B. Eilam & J. K. Gilbert (Hrsg.), *Science teachers' use of visual representations* (S. 3–28). Springer. https://doi.org/10.1007/978-3-319-06526-7_1
- Eilam, B., & Poyas, Y. (2008). Learning with multiple representations: Extending multimedia learning beyond the lab. *Learning and Instruction*, *18*(4), 368–378. <https://doi.org/10.1016/j.learninstruc.2007.07.003>

- Eilam, B., Poyas, Y., & Hashimshoni, R. (2014). Representing visually: What teachers know and what they prefer. In B. Eilam & J. K. Gilbert (Hrsg.), *Science teachers' use of visual representations* (S. 53–83). Springer. https://doi.org/10.1007/978-3-319-06526-7_3
- Farheen, A., & Lewis, S. E. (2021). The impact of representations of chemical bonding on students' predictions of chemical properties. *Chemistry Education Research and Practice*, 22(4), 1035–1053. <https://doi.org/10.1039/D1RP00070E>
- Florian, C., Schmiemann, P., & Sandmann, A. (2015). Aufgaben im Zentralabitur Biologie – eine kategoriengestützte Analyse charakteristischer Aufgabenmerkmale schriftlicher Abituraufgaben. *Zeitschrift für Didaktik der Naturwissenschaften*, 21(1), 69–86. <https://doi.org/10.1007/s40573-015-0026-8>
- Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E., & Pekrun, R. (Hrsg.). (2009). *PISA 2006 Skalenhandbuch: Dokumentation der Erhebungsinstrumente*. Waxmann.
- Galla, B. M., Wood, J. J., Tsukayama, E., Har, K., Chiu, A. W., & Langer, D. A. (2014). A longitudinal multilevel model analysis of the within-person and between-person effect of effortful engagement and academic self-efficacy on academic performance. *Journal of School Psychology*, 52(3), 295–308. <https://doi.org/10.1016/j.jsp.2014.04.001>
- Garner, R., Alexander, P. A., Gillingham, M. G., Kulikowich, J. M., & Brown, R. (1991). Interest and learning from text. *American Educational Research Journal*, 28(3), 643–659. <https://doi.org/10.3102/00028312028003643>
- Grembowski, D., Patrick, D., Diehr, P., Durham, M., Beresford, S., Kay, E., & Hecht, J. (1993). Self-efficacy and health behavior among older adults. *Journal of Health and Social Behavior*, 34(2), 89–104. <https://doi.org/10.2307/2137237>
- Halverson, K. L. (2011). Improving tree-thinking one learnable skill at a time. *Evolution: Education and Outreach*, 4, 95–106. <https://doi.org/10.1007/s12052-010-0307-0>
- Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology*, 24(2), 95–123. <https://doi.org/10.1006/ceps.1998.0987>
- Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32(6), 717–742. <https://doi.org/10.1006/jmla.1993.1036>
- Herrlinger, S., Höffler, T. N., Opfermann, M., & Leutner, D. (2017). When do pictures help learning from expository text? Multimedia and modality effects in primary schools. *Research in Science Education*, 47(3), 685–704. <https://doi.org/10.1007/s11165-016-9525-y>
- Heßler, M. (2006). Einleitung. Annäherungen an Wissenschaftsbilder. In M. Heßler (Hrsg.), *Konstruierte Sichtbarkeiten: Wissenschafts- und Technikbilder seit der Frühen Neuzeit* (S. 11–37). Wilhelm Fink Verlag.
- Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, 17, 63–84. <https://doi.org/10.1016/j.edurev.2015.11.002>

- Jansen, M., Scherer, R., & Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology*, *41*, 13–24. <https://doi.org/10.1016/j.cedpsych.2014.11.002>
- Jian, Y.-C. (2017). Eye-movement patterns and reader characteristics of students with good and poor performance when reading scientific text with diagrams. *Reading and Writing*, *30*(7), 1447–1472. <https://doi.org/10.1007/s11145-017-9732-6>
- Jian, Y.-C. (2020). Teaching fourth-grade students of different reading abilities to read biological illustrations and integrate in-text information: An empirical experiment. *Research in Science Education*, *50*(6), 2269–2282. <https://doi.org/10.1007/s11165-018-9778-8>
- Joët, G., Usher, E. L., & Bressoux, P. (2011). Sources of self-efficacy: An investigation of elementary school students in France. *Journal of Educational Psychology*, *103*(3), 649–663. <https://doi.org/10.1037/a0024048>
- 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>
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, *38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4
- Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (2. Aufl., S. 247–262). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.013>
- Kirbulut, Z. D., & Uzuntiryaki-Kondakci, E. (2019). Examining the mediating effect of science self-efficacy on the relationship between metavariablen and science achievement. *International Journal of Science Education*, *41*(8), 995–1014. <https://doi.org/10.1080/09500693.2019.1585594>
- Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, *12*(1), 1–10. [https://doi.org/10.1016/S0959-4752\(01\)00014-7](https://doi.org/10.1016/S0959-4752(01)00014-7)
- Klann, K. (2018). *Multiple externe Repräsentationen in Biologieaufgaben: Eine Eye-Tracking-Studie zu Redundanz in textlichen und bildlichen Repräsentationen* [Unveröffentlichte Masterarbeit]. Christian-Albrechts-Universität zu Kiel.
- Klassen, R. M. (2004). A cross-cultural investigation of the efficacy beliefs of South Asian immigrant and Anglo Canadian nonimmigrant early adolescents. *Journal of Educational Psychology*, *96*(4), 731–742. <https://doi.org/10.1037/0022-0663.96.4.731>
- Kleickmann, T., Großschedl, J., Harms, U., Heinze, A., Herzog, S., Hohenstein, F., Köller, O., Kröger, J., Lindmeier, A., Loch, C., Mahler, D., Möller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V., & Zimmermann, F. (2014). Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer – Testentwicklung im Rahmen des Projekts KiL. *Unterrichtswissenschaft*, *42*(3), 280–288.

- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2004). *Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie: Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 05.02.2004*, https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1989/1989_12_01-EPA-Biologie.pdf
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2005). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss: Beschluss vom 16.12.2004*. Luchterhand.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2006). *Vereinbarung zur Gestaltung der gymnasialen Oberstufe in der Sekundarstufe II: Beschluss der Kultusministerkonferenz vom 07.07.1972 i.d.F. vom 02.06.2006*, <https://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/1999/Vereinb-z-Gestalt-d-gymOb-i-d-SekII.pdf>
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2020). *Bildungsstandards im Fach Biologie für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.06.2020*, https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2020/2020_06_18-BildungsstandardsAHR_Biologie.pdf
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2021). *Vereinbarung zur Gestaltung der gymnasialen Oberstufe und der Abiturprüfung: Beschluss der Kultusministerkonferenz vom 07.07.1972 i. D. F. vom 18.02.2021*, https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1972/1972_07_07-VB-gymnasiale-Oberstufe-Abiturpruefung.pdf
- Kottmeyer, A. M., Van Meter, P., & Cameron, C. (2020). Diagram comprehension ability of college students in an introductory biology course. *Advances in Physiology Education*, 44(2), 169–180. <https://doi.org/10.1152/advan.00146.2018>
- Kozma, R., Chin, E., Russell, J., & Marx, N. (2000). The roles of representations and tools in the chemistry laboratory and their implications for chemistry learning. *Journal of the Learning Sciences*, 9(2), 105–143. https://doi.org/10.1207/s15327809jls0902_1
- Kozma, R. B., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, 34(9), 949–968. [https://doi.org/10.1002/\(SICI\)1098-2736\(199711\)34:9<949::AID-TEA7>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1098-2736(199711)34:9<949::AID-TEA7>3.0.CO;2-U)
- Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. K. Gilbert (Hrsg.), *Visualization in science education* (S. 121–145). Springer. https://doi.org/10.1007/1-4020-3613-2_8
- Kragten, M., Admiraal, W., & Rijlaarsdam, G. (2013). Diagrammatic literacy in secondary science education. *Research in Science Education*, 43(5), 1785–1800. <https://doi.org/10.1007/s11165-012-9331-0>
- Krell, M. (2015). Evaluating an instrument to measure mental load and mental effort using Item Response Theory. *Science Education Review Letters, Research Letters 2015*, 1–6. <https://doi.org/10.18452/8212>

- Krell, M. (2017). Evaluating an instrument to measure mental load and mental effort considering different sources of validity evidence. *Cogent Education*, 4(1), Article 1280256. <https://doi.org/10.1080/2331186X.2017.1280256>
- Krey, O., & Schwanewedel, J. (2018). Lernen mit externen Repräsentationen. In D. Krüger, I. Parchmann, & H. Schecker (Hrsg.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (S. 159–175). Springer. https://doi.org/10.1007/978-3-662-56320-5_10
- Lachmayer, S. (2008). *Entwicklung und Überprüfung eines Strukturmodells der Diagrammkompetenz für den Biologieunterricht* [Dissertation, Christian-Albrechts-Universität zu Kiel]. MACAU. https://macau.uni-kiel.de/receive/diss_mods_00003041
- Lachmayer, S., Nerdel, C., & Prechtel, H. (2007). Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 145–160.
- Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching*, 44(5), 706–724. <https://doi.org/10.1002/tea.20172>
- Lemke, J. L. (1998). *Teaching all the languages of science: words, symbols, images, and actions* [Paper presentation]. International Conference on Ideas for a Scientific Culture, Barcelona, Spanien.
- Lemke, J. L. (2004). The literacies of science. In E. W. Saul (Hrsg.), *Crossing borders in literacy and science instruction: Perspectives on theory and practice* (S. 33–47). International Reading Association.
- Lent, R. W., Brown, S. D., & Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology*, 33(3), 265–269. <https://doi.org/10.1037/0022-0167.33.3.265>
- Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1991). Mathematics self-efficacy: Sources and relation to science-based career choice. *Journal of Counseling Psychology*, 38(4), 424–430. <https://doi.org/10.1037/0022-0167.38.4.424>
- 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. <https://doi.org/10.3758/s13428-013-0334-1>
- Lindner, M. A., Ihme, J. M., Saß, S., & Köller, O. (2018). How representational pictures enhance students' performance and test-taking pleasure in low-stakes assessment. *European Journal of Psychological Assessment*, 34(6), 376–385. <https://doi.org/10.1027/1015-5759/a000351>
- Lowe, R. (2000). *Visual literacy and learning in science* (ED463945). ERIC. <https://files.eric.ed.gov/fulltext/ED463945.pdf>
- Magnus, L., Schütte, K., & Schwanewedel, J. (2019). *Category system for the analysis of redundancy in multiple external representations in scientific biology journals* [Unpublished material]. IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel.

- Mahler, D., Bock, D., Schauber, S., & Harms, U. (2022). *Using longitudinal models to unfold preservice science teachers' development of content knowledge and pedagogical content knowledge* [Manuscript in preparation].
- Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*(2), 331–353. <https://doi.org/10.1037/edu0000281>
- Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology, 41*, 172–187. <https://doi.org/10.1016/j.cedpsych.2015.01.004>
- Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education, 81*(3), 356–384. <https://doi.org/10.1080/00220973.2012.727885>
- Matsui, T., Matsui, K., & Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior, 37*(2), 225–238. [https://doi.org/10.1016/0001-8791\(90\)90042-Z](https://doi.org/10.1016/0001-8791(90)90042-Z)
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139164603>
- Mayer, R. E. (2005). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 183–200). Cambridge University Press.
- Mayer, R. E. (2011). Instruction based on visualizations. In R. E. Mayer & P. A. Alexander (Hrsg.), *Handbook of research on learning and instruction* (S. 427–445). Routledge.
- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (2. Aufl., S. 43–71). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- Mayer, R. E. (2019). Taking a new look at seductive details. *Applied Cognitive Psychology, 33*(1), 139–141. <https://doi.org/10.1002/acp.3503>
- Mayer, R. E. (2020). *Multimedia learning* (3. Aufl.). Cambridge University Press. <https://doi.org/10.1017/9781316941355>
- Mayer, R. E., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (2. Aufl., S. 279–315). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.015>
- Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187–198. <https://doi.org/10.1037/0022-0663.93.1.187>
- Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (11., aktual. und überarb. Aufl.). Beltz.

- Millar, M., & Schrier, T. (2015). Digital or printed textbooks: Which do students prefer and why? *Journal of Teaching in Travel & Tourism*, *15*(2), 166–185. <https://doi.org/10.1080/15313220.2015.1026474>
- Miller, B. W., Cromley, J. G., & Newcombe, N. S. (2016). Improving diagrammatic reasoning in middle school science using conventions of diagrams instruction. *Journal of Computer Assisted Learning*, *32*(4), 374–390. <https://doi.org/10.1111/jcal.12143>
- NGSS Lead States. (2013). *Next generation science standards: for states, by states*. The National Academies Press.
- Nitz, S. (2012). *Fachsprache im Biologieunterricht: Eine Untersuchung zu Bedingungsfaktoren und Auswirkungen* [Dissertation, Christian-Albrechts-Universität zu Kiel]. MACAU. https://macau.uni-kiel.de/receive/diss_mods_00008550
- Nitz, S., Ainsworth, S. E., Nerdel, C., & Prechtel, H. (2014). Do student perceptions of teaching predict the development of representational competence and biological knowledge? *Learning and Instruction*, *31*, 13–22. <https://doi.org/10.1016/j.learninstruc.2013.12.003>
- Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, *87*(2), 224–240. <https://doi.org/10.1002/sce.10066>
- Novick, L. R., Stull, A. T., & Catley, K. M. (2012). Reading phylogenetic trees: The effects of tree orientation and text processing on comprehension. *BioScience*, *62*(8), 757–764. <https://doi.org/10.1525/bio.2012.62.8.8>
- Obrentz, S. B. (2012). *Predictors of science success: The impact of motivation and learning strategies on college chemistry performance* [Doctoral dissertation, Georgia State University]. ScholarWorks @ Georgia State University. https://scholarworks.gsu.edu/epse_diss/77
- Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (2. Aufl., S. 27–42). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.004>
- Paivio, A. (1986). *Mental Representations: A dual coding approach*. Oxford University Press.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255–287. <https://doi.org/10.1037/h0084295>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, *66*(4), 543–578. <https://doi.org/10.3102/00346543066004543>
- Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performances: The need for specificity of assessment. *Journal of Counseling Psychology*, *42*(2), 190–198. <https://doi.org/10.1037/0022-0167.42.2.190>
- Partin, M. L., & Haney, J. J. (2012). The CLEM model: Path analysis of the mediating effects of attitudes and motivational beliefs on the relationship between perceived learning environment and course performance in an undergraduate non-major biology course. *Learning Environments Research*, *15*(1), 103–123. <https://doi.org/10.1007/s10984-012-9102-x>

- Peirce, C. S. S. (1906). Prolegomena to an apology for pragmatism. *The Monist*, 16(4), 492–546. <https://doi.org/10.5840/monist190616436>
- Pociask, F. D., & Morrison, G. R. (2008). Controlling split attention and redundancy in physical therapy instruction. *Educational Technology Research and Development*, 56(4), 379–399. <https://doi.org/10.1007/s11423-007-9062-5>
- Pozzer-Ardenghi, L., & Roth, W.-M. (2010). Toward a social practice perspective on the work of reading inscriptions in science texts. *Reading Psychology*, 31(3), 228–253. <https://doi.org/10.1080/02702710903256361>
- Reder, L. M., & Anderson, J. R. (1980). A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning and Verbal Behavior*, 19(2), 121–134. [https://doi.org/10.1016/S0022-5371\(80\)90122-X](https://doi.org/10.1016/S0022-5371(80)90122-X)
- Reder, L. M., & Anderson, J. R. (1982). Effects of spacing and embellishment on memory for the main points of a text. *Memory & Cognition*, 10(2), 97–102. <https://doi.org/10.3758/BF03209210>
- Roth, W.-M., Bowen, G. M., & McGinn, M. K. (1999). Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching*, 36(9), 977–1019. [https://doi.org/10.1002/\(SICI\)1098-2736\(199911\)36:9<977::AID-TEA3>3.0.CO;2-V](https://doi.org/10.1002/(SICI)1098-2736(199911)36:9<977::AID-TEA3>3.0.CO;2-V)
- Saß, S., Wittwer, J., Senkbeil, M., & Köller, O. (2012). Pictures in test items: Effects on response time and response correctness. *Applied Cognitive Psychology*, 26(1), 70–81. <https://doi.org/10.1002/acp.1798>
- Schnotz, W. (1994). Wissenserwerb mit logischen Bildern. In B. Weidenmann (Hrsg.), *Wissenserwerb mit Bildern: Instruktionale Bilder in Printmedien, Film/Video und Computerprogrammen* (S. 95–147). Hans Huber.
- Schnotz, W. (1996). Lesen als Textverarbeitung. In H. Günther & O. Ludwig (Hrsg.), *Schrift und Schriftlichkeit: Ein interdisziplinäres Handbuch internationaler Forschung* (Bd. 2, S. 972–982). Walter de Gruyter.
- Schnotz, W. (2001). Wissenserwerb mit Multimedia. *Unterrichtswissenschaft*, 29(4), 292–318. <https://doi.org/10.25656/01:7717>
- Schnotz, W. (2003). Bild- und Sprachverarbeitung aus psychologischer Sicht. In K. Sachs-Hombach (Hrsg.), *Was ist Bildkompetenz? Studien zur Bildwissenschaft* (S. 25–42). Deutscher Universitätsverlag. https://doi.org/10.1007/978-3-663-11814-5_3
- Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 49–69). Cambridge University Press.
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (2. Aufl., S. 72–103). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.006>
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction*, 13(2), 141–156. [https://doi.org/10.1016/s0959-4752\(02\)00017-8](https://doi.org/10.1016/s0959-4752(02)00017-8)
- Schnotz, W., Horz, H., McElvany, N., Schroeder, S., Ullrich, M., Baumert, J., Hachfeld, A., & Richter, T. (2010). Das BITE-Projekt: Integrative Verarbeitung von Bildern

- und Texten in der Sekundarstufe I. In E. Klieme, D. Leutner, & M. Kenk (Hrsg.), *Kompetenzmodellierung: Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. Zeitschrift für Pädagogik*, 56. Beiheft (S. 143–153). Beltz. <https://doi.org/10.25656/01:3388>
- Schöber, C., Schütte, K., Köller, O., McElvany, N., & Gebauer, M. M. (2018). Reciprocal effects between self-efficacy and achievement in mathematics and reading. *Learning and Individual Differences*, 63, 1–11. <https://doi.org/10.1016/j.lindif.2018.01.008>
- Schönborn, K. J., & Anderson, T. R. (2006). The importance of visual literacy in the education of biochemists. *Biochemistry and Molecular Biology Education*, 34(2), 94–102. <https://doi.org/10.1002/bmb.2006.49403402094>
- Schönborn, K. J., Anderson, T. R., & Grayson, D. J. (2002). Student difficulties with the interpretation of a textbook diagram of immunoglobulin G (IgG). *Biochemistry and Molecular Biology Education*, 30(2), 93–97. <https://doi.org/10.1002/bmb.2002.494030020036>
- Schroeder, S., Richter, T., McElvany, N., Hachfeld, A., Baumert, J., Schnotz, W., Horz, H., & Ullrich, M. (2011). Teachers' beliefs, instructional behaviors, and students' engagement in learning from texts with instructional pictures. *Learning and Instruction*, 21(3), 403–415. <https://doi.org/10.1016/j.learninstruc.2010.06.001>
- Schunk, D. H., & DiBenedetto, M. K. (2016). Self-efficacy theory in education. In K. R. Wentzel & D. B. Miele (Hrsg.), *Handbook of motivation at school* (S. 34–54). Routledge. <https://www.routledgehandbooks.com/doi/10.4324/9781315773384.ch3>
- Schunk, D. H., & Usher, E. L. (2012). Social cognitive theory and motivation. In R. M. Ryan (Hrsg.), *The Oxford handbook of human motivation* (S. 13–27). Oxford University Press.
- Schütte, K., Frenzel, A. C., Asseburg, R., & Pekrun, R. (2007). Schülermerkmale, naturwissenschaftliche Kompetenz und Berufserwartung. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme, & R. Pekrun (Hrsg.), *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 125–146). Waxmann. <https://epub.ub.uni-muenchen.de/61990/>
- Schwarzer, R., & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (S. 28–53), *Zeitschrift für Pädagogik*, 44. Beiheft. Beltz.
- Seufert, T. (2003). Kohärenzbildung beim Wissenserwerb mit multiplen Repräsentationen. In K. Sachs-Hombach (Hrsg.), *Was ist Bildkompetenz? Studien zur Bildwissenschaft* (S. 117–129). Deutscher Universitäts-Verlag. https://doi.org/10.1007/978-3-663-11814-5_8
- Slough, S. W., McTigue, E. M., Kim, S., & Jennings, S. K. (2010). Science textbooks' use of graphical representation: A descriptive analysis of four sixth grade science texts. *Reading Psychology*, 31(3), 301–325. <https://doi.org/10.1080/02702710903256502>
- Stäudel, L., Franke-Braun, G., & Parchmann, I. (2008). Sprache, Kommunikation und Wissenserwerb im Chemieunterricht. *Unterricht Chemie*, 19(106/107), 4–9.

- Sweller, J. (2005a). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 19–30). Cambridge University Press.
- Sweller, J. (2005b). The redundancy principle in multimedia learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 159–167). Cambridge University Press.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. <https://doi.org/10.1007/978-1-4419-8126-4>
- Taskin, V., Bernholt, S., & Parchmann, I. (2017). Student teachers' knowledge about chemical representations. *International Journal of Science and Mathematics Education*, *15*(1), 39–55. <https://doi.org/10.1007/s10763-015-9672-z>
- Tonyali, B., Ropohl, M., & Schwanewedel, J. (2019). Feedback an Lehramtsanwärterinnen und -anwärter zum Einsatz externer Repräsentationen im naturwissenschaftlichen Unterricht. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe: Gesellschaft für Didaktik der Chemie und Physik – Jahrestagung in Kiel 2018* (S. 803–806). Universität Regensburg.
- Tonyali, B., Ropohl, M., & Schwanewedel, J. (2021). Optimierung von Arbeitsblättern durch Feedback in Chemie. In S. Habig (Hrsg.), *Naturwissenschaftlicher Unterricht und Lehrerbildung im Umbruch? Gesellschaft für Didaktik der Chemie und Physik online Jahrestagung 2020* (S. 721–724). Universität Duisburg-Essen.
- Trafton, J. G., Marshall, S., Mintz, F., & Trickett, S. B. (2002). Extracting explicit and implicit information from complex visualizations. In M. Hegarty, B. Meyer & N. H. Narayanan (Hrsg.), *Diagrammatic representation and inference. Diagrams 2002. Lecture Notes in Computer Science* (Bd. 2317, S. 206–220). Springer. https://doi.org/10.1007/3-540-46037-3_22
- Treagust, D. F., & Tsui, C.-Y. (2013). Conclusion: Contributions of multiple representations to biological education. In D. F. Treagust & C.-Y. Tsui (Hrsg.), *Multiple representations in biological education* (S. 349–367). Springer. https://doi.org/10.1007/978-94-007-4192-8_19
- Tsui, C.-Y., & Treagust, D. F. (2013). Introduction to multiple representations: Their importance in biology and biological education. In D. F. Treagust & C.-Y. Tsui (Hrsg.), *Multiple representations in biological education* (S. 3–18). Springer. https://doi.org/10.1007/978-94-007-4192-8_1
- Uçar, F. M., & Sungur, S. (2017). The role of perceived classroom goal structures, self-efficacy, and engagement in student science achievement. *Research in Science & Technological Education*, *35*(2), 149–168. <https://doi.org/10.1080/02635143.2017.1278684>
- Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research*, *78*(4), 751–796. <https://doi.org/10.3102/0034654308321456>
- Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

- van Merriënboer, J. J. G., & Kester, L. (2005). The four-component instructional design model: Multimedia principles in environments for complex learning. In R. E. Mayer (Hrsg.), *The Cambridge handbook of multimedia learning* (S. 71–93). Cambridge University Press.
- Villavicencio, F. T., & Bernardo, A. B. I. (2013). Negative emotions moderate the relationship between self-efficacy and achievement of Filipino students. *Psychological Studies*, 58(3), 225–232. <https://doi.org/10.1007/s12646-013-0193-y>
- VolkswagenStiftung. (2021, 25. Juni). *VolkswagenStiftung fördert mit 15 Mio. Euro vier neue Zentren für Wissenschaftskommunikationsforschung* [Pressemitteilung], <https://www.volkswagenstiftung.de/aktuelles-presse/presse/volkswagenstiftung-f%C3%B6rdert-mit-15-mio-euro-vier-neue-zentren-f%C3%BCr-wissenschaftskommunikationsforschung>
- von Kotzebue, L., & Nerdel, C. (2012). Professionswissen von Biologielehrkräften zum Umgang mit Diagrammen. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 181–200.
- Webb-Williams, J. (2018). Science self-efficacy in the primary classroom: Using mixed methods to investigate sources of self-efficacy. *Research in Science Education*, 48(5), 939–961. <https://doi.org/10.1007/s11165-016-9592-0>
- Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology*, 4(433), 1–22. <https://doi.org/10.3389/fpsyg.2013.00433>
- Williams, T., & Williams, K. (2010). Self-efficacy and performance in mathematics: Reciprocal determinism in 33 nations. *Journal of Educational Psychology*, 102(2), 453–466. <https://doi.org/10.1037/a0017271>
- Winn, B. (1987). Charts, graphs, and diagrams in educational materials. In D. M. Willows & H. A. Houghton (Hrsg.), *The psychology of illustration* (S. 152–198). Springer.
- Yarden, A., Norris, S. P., & Phillips, L. M. (2015). *Adapted primary literature: The use of authentic scientific texts in secondary schools* (Bd. 22). Springer. <https://doi.org/10.1007/978-94-017-9759-7>
- Yore, L. D., & Hand, B. (2010). Epilogue: Plotting a research agenda for multiple representations, multiple modality, and multimodal representational competency. *Research in Science Education*, 40(1), 93–101. <https://doi.org/10.1007/s11165-009-9160-y>
- Yore, L. D., Pimm, D., & Tuan, H.-L. (2007). The literacy component of mathematical and scientific literacy. *International Journal of Science and Mathematics Education*, 5(4), 559–589. <https://doi.org/10.1007/s10763-007-9089-4>
- Ziepprecht, K. (2016). *Strategien und Kompetenzen von Lernenden beim Erschließen von biologischen Informationen aus unterschiedlichen Repräsentationen*. Logos.
- Ziepprecht, K., Jäger, D. S., & Schwanewedel, J. (2017). Charakteristika von Repräsentationen im Biologieunterricht aus Sicht von Schülerinnen und Schülern der Sekundarstufe II. *Zeitschrift für Didaktik der Biologie (ZDB) - Biologie Lehren und Lernen*, 21(1), 23–41. <https://doi.org/10.4119/zdb-1632>

Literatur

Zusho, A., Pintrich, P. R., & Coppola, B. (2003). Skill and will: The role of motivation and cognition in the learning of college chemistry. *International Journal of Science Education*, 25(9), 1081–1094. <https://doi.org/10.1080/0950069032000052207>

ABBILDUNGSVERZEICHNIS

Abbildung 2.2.1	Klassifikation von Repräsentationen basierend auf verschiedenen Zeichensystemen (nach Schnotz 1994, 2001).	12
Abbildung 2.3.1	Informationsfluss durch das menschliche Gedächtnis. Vereinfachte Darstellung basierend auf dem von Atkinson und Shiffrin (1968) vorgeschlagenen Mehrspeichermodell.	15
Abbildung 2.3.2	<i>Cognitive Theory of Multimedia Learning</i> verändert nach Mayer (2014, S. 52).	17
Abbildung 2.3.3	Grundlegende Aspekte des Lernens mit MER, die das <i>DeFT (Design, Functions, Tasks) Framework</i> von Ainsworth (2006) umfasst. Die pädagogischen Funktionen wurden bereits in Ainsworth (1999) vorgestellt.	25
Figure 4.3.1	Exemplary test material closely resembling parts of the actual material for the no-redundancy group in Task A. Textual information that was required to solve the task is highlighted in red, relevant depictive information is highlighted in blue. No highlights were included in the material of the actual task provided to the students.	71
Figure 4.3.2	Exemplary test material closely resembling parts of the actual material for the redundancy group in Task A. Textual information that was required to solve the task is highlighted in red, relevant (originally) depictive information is highlighted in blue. No highlights were included in the material of the actual task provided to the students.	72
Figure 4.3.3	Exemplary test material closely resembling parts of the actual material for the text-only group in Task A. Textual information that was required to solve the task is highlighted in red, relevant originally depictive information is highlighted in blue. No highlights were included in the material of the actual task provided to the students.	73
Figure 4.4.1	Mean subtest scores of different information units of Task A and Task B by group collapsed across classroom. The mean subtest scores are presented for relevant information units (originally) represented by figures or texts of Task A and for information units (originally) represented by figures and tables or by texts of Task B. Fig = Figure; Tab = Table. Error bars represent standard errors. All highlighted significance levels for pairwise comparisons were adjusted using Bonferroni correction. ** $p < .01$. *** $p < .001$.	79

Figure 4.9.1	Boxplots of test performance for the entire Task A separated by group. Boxes show the middle 50% of scores, and whiskers show the top and bottom 25%. The bold horizontal line in each box depicts the median.	95
Figure 4.9.2	Boxplots of test performance examining the (originally) descriptive information units of Task A separated by group. Boxes show the middle 50% of scores, and whiskers show the top and bottom 25%. The bold horizontal line in each box depicts the median.	95
Figure 4.9.3	Boxplots of test performance examining the (originally) descriptive information units of Task A separated by group. Boxes show the middle 50% of scores, whiskers show the top and bottom 25%, and the individual dot represents outliers (scores of four students). The bold horizontal line in each box depicts the median.	96
Figure 4.9.4	Boxplots of test performance for the entire Task B separated by group. Boxes show the middle 50% of scores, whiskers show the top and bottom 25%, and the individual dot represents outliers (scores of two students). The bold horizontal line in each box depicts the median.	96
Figure 4.9.5	Boxplots of test performance examining the (originally) descriptive information units of Task B separated by group. Boxes show the middle 50% of scores, whiskers show the top and bottom 25%, and the individual dots represent outliers (each dot represents scores of two students). The bold horizontal line in each box depicts the median.	97
Figure 4.9.6	Boxplots of test performance examining the (originally) descriptive information units of Task B separated by group. Boxes show the middle 50% of scores, whiskers show the top and bottom 25%, and the individual dots represent outliers (each dot at a score = 1 represents results of two students, the dot at score = 0 the result of one student). The bold horizontal line in each box depicts the median.	97
Abbildung 5.3.1	Testleistung in Aufgabe A als Funktion von Selbstwirksamkeitserwartung und Experimentalgruppe (simple slopes).	114

TABELLENVERZEICHNIS

Tabelle 2.2.1	Charakteristika und Beispiele deskriptionaler und depiktionaler Repräsentationen (basierend auf Schnotz, 2001)	13
Tabelle 3.1	Übersicht über die durchgeführten Studien des Promotionsprojekts	60
Table 4.4.1	Descriptive statistics and ANOVA results for the covariates	77
Table 4.9.1	Excerpt of the category scheme for the demonstration example (cf. Figures 4.3.1–4.3.3)	98
Table 4.9.2	Pairwise comparisons between experimental groups for Task A and B	99
Table 4.9.3	Descriptive data and pairwise comparisons between experimental groups for Task A and B with cognitive abilities as covariate	100
Tabelle 5.3.1	Mittelwerte und Standardabweichungen nach Experimentalgruppe	112
Tabelle 5.3.2	Regressionsmodelle 2 und 3 zur Vorhersage der Testleistung von Aufgabe A	117

ABKÜRZUNGSVERZEICHNIS²⁶

APL	<i>Adapted Primary Literature</i>
BITE	Entwicklung und Überprüfung von Kompetenzmodellen zur integrativen Verarbeitung von Texten und Bildern
CLT	<i>Cognitive Load Theory</i>
COVID-19	<i>Coronavirus Disease 2019</i>
CTML	<i>Cognitive Theory of Multimedia Learning</i>
DCT	<i>Dual Coding Theory</i>
DeFT	<i>Design, Functions, Tasks</i>
DNA	<i>deoxyribonucleic acid</i>
ER	externe Repräsentation; engl. <i>external representation</i>
FiRe ₂	Feedback im Referendariat zu externen Repräsentationen
ITPC	<i>Integrated Model of Text and Picture Comprehension</i>
KeiLa	Kompetenzentwicklung in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen
KiL	Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen
KMK	Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland
MER	multiple externe Repräsentation; engl. <i>multiple external representation</i>
mRNA	<i>messenger ribonucleic acid</i>

²⁶ In diesem Abkürzungsverzeichnis werden die im übergreifenden Rahmen dieser Dissertation genutzten Abkürzungen aufgelistet, die von gängigen in der deutschen Sprache verwendeten Abkürzungen abweichen.

NGSS	<i>Next Generation Science Standards</i>
PISA	<i>Programme for International Student Assessment</i>
SE	<i>self-efficacy belief</i>
SWE	Selbstwirksamkeitserwartung

DANKSAGUNG

Während meiner Promotionszeit habe ich auf vielfältige Weise Unterstützung und die Möglichkeit zum Austausch mit vielen Menschen erhalten und so möchte ich mich im Folgenden bei all denen bedanken, die diese Jahre begleitet haben.

Ein großer Dank gilt meiner Doktormutter Prof. Dr. Julia Schwanewedel. Angefangen von der Möglichkeit, mit fachbiologischem Studium dieses Promotionsprojekt in der Biologiedidaktik bei ihr beginnen zu dürfen, danke ich für die Förderung meines Wissens über Biologiedidaktik und biologiedidaktische Forschung, die motivierende und konstruktive Unterstützung meiner Arbeit, spannende fachliche Diskussionen, den Raum, den sie mir für meine Weiterentwicklung gegeben hat, und das in meine Arbeit gesetzte Vertrauen.

Ein besonderer Dank gilt ebenfalls meiner Mentorin Dr. Kerstin Schütte, die diese Arbeit mit unermüdlichem Einsatz begleitet und mit hoher fachlicher Expertise bereichert hat. Ich danke für ihre große Bereitschaft, mit mir die kleinen und großen Fragen zu diskutieren, und dafür, dass sie mir stets das Gefühl gab, dass ich mich an sie wenden kann, wenn ich mit einer Sache nicht weiterkam.

Meinen beiden Betreuerinnen möchte ich sagen: Mit Freude und Wehmut denke ich an unsere langen und doch kurzweiligen Treffen zurück, die immer konstruktiv und wertschätzend waren, in denen ich unglaublich viel gelernt habe und die mich angeregt haben, weiterzudenken. Neben allem anderen danke ich euch dafür, dass ihr mich bestärkt habt in dem, was ich tat. Danke für den großen Rückhalt, den ihr mir gegeben habt!

Prof. Dr. Ute Harms danke ich für die freundliche Aufnahme in ihre Abteilung und das Ermöglichen der Vertragsverlängerung auch nach dem Standortwechsel meiner Doktormutter.

Mein großer Dank gilt auch allen Abteilungsmitgliedern aus der Bio-Didaktik und „Assoziierten“ aus anderen Abteilungen, mit denen ich die Jahre am IPN und darüber hinaus verbracht habe: Danke für offene Ohren und große Hilfsbereitschaft, für fachliche Diskussionen und Zerstreuung zum jeweils richtigen Zeitpunkt, für Kuchen und erholsame Mittagspausen, für Zimmerpflanzen und Wasserflaschen, für all die guten Zeiten auf Konferenzen und im Büro, für das Herausholen aus Coronatiefen zu schönen Abstandsmittagspausen, für das Auge für den Mitmenschen und natürlich für jede einzelne gemeinsame Freizeitaktivität, die aufzuzählen den Rahmen sprengen würde. Ich bin zutiefst dankbar für

jede Freundschaft, die in den Jahren entstanden ist. Darüber hinaus danke ich allen Mitarbeiter:innen am IPN, die zu einem so freundlichen Miteinander auf den Fluren auch über Abteilungsgrenzen hinweg beigetragen haben.

Ich möchte mich bei meinen Kolleg:innen an der HU Berlin für die freundliche Aufnahme in die Abteilung und schöne gemeinsame Videokonferenzen bedanken. Ein großes Dankeschön geht zudem an meine Kolleginnen an der Universität Hamburg für ihre Hilfsbereitschaft und das Verständnis und die Unterstützung besonders in den letzten Monaten sowie für die gemeinsamen digitalen Treffen, auf die ich mich jedes Mal freue.

Katharina Klann danke ich für die sorgfältige Anfertigung ihrer Masterarbeit, die spannende qualitative Einblicke geliefert hat. Allen beteiligten studentischen Hilfskräften danke ich für ihre gewissenhafte Arbeit. Darüber hinaus gilt mein Dank allen Dozierenden, die freundlicherweise die Datenerhebung in ihren Kursen ermöglicht haben, sowie allen Studienteilnehmer:innen für ihre Bereitschaft zur Mitwirkung.

Auch meiner Familie und meinen weiteren Freund:innen danke ich sehr für ihre große Unterstützung, für das Nachfragen, Zuhören, aber auch Ablenken und das Mitfiebern und Mitfreuen.

ERKLÄRUNG

Hiermit versichere ich an Eides statt, dass diese Dissertation – abgesehen von der Beratung durch meine Betreuerinnen – nach Inhalt und Form meine eigene Arbeit ist und von mir nur mit den angegebenen Hilfsmitteln verfasst wurde. Ein Teil dieser Arbeit (Kapitel 4 und 5, siehe Angaben in Tabelle 3.1 bzw. zu Beginn von Kapitel 4 und 5) wurde bereits in wissenschaftlichen Zeitschriften publiziert. Diese Dissertation hat aber weder ganz noch zum Teil einer anderen Stelle im Rahmen eines Promotionsverfahrens vorgelegen. Diese Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden. Mir wurde kein akademischer Grad entzogen.

Kiel, _____

