



Concept Maps im Biologieunterricht: Kognitive und emotionale Wirkung von Trainings- maßnahmen und visuellen Gestaltungsmerkmalen

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Sina Joana Lenski, M. Sc.

BONN

2022

Berichterstatter (Gutachter): Prof. Dr. Jörg Großschedl
Prof. Dr. Werner Rieß

„Concepts are the glue that holds the world together“

Gregory Murphy (Kognitionslinguist)

Zusammenfassung

Die Komplexität und die große Vielfalt von Systemelementen der im Biologieunterricht behandelten Lerninhalte, kann für Lernende eine große Herausforderung darstellen. Um sie dabei zu unterstützen, Zusammenhänge zu identifizieren und zu verstehen, werden graphische Lernmethoden zur Organisation und Darstellung von Wissen für den Biologieunterricht empfohlen.

Concept Maps bilden Informationen über Begriffsnetze ab und stellen eine solche Lernmethode dar. Trotz ihrer einfachen syntaktischen Struktur können *Concept Maps* Lernende, die mit der Methode nicht vertraut sind jedoch überfordern und somit das Lernen behindern. Auf dem Forschungsgebiet besteht allerdings Unklarheit darüber, ob und in welchem Umfang das Erstellen (die Konstruktion) und Betrachten (die Rezeption) von *Concept Maps* trainiert werden muss. Neben einem Methodentraining könnte auch eine Anpassung der Gestaltung von *Concept Maps* das Lernen mit diesen unterstützen. Der Multimedia-Forschung zufolge könnte die Lernwirksamkeit auch durch den Einbezug von Bildelementen verbessert werden. Ein neuer Forschungszweig, nimmt an, dass spezifische Designmanipulationen, etwa menschliche Merkmale und die Verwendung bestimmter Farben und Formen (*Emotional Design*), den affektiven Zustand der Lernenden beeinflussen und das Lernen fördern.

Mit dem Ziel, lernförderliche Bedingungen für die Anwendung von *Concept Maps* im Biologieunterricht zu identifizieren wurden in der vorliegenden Dissertation zwei Projekte durchgeführt, die sich zum einen der Relevanz von *Concept Map*-Trainingsmaßnahmen und zum anderen der Wirkung von *Emotional Design*-Abbildungen widmeten.

Im Rahmen von **Projekt 1** wurden zwei *Concept Map*-Trainingsmaßnahmen mit unterschiedlichen Schwerpunkten (Rezeption oder Konstruktion) entwickelt und im Rahmen einer quasi-experimentellen Studie im 3x2-faktoriellen Design ($N = 167$) untersucht. Hierbei wurde ihre kognitive, metakognitive und emotionale Wirkung in einer den Trainingsmaßnahmen nachgelagerten Lernphase, in der entweder durch die Rezeption oder die Konstruktion gelernt wurde, mit der einer Kurzeinführung zu *Concept Maps* verglichen. Zudem wurden die Blickbewegungscharakteristika ($N = 52$) von trainierten und untrainierten Lernenden mittels *Eye-Tracking*-Messungen untersucht. Die Ergebnisse zeigen, dass ein Konstruktionstraining die Lernleistung und die Fähigkeit *Concept Maps* zu

erstellen erhöhen kann und sich die in der Konstruktion gewonnene Expertise auf das Rezipieren von *Concept Maps* übertragen lässt. Der durch das Konstruktionsstraining gewonnene Expertiseunterschied wurde auch im Rahmen von Blickbewegungsuntersuchungen beobachtet und trat unabhängig von der Struktur der *Concept Maps*, also sowohl bei hierarchisch als auch bei netzartig strukturierten *Concept Maps*, auf. Außerdem schätzten Lernende ihre eigenen Leistungen nach der Teilnahme an einem Konstruktionsstraining realistischer ein, als nach der Teilnahme an einem Rezeptionstraining. Keine der *Concept Map*-Trainingsmaßnahmen führte hingegen dazu, dass Lernende mehr Freude bei der Verwendung von *Concept Maps* empfanden.

In **Projekt 2** wurde die Wirkung von *Emotional Design-Concept Maps*, mit *Concept Maps* in die neutrale Abbildungen integriert wurden und mit rein textbasierten *Concept Maps* verglichen. In einer experimentellen Studie ($N = 249$) wurde der Einfluss dieser unterschiedlich gestalteten *Concept Maps* auf kognitive und affektive Variablen untersucht. Eine lernförderliche Wirkung wurde nicht gefunden. Es konnte jedoch gezeigt werden, dass *Emotional Design*-Abbildungen in *Concept Maps* zu einer Abnahme der wahrgenommenen Aufgabenschwierigkeit und einer Reduktion des negativen Affekts führen. Die in den *Concept Maps* verwendeten *Emotional Design*-Abbildungen wurden zudem als angenehmer empfunden (höhere Valenz). Psychophysiologische Messungen ($N = 42$) ergaben, dass Lernende durch *Emotional Design*-Abbildungen nicht stärker aktiviert wurden oder diese mehr Aufmerksamkeit generierten als die anderen *Concept Map*-Designs. Die gewonnenen Ergebnisse tragen nicht nur zu einem wissenschaftlichen Erkenntnisgewinn bei, sie ermöglichen auch die Ableitung schulpraktischer Implikationen in Bezug auf die Nutzung von *Concept Maps* im Biologieunterricht.

Abstract

The complexity and wide variety of system elements covered in biology classes, can be very challenging for learners. To help them identify and understand relationships, graphical learning methods for organizing and representing knowledge are recommended for biology classes.

Concept maps represent such a learning method. However, despite their simple syntactic structure, *concept maps* can overwhelm learners who are unfamiliar with using this method and thus hinder learning. In the research field, it is unclear so far whether and to what extent the construction and the study of *concept maps* needs to be trained.

Besides a method-training, another possibility to support learners in the use of *concept maps*, is to adapt the concept map design. According to multimedia research, learning effectiveness could be improved by including visual elements. A new branch of research assumes that specific design manipulations, such as human features and the use of certain colors and shapes (*emotional design*), influence the affective state of learners and promote learning.

This thesis aimed at identifying conditions which support learning with *concept maps* in biology classes. Therefore, two projects were conducted. One investigated the relevance of concept map training while the other examined the effect of *emotional design* pictures.

In **Project 1**, two concept map trainings focusing either the study or the construction were developed and examined in a quasi-experimental study using a 3x2 factorial design ($N = 167$). The effects of these trainings on cognition, metacognition, and affect were compared with those of a brief introduction to *concept maps*. In addition, eye movement characteristics ($N = 52$) of trained and untrained learners were examined using *eye-tracking* measures. The results showed that a construction-training can increase learning performance and the ability to create *concept maps*. Moreover, this expertise can be transferred to the study of *concept maps*. The difference in expertise gained through a construction training was also observed in eye movement studies and occurred regardless of the structure of *concept maps*, i.e., for both hierarchical and network structured *concept maps*. In addition, learners rated their own performance more accurately after participating in a construction training compared to learners who participated in a study-training.

In contrast, none of the concept map-trainings increase the enjoyment of using *concept maps*.

Project 2 compared the effects of *emotional design-concept maps* with *concept maps* that incorporated neutral pictures, and with text-only *concept maps*. In an experimental study ($N = 249$), the influence of these differently designed *concept maps* on cognitive and affective variables was investigated. A learning-promoting effect was not found. However, it was shown that *emotional design* pictures in *concept maps* lead to a decrease in perceived task difficulty and a reduction in negative affect. The *emotional design* pictures used in the *concept maps* were also perceived as more pleasant (higher valence). Psychophysiological measurements ($N = 42$) revealed that learners were not more activated by *emotional design* pictures or that they generated more attention than the other concept map designs.

The results obtained not only contribute to scientific knowledge, but also allow giving recommendations for implications of *concept maps* in biology classes.

Inhaltsverzeichnis

Inhaltsverzeichnis	VIII
Einleitung	1
Theoretischer Rahmen	4
1 Lernen und Gedächtnis.....	4
1.1 Die Struktur des Gedächtnisses	4
1.2 Die begrenzte Kapazität des Arbeitsgedächtnisses und dessen Auslastung	8
2 Lernen mit Concept Maps	13
2.1 Concept Maps	13
2.2 Die Lernwirksamkeit von Concept Maps.....	17
2.3 Die empirische Befundlage zur Lernwirksamkeit von Concept Maps	20
2.4 Der Einfluss der Vertrautheit mit Concept Maps auf die Lernwirksamkeit	22
3 Lernen mit Multimedia	26
3.1 Die Definition von Multimedia für den Lernkontext	26
3.2 Die Grundannahmen der Cognitive Theory of Multimedia Learning	26
3.3 Cognitive Theory of Multimedia Learning	28
3.4 Cognitive-Affective Theory of Learning with Media	30
4 Emotionen	32
4.1 Gefühl, Affekt, Emotion und Stimmung - Ein Abgrenzungsversuch.....	32
4.2 Die Klassifizierung/Taxonomie von Emotionen	35
4.3 Emotionen, kognitive Prozesse und Lernen	37
4.4 Die empirische Befundlage zu Emotionen und Lernerfolg	39
5 Emotional Design.....	41
5.1 Emotional Design, kognitive Prozesse und Lernen.....	43
5.2 Die empirische Befundlage zum Emotional Design	44

6	Problemstellung	48
	6.1 Die Relevanz von Concept Map-Trainingsmaßnahmen.....	48
	6.2 Die Wirkung von Emotional Design-Abbildungen.....	51
7	Projektübersicht	53
	Eigene empirische Studien.....	54
8	Die Relevanz von Concept Map-Trainingsmaßnahmen	54
	8.1 Comparing construction and study of concept maps – An intervention study on cognitive, metacognitive and emotional effects of training & learning (Publikation I).....	54
	8.1.1 Zusammenfassung, Fragestellungen und methodischer Zugang	54
	8.1.2 Eigenanteile.....	55
	8.1.3 Publikation I: Originalmanuskript.....	56
	8.2 Biologie lernen mit Concept Maps: Lässt sich die Expertise im Umgang mit Concept Maps von den Augen ablesen? (Publikation II).....	85
	8.2.1 Zusammenfassung, Fragestellungen und methodischer Zugang	85
	8.2.2 Eigenanteile.....	85
	8.2.3 Publikation II: Originalmanuskript.....	87
9	Die Wirkung von Emotional Design-Abbildungen.....	106
	9.1 Emotional Design in concept maps – No support but also no burdens (Publikation III).....	106
	9.1.1 Zusammenfassung, Fragestellungen und methodischer Zugang	106
	9.1.2 Eigenanteile.....	107
	9.1.3 Publikation III: Originalmanuskript	108
	9.2 Emotional Design pictures – Pleasant but too weak to evoke arousal and attract attention? (Publikation IV).....	145
	9.2.1 Zusammenfassung, Fragestellungen und methodischer Zugang	145
	9.2.2 Eigenanteile.....	146
	9.2.3 Publikation IV: Originalmanuskript.....	147
	Gesamtdiskussion und Ausblick	180

10	Gesamtdiskussion	180
10.1	Die Relevanz von Concept Map-Trainingsmaßnahmen (Projekt 1)	181
10.2	Zusammenfassung und Ausblick in Bezug auf Concept Map-Trainingsmaßnahmen	186
10.3	Praktische Implikationen in Bezug auf Concept Map-Trainingsmaßnahmen	188
10.4	Die Wirkung von Emotional Design-Abbildungen (Projekt 2)	190
10.5	Zusammenfassung und Ausblick in Bezug auf Emotional Design-Abbildungen	192
10.6	Praktische Implikationen in Bezug auf Emotional Design-Abbildungen	195
11	Fazit	195
	Literaturverzeichnis	197

Einleitung

In unserer heutigen schnelllebigen Welt werden wir vielen verschiedenen Informationen und Eindrücken ausgesetzt, die wir aufnehmen und miteinander in Verbindung setzen müssen, um Zusammenhänge zu verstehen. Die eingehenden Informationen müssen auf die zentralen Aspekte reduziert und auf andere Kontexte übertragen werden. Erst wenn die Zusammenhänge durchdrungen wurden, ist es möglich qualifizierte Bewertungen vorzunehmen. Dies gilt aber nicht nur für den Alltag, insbesondere im Lernkontext ist es von großer Bedeutung Zusammenhangswissen zu erlangen und kommunizieren zu können (Mandl & Fischer, 2000). Im Gegensatz zu isoliertem Faktenwissen, bietet vernetztes Wissen Anknüpfungspunkte für die Aufnahme von neuen fachlichen und überfachlichen Informationen in den eigenen Wissensbestand (Stracke, 2004). Im schulischen Kontext helfen vernetzte Wissensbestände, die beispielsweise biologische, chemische, physikalische und mathematische Inhalte umfassen, dabei naturwissenschaftliche Phänomene zu verstehen.

Die Erlangung von Zusammenhangswissen wird insbesondere im Biologieunterricht, welcher sich mit Lebewesen und Lebensvorgängen befasst, gefordert (Killermann, 1995, 2000). Biologische Sachverhalte bestehen häufig aus Einzelelementen, die miteinander in Verbindung stehen und größere Systeme bilden, welche wiederum mit anderen Systemen interagieren (Sommer & Lücken, 2010). Ein Beispiel hierfür stellt das Ökosystem See dar. Um etwa zu verstehen, warum das Füttern von Enten in einem See tödlich für die darin lebenden Fische sein kann, muss zunächst verstanden werden, dass dieses Ökosystem unterschiedliche Lebewesen umfasst. Zu diesen gehören neben Fischen auch Pflanzen und Bakterien. Durch Futterreste gelangen große Nährstoffmengen in das Gewässer, was zu einem rasanten Anstieg des Algenwachstums führen kann. Wenn diese Algen absterben, führt der damit verbundene Abbauprozess zu einer Absenkung der Sauerstoffsättigung, was unter anderem das Sterben von Fischen bewirken kann.

Schüler:innen müssen somit nicht nur biologisches Fachwissen über Einzelelemente erwerben, sie müssen diese Einzelelemente zudem miteinander in Beziehung setzen. Auch ihre Rolle im System und die Wechselwirkungen mit anderen Systemen müssen für ein umfängliches Verständnis nachvollzogen werden. Sowohl die *PISA*-Studien (OECD,

2016, 2019) als auch die *TIMSS*-Studie (Wendt et al., 2016) zeigen, dass Schüler:innen in naturwissenschaftlichen Schulfächern über ein gutes Faktenwissen verfügen. Ihnen mangelt es jedoch an der Kompetenz, naturwissenschaftliche Zusammenhänge zu verstehen und zu begründen (ebd.). Die Bildungs- und Lernforschung beschäftigt sich daher intensiv mit der Frage, wie lernförderliche Lernarrangements gestaltet sein müssen.

Ein zentraler Forschungsgegenstand zur Schaffung eines lernförderlichen Lernarrangements ist die Entwicklung effizienter Lernmethoden, welche Schüler:innen beim Generieren von Zusammenhangswissen unterstützen. Basierend auf kognitivistischen Lerntheorien wurden in den 1970er Jahren sogenannte *Concept Maps* entwickelt (Novak, 1990). Hierbei handelt es sich um eine Lernmethode zur graphischen Darstellung komplexer Zusammenhänge (Mandl & Fischer, 2000). Bei dieser Methode werden Inhalte auf ihre Kernaspekte reduziert und der Fokus auf die Verbindung dieser Aspekte gelegt. Sie dient im Wesentlichen dazu, Lernende in der Strukturierung, Organisation und Vernetzung ihrer Wissensbestände zu unterstützen (ebd.).

Die Lernwirksamkeit dieser Methode wurde bereits intensiv erforscht (Schroeder et al., 2017). Trotz dieser Bemühungen existieren noch Unklarheiten und Unstimmigkeiten bezüglich der Lernwirksamkeit von *Concept Maps*. Es wird beispielsweise angenommen, dass Lernende, die in der Nutzung von *Concept Maps* unerfahren sind, viele kognitive Ressourcen für die Umsetzung der Methode gebunden werden, was wiederum die für den Lerninhalt zur Verfügung stehenden Ressourcen reduziert. Wieviel Übung in der Verwendung dieser Methode notwendig ist, ist Gegenstand kontroverser Diskussionen (Karpicke & Blunt, 2011; Mintzes et al., 2011; Schroeder et al., 2017).

Nach der *Cognitive Load Theory*, welche sich mit der kognitiven Belastung bei Lernprozessen befasst (Sweller et al., 1998), kann der Lernerfolg auch durch die suboptimale Gestaltung von Lernmaterialien reduziert werden (ebd.). Eine lernförderliche Gestaltung kann laut der Forschung zum Lernen mit Multimedia durch die kombinierte Darstellung von Bild- und Textelementen erzielt werden (Mayer, 2002; Mayer & Chandler, 2001; Mayer & Estrella, 2014). Zur Verbesserung des Lernerfolgs wird in der aktuellen Forschung auf diesem Gebiet im Rahmen der Entwicklung des sogenannten *Emotional Design* der Fokus auf die Nutzung affektiver und motivationaler Eigenschaften von Lernmaterial gelegt (Mayer & Estrella, 2014; Moreno, 2006; Um et al., 2012).

Ausgehend von den skizzierten Befunden wurden im Rahmen dieser Dissertation mehrere Studien durchgeführt, welche sich der Vertrautheit in der Verwendung von *Concept Maps* sowie unterschiedlichen Arten der visuellen Gestaltung von *Concept Maps* im Biologieunterricht widmeten. Hierbei wurden substantielle Erkenntnisse in Bezug auf den lernförderlichen Einsatz von *Concept Maps* gewonnen, welche im Verlauf dieser Dissertation vorgestellt werden.

Im Folgenden werden zunächst die theoretischen Grundlagen der Studien betrachtet, die dieser Dissertation zugrunde liegen und zentrale Forschungsbefunde referiert (**Kapitel 1 bis 5**). Basierend auf einer überblicksartigen Zusammenfassung werden anschließend die Ziele dieser Arbeit vorgestellt (**Kapitel 6**) und ein Gesamtüberblick über die durchgeführten Studien gegeben (**Kapitel 7**). In den darauffolgenden Kapiteln werden Fragestellungen und forschungstechnische Zugänge zu den einzelnen Studien präsentiert und die Publikationen im Original zur Verfügung gestellt (**Kapitel 8 und 9**). Im Rahmen einer Gesamtdiskussion werden die Befunde studienübergreifend interpretiert und diskutiert (**Kapitel 10**). Zudem werden aus den gewonnenen Erkenntnissen Implikationen für die Schulpraxis abgeleitet und Impulse für künftige Forschung gesetzt. Die Arbeit schließt mit einem Fazit (**Kapitel 11**).

Theoretischer Rahmen

1 Lernen und Gedächtnis

1.1 Die Struktur des Gedächtnisses

Nach Brand und Markowitsch (2006) wird unter dem Begriff des Lernens die Aufnahme und langfristige Speicherung von Informationen verstanden. Das Gedächtnis und dessen Struktur spielen bei Aufnahme, Verarbeitung und Speicherung von Lerninhalten eine entscheidende Rolle.

Eins der bekanntesten Modelle, das die Funktionsweise des Gedächtnisses versucht zu erklären, ist das *Multi-store Model* (Atkinson & Shiffrin, 1968). Es umfasst drei Komponenten: das sensorische Register (Ultrakurzzeitgedächtnis), das Kurzzeitgedächtnis und das Langzeitgedächtnis (siehe **Abbildung 1**). Der Begriff *Kurzzeitgedächtnis* wurde in der neueren Literatur durch *Arbeitsgedächtnis* abgelöst, um die aktive Rolle dieser Gedächtnisform bei der Informationsverarbeitung hervorzuheben¹. Nach diesem Modell gelangen alle Informationen zunächst in Form von (Umwelt-)reizen in das sensorische Register. Grundvoraussetzung für die Aufnahme von Informationen ist, dass sie die für den jeweiligen Sinn spezifische Wahrnehmungsschwelle überschreiten. Das sensorische Register stellt somit eine Art Puffer dar, bei dem angenommen wird, dass Informationen zwar nur eine sehr kurze Verweildauer haben, das Register an sich jedoch über eine hohe Kapazität verfügt. Bewusst wahrgenommene Reize werden weiterverarbeitet und gelangen in das Arbeitsgedächtnis. Alle anderen Reize werden herausgefiltert bzw. vergessen. Das Arbeitsgedächtnis wird mit einer Art mentalem Arbeitsplatz des Denkens verglichen (Miller & Cohen, 2001) und Informationen haben hier eine höhere Verweildauer als im sensorischen Register. Wenn es Lernenden gelingt die neuen Informationseinheiten im Arbeitsgedächtnis zu halten oder mit bereits vorhandenem Vorwissen zu verknüpfen (*Elaboration*), besteht die Möglichkeit diese Inhalte ins Langzeitgedächtnis zu überführen (Atkinson & Shiffrin, 1968). Wenn diese Überführung jedoch misslingt, werden die

¹ Ein umfassender Überblick zu der Unterscheidung zwischen Kurzzeit- und Arbeitsgedächtnis und die Verwendung beider Begrifflichkeiten findet sich beispielsweise in Aben et al. (2013).

Inhalte vergessen. Inhalte, bei denen dieser Überförungsprozess hingegen gelingt, können anschließend aus dem Langzeitgedächtnis abgerufen werden, ohne dass eine zusätzliche Belastung für das Arbeitsgedächtnis entsteht. Dieser Umstand erklärt, warum ein hohes Vorwissen so prädiktiv und zentral für den Lernerfolg ist.



Abbildung 1. Das *Multi-store Model* nach Atkinson und Shiffrin (1968), angelehnt an eine Darstellung von Kelleher und Dobnik (2019)².

Das Langzeitgedächtnis ist dynamisch und verändert sich mit der Zeit. Es scheint zumindest theoretisch keine kapazitative Beschränkung zu haben (Brady et al., 2011; Brady et al., 2008), bei fehlender Nachfrage, also dem Ausbleiben von Informationsabrufen, können Erinnerungen jedoch verblassen. Wissen wird im Langzeitgedächtnis in zwei verschiedenen Gedächtnissystemen gespeichert (Strobach, 2020). Das deklarative Gedächtnis speichert bewusstes Wissen, was durch Sprache ausgedrückt werden kann (Schirp, 2007; Strobach, 2020). Hier werden etwa Erinnerungen an den Urlaub in Italien (episodisches Wissen), die Vokabeln für die nächste Englischklausur und Faktenwissen über unterschiedliche Gedächtnissysteme gespeichert (semantisches Wissen). Im non-deklarativen (impliziten) Gedächtnis wird unbewusstes Wissen, welches nicht sprachlich, sondern durch Handlungen ausgedrückt wird, gespeichert (ebd.). Hierzu zählt unter anderem

² Für die Darstellung verschiedener Gedächtnismodelle (**Abbildung 1, 3, 5 und 6**) wurde eine ähnliche Art der Gestaltung in Bezug auf räumliche Anordnung sowie Färbung der unterschiedlichen Komponenten gewählt. Auf diese Weise sollen Ähnlichkeiten und Unterschiede verdeutlicht werden. Zudem handelt sich um mögliche Darstellungsformen, welche keinen Anspruch auf Vollständigkeit oder Überlegenheit gegenüber anderen Arten der Darstellung erheben. An dieser Stelle soll zudem betont werden, dass sich Begriffsdefinitionen trotz der Nutzung ähnlicher Begrifflichkeiten zwischen den verschiedenen Modellen unterscheiden können.

das prozedurale Wissen (Strobach, 2020). Dieses Wissen ist unbewusst und wird beispielsweise im Rahmen der klassischen *Konditionierung*³ oder durch den Prozess des *Primings*⁴ erworben (ebd.). In dieser Untereinheit befindet sich das Wissen darüber, wie ein Instrument gespielt wird oder wie man eine Schleife bindet. Wissen, das man benötigt, um ein Periodensystem zu interpretieren oder die Evolutionstheorie zu verstehen, wird konzeptuelles Wissen genannt. Dieses Wissen kann sowohl implizit als auch verbalisiert vorliegen (Myers & Wilson, 2014; Schirp, 2007). Es ist erweiter- und wandelbar (ebd.).

Für die Speicherung von episodischen Gedächtnisinhalten ist eine einmalige Erfahrung ausreichend (z. B. eine Reise nach Italien) und eine Wiederholung ähnlicher Ereignisse führt zum Verblässen einzelner Erinnerungen (Strobach, 2020). Mehrmalige Wiederholungen haben bei semantischem Wissen den gegenteiligen Effekt. Hier tritt eine Verfestigung des Wissens durch Wiederholungen ein. Bezüglich der Art der Speicherung von semantischen Gedächtnisinhalten⁵, wird davon ausgegangen, dass Begriffe in Form von gleichwertigen Kategorien (Knoten) im Gedächtnis repräsentiert werden (ebd.). Diese Knoten werden nach Collins und Quillian (1969) in einem hierarchisch strukturierten semantischen Netzwerk organisiert, wo sie durch sogenannte Kanten miteinander verbunden sind (siehe **Abbildung 2**). Den Knoten werden Eigenschaften und Merkmale zugeordnet. Es findet zur Einordnung vom Allgemeinen (siehe **Abbildung 2**, Ebene 3) hin zum Spezifischen (siehe **Abbildung 2**, Ebene 1) eine Unterteilung in verschiedene Ebenen statt. Die Eigenschaften, die einem Knoten auf oberster Ebene zugeordnet werden, gelten auch für alle nachfolgenden Ebenen und deren Knoten. Die Eigenschaften werden also „vererbt“. Beispielhaft wird im Folgenden ein semantisches Netzwerk zum Thema *Tier* betrachtet (Collins & Quillian, 1969). Das Merkmal *kann sich bewegen*, welches für den Knoten *Tier* gilt, ist somit auch auf den Begriff *Vogel* übertragbar, da dieser sich auf einer tieferen Ebene befindet. Für den *Vogel* kommt wiederum das neue Merkmal *hat*

³ Klassische Konditionierung ist definiert als die Fähigkeit von Lebewesen, die prädiktive Beziehung zwischen einem Signal und einem (als wichtig erachteten) Ereignissen zu erkennen und basierend darauf zu lernen (Ferrandiz, 2004).

⁴ Durch *Priming* werden verknüpfte Informationen zusammen abgerufen, sobald auf eine davon zugegriffen wird (Myers & Wilson, 2014). *Priming* ist somit das Resultat aus einer gemeinsamen vorherigen Aktivierung (Strobach, 2020). Umweltreize, die *Priming* auslösen können, werden auch als Abrufhilfen bezeichnet.

⁵ Informationen zur Speicherung episodischer Gedächtnisinhalte werden im Rahmen dieser Arbeit nicht thematisiert.

Flügel hinzu. Umso stärker ein Knoten aktiviert wird, desto größer ist die Wahrscheinlichkeit, dass die enthaltene Information aus dem Langzeitgedächtnis abgerufen werden kann. Ein zusätzlicher Effekt, der semantischen Netzwerken beigemessen wird, ist die Aktivierungsausbreitung (*Spreading-activation Theory*; Collins & Loftus, 1975). Laut dieser Theorie werden je nach Aktivierungsgrad eines Knotens auch umliegende Knoten, abhängig von der Ausprägung ihrer Verknüpfung, abgerufen. Auf diesem Prinzip beruht auch der Prozess des *Primings*. In der neueren Gedächtnisforschung wird die Annahme, dass semantische Netzwerke hierarchisch strukturiert sind, aufgrund empirischer Befunde zunehmend aufgeweicht und die Vermutung, dass semantische Netzwerke nicht-hierarchisch organisiert sind, nimmt zu (Strobach, 2020).

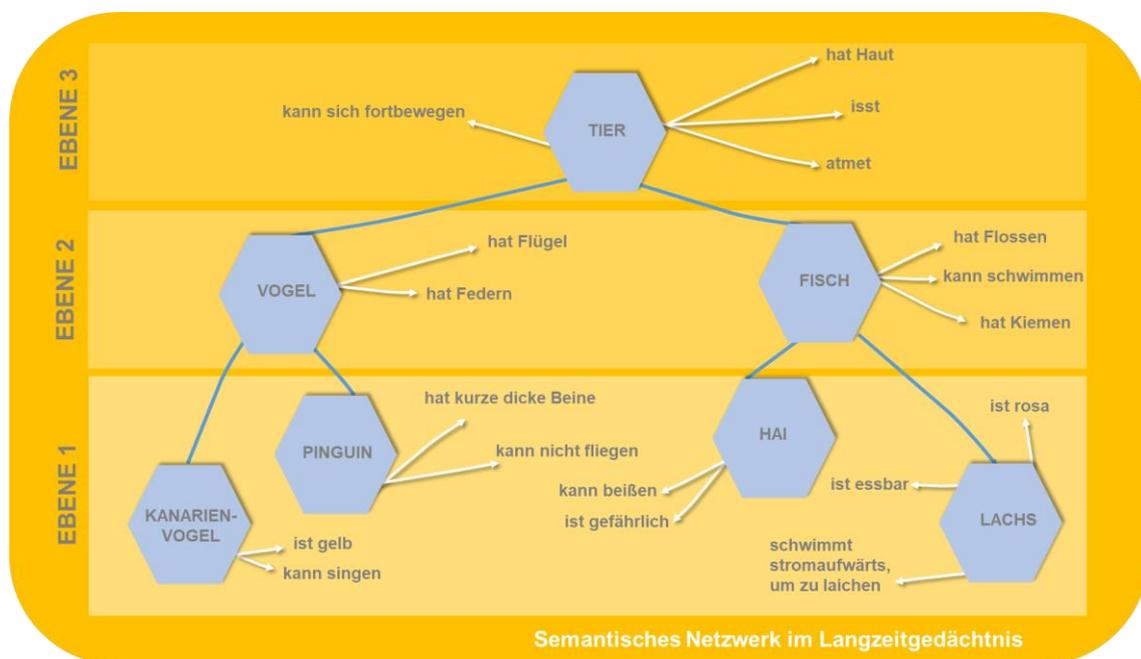


Abbildung 2. Beispielhafte Darstellung eines hierarchisch organisierten semantischen Netzwerkes, angelehnt an Collins und Quillian, 1969.

Die vorgestellten Annahmen über die verschiedenen Gedächtnissysteme ermöglichen Schlussfolgerungen in Bezug auf den Lernprozess. Da unter Lernen die Überführung von Informationen aus dem Arbeits- in das Langzeitgedächtnis verstanden wird, handelt es sich um ein lernförderliches Lernarrangement, wenn diese Überführung unterstützt wird (Tobinski, 2017). Ein weiterer wichtiger Aspekt ist die Festigung von vorhandenen Wissensstrukturen im Gedächtnis, was beispielsweise durch Elaboration, Wiederholung und

Organisation von Lerninhalten gelingt und bedeutungsvolles Lernen ermöglicht (Novak, 1990; Novak & Cañas, 2004; Novak & Cañas, 2006)⁶.

1.2 Die begrenzte Kapazität des Arbeitsgedächtnisses und dessen Auslastung

Eine zentrale Annahme für das Lernen ist, dass die Kapazität des Arbeitsgedächtnisses begrenzt ist (Baddeley, 1992)⁷. Im Rahmen des *Model of Working Memory* spezifiziert Baddeley (1992) die Eigenschaften des Arbeitsgedächtnisses⁸. Nach diesem Modell besteht es aus zwei Speichersystemen sowie dem episodischen Puffer (*Episodic Buffer*)⁹. Diese Systeme sind die phonologische Schleife (*Phonological Loop*) und der visuell-räumliche Notizblock (*Visuospatial Scratchpad*). Diese Speichersysteme sind über den episodischen Puffer und das Kontrollsystem (zentrale Exekutive) miteinander verbunden (siehe **Abbildung 3**). Es findet somit eine modalitätsspezifische Verarbeitung, also eine Unterscheidung nach der Art der Sinneswahrnehmung statt. Unterschieden wird zwischen auditiven Informationen (phonologische Schleife) und visuell-räumlichen Informationen (visuell-räumlicher Notizblock). Die phonologische Schleife gliedert sich ihrerseits in zwei Einheiten, wobei eine Einheit der kurzzeitigen Speicherung (phonologischer Speicher; ca. ein bis zwei Sekunden) und die andere der Wiederholung und Aufrechterhaltung von Informationen dient (artikulatorische Schleife). Durch die permanente Wiederholung und das stille Sprechen (subvokale Artikulierung) der Informationen wird der Zerfall der Gedächtnisspur geschützt und die Informationen werden aufrechterhalten. Der visuell-räumliche Notizblock ist zuständig für das kurzfristige Behalten und Manipulieren von visuellen und räumlichen Informationen. Auch hier werden Untereinheiten für

⁶ Nach der konstruktivistischen Lerntheorie von Ausubel (1968) nimmt das vorhandene Vorwissen eine große Rolle für das bedeutungsvolle Lernen (*Meaningful Learning*) ein. Bedeutungsvolles Lernen erfolgt dann, wenn neues Wissen in bereits bestehende Wissensstrukturen integriert wird. Deshalb wird dem Auswendiglernen (*Rote Learning*) nach dieser Theorie eine geringere Lernwirksamkeit zugesprochen (ebd.).

⁷ Es wird angenommen, dass sieben *Chunks* gleichzeitig im Arbeitsgedächtnis verarbeitet werden können (Sweller et al., 1998). Unter *Chunks* werden nach Steiner (2006) übergeordnete Informationseinheiten, bestehend aus verschiedenen Schemata verstanden.

⁸ Neben den Modellen von Baddeley (z. B. 1992, 2001) wurden noch weitere Arbeitsgedächtnismodelle entwickelt, auf die im Rahmen dieser Arbeit lediglich verwiesen werden soll. Bekannte Modelle stellen beispielsweise das *Memory Model* von Case (1985) sowie das Prozessmodell von Cowan (1998) dar.

⁹ Der episodische Puffer wurde nachträglich im Rahmen einer Umformulierung des Modells ergänzt (Baddeley, 2000, 2001). Die Ergänzung des Modells erfolgte, da empirische Forschungsergebnisse sich nicht durch das ursprüngliche Modell plausibel erklären und der Funktion der zentralen Exekutive zuordnen ließen (Zoelch et al., 2019). Die Tragfähigkeit des ergänzten Modells gilt es noch zu erforschen (Schmid et al., 2008; Zoelch et al., 2019).

Objektinformationen (z. B. Form und Farbe) und für räumliche Informationen (z. B. Rotation) angenommen (Baddeley, 1992, 2006). Der episodische Puffer stellt ein multimodales Speichersystem mit begrenzter Kapazität dar, das sowohl visuelle als auch phonologische Informationen speichern kann. In der neueren Literatur wird ihm auch eine Schnittstellenfunktion zwischen zentraler Exekutive und dem Langzeitgedächtnis zugeschrieben (Baddeley, 2001, 2012). Die zentrale Exekutive kann als übergeordnetes Kontroll- und Steuerelement verstanden werden. Hier werden die Informationen aus der phonologischen Schleife und dem visuell-räumlichen Notizblock überwacht. Außerdem findet hier die Aufmerksamkeitslenkung bzw. die Ver- und Zuteilung von Aufgaben statt.

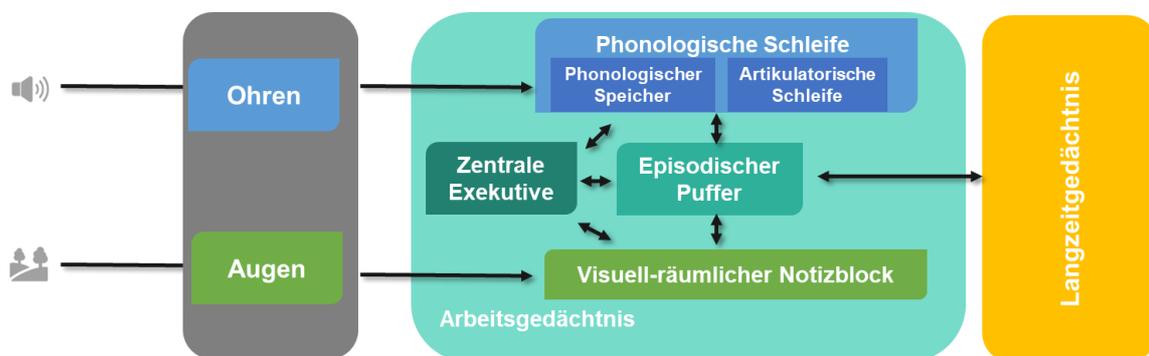


Abbildung 3. Das *Working Memory Model* von Baddeley (2012), angelehnt an McAfoose und Baune (2009).

Nach Sweller et al. (1998) wird davon ausgegangen, dass verschiedene Aspekte des Lernens zur Aus- oder sogar Überlastung des Arbeitsgedächtnisses führen können. Eine solche Überschreitung der Kapazität kann in einem sogenannten *Cognitive Overload* münden. Den unterschiedlichen Arten kognitiver Belastung widmet sich die *Cognitive Load Theory* (CTL; z. B. Sweller, 2005). Nach dieser Theorie werden drei verschiedene Arten kognitiver Belastung (*Cognitive Load*) unterschieden: intrinsische, extrinsische und die lernbezogene Belastung (*Intrinsic*, *Extraneous* und *Germane Cognitive Load*).

Die intrinsische Belastung (*Intrinsic Load*) ergibt sich aus der Schwierigkeit und Komplexität der Aufgabe sowie der Expertise der Lernenden. Die Komplexität hängt unter anderem von der Anzahl der miteinander verbundenen Elemente ab, welche gleichzeitig verarbeitet werden müssen (Elementinteraktivität; Kalyuga & Renkl, 2010). Während die Grammatik einer Fremdsprache eine hohe Elementinteraktivität aufweist, da für ihr Verständnis viele Informationen gleichzeitig im Arbeitsgedächtnis verarbeitet werden müssen, weist das Lernen von Vokabeln eine niedrige Elementinteraktivität auf (Ginns,

2005). Hierbei sollte berücksichtigt werden, dass umso größer das inhaltsbezogene Vorwissen des Lernenden ist, desto geringer ist die Elementinteraktivität und somit die intrinsische Belastung (Schnotz & Kürschner, 2007; Sweller et al., 1998).

Die Aktivierung des Vorwissens der Lernenden, oder die Vereinfachung bzw. Anpassung der Aufgabenschwierigkeit an die Fähigkeiten der Lernenden kann eine Reduzierung der intrinsischen Belastung darstellen (Klepsch & Seufert, 2020). Zur Reduzierung der intrinsischen kognitiven Belastung eignen sich nach Mayer und Moreno (2003) Trainingsmaßnahmen vor dem Lernen. Auch die Steigerung der Anforderung vom Einfachen hin zum Komplexen wird empfohlen (van Merriënboer et al., 2006). Durch die kontinuierliche Steigerung sollen sich die Ressourcen für die lernbezogene Belastung vergrößern (ebd.). Während die intrinsische Belastung nicht direkt durch die Darstellung des Lernmaterials verändert werden kann, bildet die suboptimale Gestaltung von Lernmaterialien die Hauptursache für eine hohe extrinsische Belastung (*Extraneous Load*). Grund hierfür ist, dass Lernende irrelevante Such-, Organisations- und Integrationsprozesse durchführen müssen (Kalyuga & Renkl, 2010). Beispiele für eine suboptimale Gestaltung sind rein textbasierte Lernmaterialien, eine zu überladene Darstellung, oder die Integration nicht lernbezogener Elemente (Harp & Mayer, 1997; Lenzner et al., 2013). Elemente, welche zwar das Interesse des Lernenden wecken, inhaltlich aber irrelevant sind, werden *Seductive Details* genannt (*Seductive Detail Effect*; z. B. Harp & Mayer, 1997). Diese lenken die Lernenden ab und verbrauchen unnötig kognitive Ressourcen. *Seductive Details* sind beispielsweise Textpassagen, die für das Verständnis der Aufgabe irrelevant sind (z. B. interessante Zusatzinformationen) oder Abbildungen ohne Lerninhalt (Fotografien, die das Interesse wecken sollen). Eine räumlich oder zeitlich getrennte Darstellung von denklogisch verbundenen Informationen kann sich ebenfalls lernhinderlich auswirken. Durch den Versuch der Lernenden, beide Informationen zu zeitgleich zu verarbeiten, wird die Aufmerksamkeit zwischen den beiden Informationsquellen aufgeteilt (*Split-Attention Effect*; Chandler & Sweller, 1991). Auch Redundanzen (die Darstellung derselben Information in unterschiedlichen Formen) können die extrinsische kognitive Belastung erhöhen, indem den Lernenden die Identifikation relevanter Informationen erschwert wird (*Redundancy Effect*; Sweller & Chandler, 1991; Sweller et al., 1998)¹⁰. Ein Beispiel für

¹⁰ Im Rahmen der Forschung auf dem Gebiet des Lernens mit Multimedia wurden zahlreiche Begebenheiten beschrieben, welche die extrinsische Belastung erhöhen können und Vorschläge gemacht diesem negativen Effekt entgegenzuwirken. An dieser Stelle wird nur eine Auswahl an Gestaltungsempfehlungen, welche auf der *Cognitive-Load Theory* basieren, betrachtet. Einen Überblick hierzu bietet Sweller et al. (1998).

eine solche Redundanz ist etwa ein selbsterklärendes Diagramm zum Thema Klimawandel, dem zusätzlich eine textuelle Erläuterung beiliegt. Nach dem Modalitätseffekt (*Modality Effect*; Sweller et al., 1998) wirkt sich das zeitgleiche Ansprechen verschiedener Sinne (zum Beispiel durch visuelle und auditive Informationen) positiv auf den Lernerfolg aus. Beispielsweise ist die Verwendung einer Graphik lernförderlicher, wenn diese mündlich erklärt wird (z. B. durch einen Audiokommentar), als durch eine textbasierte Legende.

Unter der lernbezogenen Belastung (*Germane Load*) versteht man keine klassische Belastung, sondern die Ressourcen, welche für den Lernprozess (etwa der Erwerb von Schemata und deren Automatisierung) zur Verfügung stehen. Dies betrifft beispielsweise die Nutzung von Lernstrategien und metakognitiven Strategien (Schnotz & Kürschner, 2007). Zwar kann der Lernprozess auch ohne die Nutzung dieser lernbezogenen Ressourcen stattfinden, durch ihre Erhöhung können die Behaltens- und Transferleistungen der Lernenden jedoch gesteigert werden (van Merriënboer et al., 2002). Da die unterschiedlichen Belastungen additiv sind (Schnotz & Kürschner, 2007), führt eine Reduzierung der extrinsischen und intrinsischen Belastung dazu, dass das Arbeitsgedächtnis mehr in lernbezogene Ressourcen investieren kann, und so der Lernprozess gefördert wird.

Basierend auf diesen Erkenntnissen lassen sich Strategien entwickeln, durch die bedeutungsvolles Lernen ermöglicht wird (Lernstrategien¹¹). Lernstrategien sind Handlungs- und Verhaltensweisen sowie Gedanken, welche Lernenden helfen, ihr Lernverhalten zu beeinflussen und zu kontrollieren (Mandl & Friedrich, 2006; Wild et al., 2006).

Hierbei wird zwischen kognitiven und metakognitiven Strategien unterschieden (Leopold & Leutner, 2002; Mandl & Friedrich, 2006)¹². Während kognitive Strategien an den Prozessen der Aufnahme und Verarbeitung von Informationen (z. B. Organisation, Strukturierung und Verknüpfung von Informationen) ansetzen (Pintrich, 1999; Weinstein & Mayer, 1986), widmen sich metakognitive Strategien übergeordneten Aspekten, wie der

¹¹ Da laut der einschlägigen Fachliteratur die Begriffe Lernstrategie, Lernmethode und Lerntechnik nicht trennscharf voneinander abzugrenzen sind und deshalb meist synonym verwendet werden (Haudeck 2008; Neuner-Anfindsen 2005), wird im Rahmen dieser Arbeit ebenso verfahren.

¹² Nach Wild und Schiefele (1994) werden drei Arten von Lernstrategien unterschieden: Kognitive, metakognitive und ressourcenbezogene Lernstrategien. Auf die letztgenannten Lernstrategien wird im Rahmen dieser Arbeit nicht eingegangen. Neben den vorgestellten Klassifikationen von Lernstrategien, existieren noch weitere Ansätze, welche beispielsweise eine Unterteilung in allgemeinen und spezifischen Lernstrategien vornehmen (Mandl & Friedrich, 2006).

Planung, der Überwachung oder der Regulation des eigenen Lernprozesses (Leopold & Leutner, 2002). Wiederholungs-, Elaborations- und Organisationsstrategien sind Beispiele für kognitive Strategien. Je nach Tiefe der Verarbeitung (*Levels of Processing*; Craik & Lockhart, 1972) kann zudem zwischen Tiefen- und Oberflächenstrategien unterschieden werden (Leopold & Leutner, 2002).

Während Tiefenstrategien das tiefere Verständnis (eines Textes) zum Ziel haben, steht bei den Oberflächenstrategien die (genaue) Wiedergabe der Informationen im Fokus (ebd.). Das wiederholte Lesen eines Lerntextes (Wiederholungsstrategie) wird beispielsweise meist den Oberflächenstrategien zugeordnet (Sumfleth et al., 2010). Elaborationsstrategien hingegen zielen darauf ab, neue Informationen in bestehende Wissensstrukturen zu integrieren und werden den Tiefenstrategien zugeordnet (Mandl & Friedrich, 2006). Ziel der Organisationsstrategien ist es Informationen auf das Wesentliche zu reduzieren und so in eine „verarbeitungsfreundliche Form“ zu bringen (Großschedl & Harms, 2011, S. 116). Dies soll eine effizientere Nutzung des in seiner Kapazität limitierten Arbeitsgedächtnisses ermöglichen. Da Organisationsstrategien das tiefere Verständnis fördern, werden sie den Tiefenstrategien zugeordnet (Leopold & Leutner, 2002).

2 Lernen mit Concept Maps

2.1 Concept Maps

Bei *Concept Maps*¹³ handelt es sich um eine Methode, mit deren Hilfe Wissensinhalte und komplexe Sachzusammenhänge graphisch in einer Art Übersichtskarte dargestellt werden können (Novak, 1990). Sie haben ihren Ursprung im Bereich der Lernpsychologie und wurden ursprünglich als Diagnoseinstrument entwickelt (Novak, 1990; Novak, 1990; Novak & Cañas, 2006). *Concept Maps* bestehen aus Begriffen (Konzepten) und deren Verbindungen (Relationen) zueinander (siehe **Abbildung 4** und **7**). Konzepte, auch Knoten genannt, beinhalten die Schlüsselbegriffe eines Themengebiets und werden graphisch durch Umrahmungen hervorgehoben. Die Konzepte werden durch beschriftete Pfeile (Relationen) miteinander verbunden (Fürstenau, 2011; Novak & Cañas, 2006; Stracke, 2004). Die Richtung der Pfeile gibt die Leserichtung der Beschriftungen vor und verdeutlicht die Art des Zusammenhangs der verbundenen Konzepte. Eine Kombination aus mindestens zwei Konzepten, welche durch eine Relation miteinander verbunden sind, bildet eine sogenannte Proposition, die kleinste Einheit einer *Concept Map* (Novak & Cañas, 2006; Novak, 1990). Beispielsweise können die Begriffe *Pflanzen* und *Fotosynthese* über die Verknüpfung *betreiben* verbunden werden und abhängig von der Pfeilrichtung kann die Proposition *Pflanzen betreiben Fotosynthese* gebildet werden. Eine *Concept Map* besteht also aus verknüpften sinnstiftenden Einheiten, welche theoretisch für sich alleine stehen können, aber auch untereinander durch Querverweise in Beziehung gesetzt werden können. Somit besteht zumindest eine indirekte Verbindung zwischen allen Konzepten einer *Concept Map* (Graf, 2014). Die Anzahl an Propositionen einer *Concept Map* ist unbegrenzt, laut Novak und Cañas (2008) werden üblicherweise fünfzehn bis fünfundzwanzig Konzepte für die Erstellung einer (guten) *Concept Map* benötigt. Zudem zeichnet sich die Qualität einer *Concept Map* durch ihre sogenannte *Propositional Coherence* aus, nach der alle aufgeführten Propositionen, für sich genommen Sinn ergeben sollen (Moon et al., 2011).

¹³ In der Fachliteratur werden die Begriffe *Concept Map(s)*, *Begriffsnetz(e)*, *Wissensnetz(e)* oder *Knowledge Map(s)* synonym verwendet. In dieser Arbeit wird ausschließlich der Begriff *Concept Map(s)* verwendet.

Eine *Concept Map* ist im Übrigen von einer *Mind Map* abzugrenzen, bei der Ideen und Assoziationen zu einem vorgegebenen Begriff oder Sachverhalt gesammelt und netzartig dargestellt werden. Da hier ausschließlich assoziative Beziehungen hergestellt werden und es keine Informationen darüber gibt, in welcher Beziehung die Begriffe zueinander stehen, bildet eine *Mind Map* keine sachlogischen Zusammenhänge, sondern lediglich assoziative Verbindungen ab (Graf, 2014).

Concept Maps können, wie von Novak und Cañas (2006) empfohlen, hierarchisch aufgebaut werden. Hier werden die Konzepte nach ihrer Relevanz angeordnet. Wichtige Begriffe befinden sich oben und spezifischere Begriffe sind weiter unten in der *Concept Map* zu finden (siehe **Abbildung 4 A**). Der allgemeinste oder wichtigste Begriff wird als Schlüsselbegriff (*Root Concept*) bezeichnet (Novak & Cañas, 2006). Im Gegensatz zu einem hierarchischen Aufbau kann die Anordnung der Konzepte auch rund um einen Hauptbegriff, welcher im Zentrum der *Concept Map* steht, erfolgen und eine netzartige Struktur ergeben (siehe **Abbildung 4 B**; Áhlberg, 2004). Entscheidend für die Wahl der Struktur, also ob ein hierarchischer oder netzartiger Aufbau gewählt wird, ist laut Stracke (2004) das Themengebiet¹⁴.

¹⁴ Auch die zugrundeliegende theoretische Fundierung beeinflusst die Wahl der *Concept Map*-Struktur. Da im Rahmen von Assoziationstheorien eine netzartige Struktur von Gedächtnisinhalten angenommen wird, werden diesem Ansatz entsprechend auch netzartig strukturierte *Concept Maps* präferiert (Lopez et al., 2014; Ruiz-Primo & Shavelson, 1996). Der Assimilationstheorie folgend, wird die hierarchische Struktur bevorzugt.

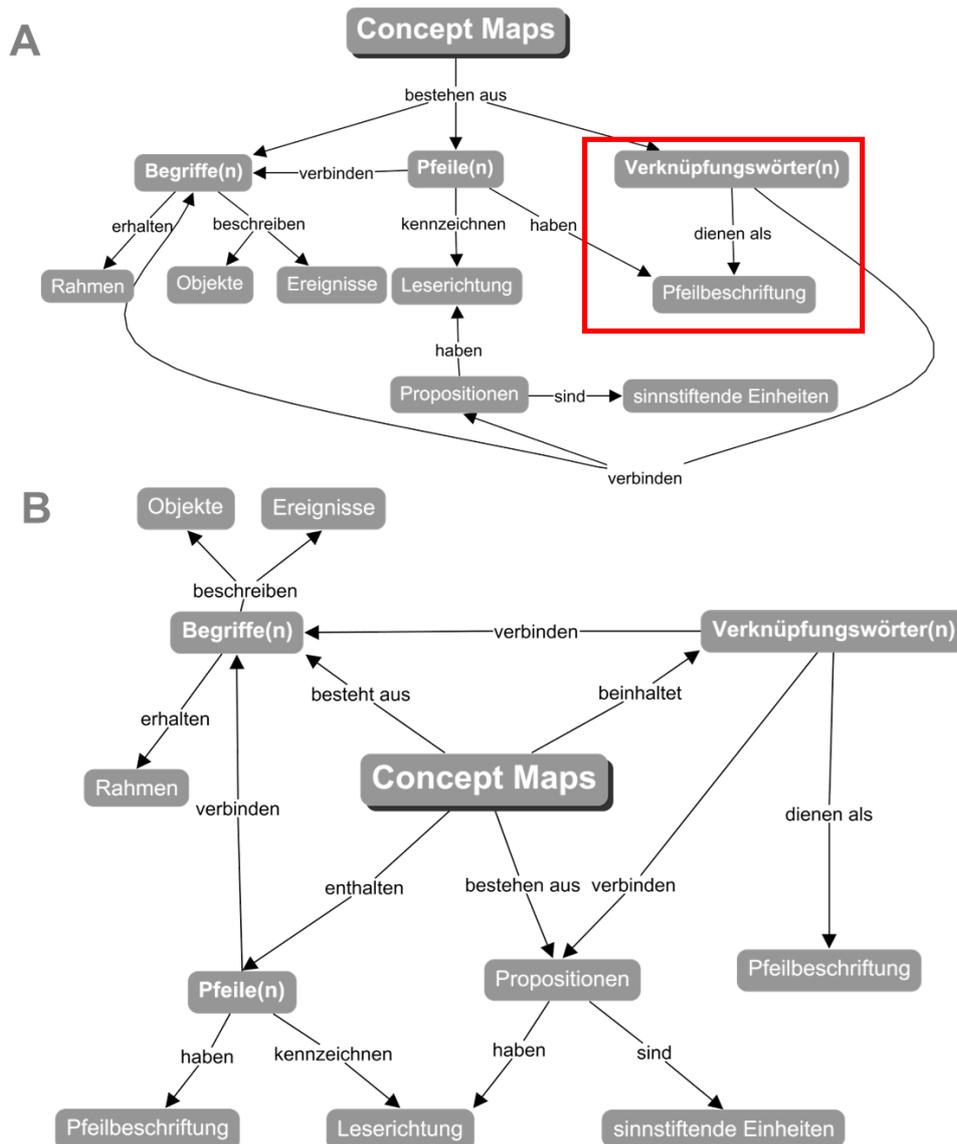


Abbildung 4. Eine beispielhafte *Concept Map* zum Thema *Concept Maps* hierarchisch (**A**) oder netzartig (**B**) strukturiert (die hierarchische *Concept Map* wurde entnommen aus Lenski und Großschedl [2021] nach Großschedl [2010]). Der rote Kasten umfasst eine beispielhafte Proposition bestehend aus den Begriffen *Verknüpfungswörter(n)* und *Pfeilbeschriftung*, welche über die semantische Beziehung *dienen als* miteinander verbunden sind¹⁵.

Die Arbeit mit *Concept Maps* kann aus der eigenen Erstellung von *Concept Maps* (Konstruktion¹⁶) oder der Betrachtung von vorgefertigten *Concept Maps* (Experten-Maps) bestehen (Rezeption). Zudem können sie sowohl als kognitive, wie auch als metakognitive

¹⁵ Alle in dieser Arbeit abgebildeten *Concept Maps* wurden mit Hilfe des Programms *CmapTools* erstellt.

¹⁶ Bis zur eigenständigen Erstellung von *Concept Maps* können Lernende verschiedene Unterstützungsmöglichkeiten angeboten werden. Bewährt hat sich beispielsweise die Verwendung von sogenannten Lücken-*Concept Maps* (*Skeleton Concept Maps*) bei denen bestimmte Elemente bereits vorgegeben sind (Chang et al., 2002; Gehl, 2013; Ruiz-Primo et al., 2001).

Lernstrategien eingesetzt werden (z. B. Novak, 1990)¹⁷. Hierdurch eröffnet sich ein breites Einsatzspektrum für die Nutzung von *Concept Maps*¹⁸.

Nach Gehl (2013) gibt es für *Concept Maps* drei Hauptanwendungsbereiche. Sie können als Lern- und Lehrmethode, als Instrument des Wissensmanagements oder zur Evaluation (Wissensdiagnose) eingesetzt werden. Das Lernen mittels *Concept Maps* kann kollaborativ (in Form von Partner- oder Gruppenarbeiten) oder eigenständig erfolgen (Haugwitz & Sandmann, 2009; Schroeder et al., 2017).

In Graf (2014) findet sich eine detaillierte Übersicht der Anwendungsmöglichkeiten von *Concept Maps*. Hier wird zwischen Lernenden (Ebene 1) und Lehrenden (Ebene 2) unterschieden. Für beide Ebenen gibt es jeweils drei Arten der Verwendung.

Der Ebene der Lernenden sind *die individuelle Prüfung von Wissensbeständen und der Wissenserwerb* (1), die *Organisationshilfe beim Erwerb konzeptbezogener Kompetenzen* (2) sowie die *Anregung fachbezogener Kommunikationskompetenzen* (3) zuzuordnen (Graf, 2014, S. 329). Unter der *individuellen Prüfung von Wissensbeständen und Wissenserwerb* werden Aspekte wie die Visualisierung des Wissenserwerbs durch Vorher-Nachher-Maps und die Evaluation des eigenen Wissens verstanden.

Der Aspekt der Organisationshilfe beim Erwerb konzeptbezogener Kompetenzen umfasst beispielsweise das strukturierte Analysieren von komplexen Beziehungen in Lerntexten und Unterrichtsmaterialien sowie deren Darstellung.

Der Einsatz von *Concept Maps* kann Lernenden somit dabei helfen, Schlüsselbegriffe in Lerntexten zu identifizieren, die Verbindung der Begriffe untereinander zu verstehen und sich diese einzuprägen. Dieser Aspekt wird von Gehl (2013) als Reduktionsfunktion zusammengefasst. Aufgrund der Tatsache, dass die Selektion der zentralen Information in allen Schulfächern wichtig ist, ist dieser Prozess im Unterricht von hoher Bedeutung. Die Strukturierungsfunktion umfasst die (An)ordnung der Begriffe und die Schaffung von Beziehungen untereinander (ebd.). Die Anregung fachbezogener Kommunikationskompetenzen, beruht auf der Annahme, dass die Kommunikation über und mittels *Concept*

¹⁷ Hierauf wird im weiteren Verlauf näher eingegangen.

¹⁸ Während manche Autor:innen den Begriff *Concept Mapping* lediglich für den Prozess der Konstruktion von *Concept Maps* verwenden (z. B. Jaafarpour et al., 2016; Sumfleth et al., 2010), wird er von anderen für die generelle Nutzung von *Concept Maps* verwendet (z. B. Akinsanya & Williams, 2004). Um Irritationen zu vermeiden, wird in dieser Arbeit auf die Nutzung des Begriffs verzichtet und stattdessen von der Konstruktion oder der Rezeption von *Concept Maps* gesprochen.

Maps zu einem besseren Verständnis von Lernhalten und deren Verfestigung führt sowie das kreative Denken anregt (Kommunikationsfunktion; Gehl, 2013).

Auf der Ebene der Lehrenden umfassen die Einsatzmöglichkeiten die *Strukturanalyse und -darstellung von Inhaltsbereichen* (4), die *Hilfe bei der Lehrplanentwicklung und Unterrichtsplanung* (5) und die *Evaluation des eigenen Unterrichts* (6). Unter dem Aspekt der Strukturanalyse und -darstellung von Inhaltsbereichen fällt beispielsweise die Erstellung von Experten-*Concept Maps* (ebd.). Auf diese Art können *Concept Maps* Lehrende bei der Strukturierung und Planung von Unterricht unterstützen, da sie eine strukturierte und übersichtliche Visualisierung von Sachverhalten ermöglichen. Im nächsten Schritt können sie die Kommunikation von Unterrichtsinhalten an die Schüler:innen erleichtern. Der Einsatz von *Concept Maps* als Diagnoseinstrument im Rahmen des Evaluationsprozesses hilft bei der Erfassung des Wissenstands der Schüler:innen (Stracke, 2004)¹⁹.

2.2 Die Lernwirksamkeit von Concept Maps

Bei den geschilderten Einsatzmöglichkeiten von *Concept Maps*, nimmt der Aspekt des Lernens und somit der Lernwirksamkeit eine wichtige Rolle ein. Um die Lernförderlichkeit von *Concept Maps* auf theoretischer Ebene zu erklären, wurden unterschiedliche Eigenschaften von *Concept Maps* betrachtet. Es wurden verschiedene mögliche Faktoren identifiziert, welche die Lernförderlichkeit begründen können (Nesbit & Adesope, 2006; Schroeder et al., 2017). Diese Faktoren lassen sich in drei Kategorien zusammenfassen: die *Abbildung der kognitiven Wissensstruktur* (1), die *Art der Darstellung* (2) und die *Verdeutlichung des eigenen Wissenstands* (3).

In Bezug auf die Abbildung der kognitiven Wissensstruktur, ermöglichen *Concept Maps* nach Novak und Cañas (2006) die Externalisierung oder Abbildung von Wissen in einer Weise, die den kognitiven Strukturen (semantischen Netzwerken) in welcher das Wissen im Gedächtnis gespeichert wird, ähnelt (vgl. **Abbildung 2** und **4**). Da das Wissen im Gedächtnis nicht direkt sichtbar gemacht werden kann, können *Concept Maps* einen indirekten Zugang schaffen. Dies ermöglicht es subjektive mentale Modelle offen zu legen

¹⁹ Im Folgenden wird der Fokus auf die Verwendung von *Concept Maps* als Lernstrategie gelegt, da dieser Aspekt Gegenstand dieser Arbeit ist. Die anderen Anwendungsgebiete werden nicht weiter thematisiert. Weiterführende Literatur zu *Concept Maps* als Diagnoseinstrument umfassen unter anderem die Arbeiten von Graf (2014), Haugwitz et al. (2010) sowie Hahn-Laudenberg (2017).

und zu kommunizieren (Jonassen et al., 1997; Shavelson et al., 2005)²⁰. Die Lernförderlichkeit ergibt sich daraus, dass die Verwendung *Concept Maps* personalisiertes Feedback und Diskussionen über den aktuellen Wissensstand Lernenden ermöglicht (Anohina - Naumeca, 2015). Der Wissens- und Erfahrungsaustausch mit anderen Lernenden kann durch die kollaborative Nutzung von *Concept Maps* oder die Besprechung selbst erstellter *Concept Maps* erleichtert werden (Nesbit & Adesope, 2006; Schaal et al., 2010).

In Bezug auf die Art der Darstellung, werden in *Concept Maps* wichtige Informationen in komprimierter Form visualisiert. Auf diese Weise können Lernende die Inhalte schnell erfassen und die Aufmerksamkeit wird auf die relevanten Aspekte gelenkt (Lambiotte et al., 1989; Nesbit & Adesope, 2006). Basierend auf Paivios *Dual Coding Theory* (1986)²¹ nehmen Adesope und Nesbit (2013) an, dass die Kombination aus verbalen Informationen (Konzepte und Relationen) und visuellen Informationen (Pfeile und Kästen) in *Concept Maps* dazu führt, dass Inhalte schneller aufgenommen werden können. Durch die kombinierte Darstellung verbaler und non-verbaler Informationen kann die Belastung des Arbeitsgedächtnisses²² reduziert (Amadiou, 2009) und so die kognitive Verarbeitung dieser Informationen verbessert werden (O'Donnell et al., 2002). Dadurch, dass in *Concept Maps* das Wissen analog zu kognitiven Strukturen organisiert ist (s.o.), sind komplexe Umkodierungsprozesse nicht notwendig. Diese Prozesse sind etwa für das Bilden semantischer Netzwerke, die auf Textinformationen basieren, zwingend (Fürstenau, 2011). Aufgrund der einfachen Syntax gibt es einen niedrigschwelligen Zugang zu den Informationen für Lernende, die (noch) über geringe Lese- und Schreibfähigkeiten verfügen (Schmid & Telaro, 1990)²³. Auch für Lernende mit geringeren kognitiven Fähigkeiten erwies sich die Verwendung von *Concept Maps* als vorteilhaft (Haugwitz et al., 2010).

Das Verständnis für komplexe Beziehungen nimmt eine Schlüsselrolle für bedeutungsvolles Lernen ein, stellt Lernende jedoch vor eine große Herausforderung (z. B. Reiss et

²⁰ Um von Lernenden erstellte *Concept Maps* in Bezug auf unterschiedliche Aspekte zu analysieren, wurden verschiedene Verfahren (*Mentale Representation Analyses*) entwickelt (Clausen, & Christian, 2012; Holley & Danserau, 1984; McClure et al., 1999).

²¹ Die *Dual Coding Theory* wird in **Kapitel 3.2** vorgestellt.

²² Genauer gesagt, wird die sogenannte extrinsische kognitive Belastung (*Extraneous Cognitive Load*) hierdurch reduziert. In **Kapitel 1.2** wurde bereits näher auf die unterschiedlichen Arten der kognitiven Belastung eingegangen.

²³ Propositionen müssen nicht grammatikalisch korrekt formuliert werden (Abrams et al., 2006). Da *Concept Maps* im Rahmen dieser Arbeit im Unterricht eingesetzt wurden, wurde auf die grammatikalisch korrekte Gestaltung der Propositionen geachtet. Die Schüler:innen sollten durch fehlerhafte Grammatik nicht irritiert bzw. verunsichert oder vom Lerngegenstand abgelenkt werden.

al., 2019). Lernende können schnell durch die Anzahl an verbundenen Informationen (Informationen mit hoher Elementinteraktivität²⁴) überfordert werden (van Gog et al., 2010). *Concept Maps* heben semantische Beziehungen besser hervor als Texte, indem sie die genauen Beziehungen zwischen den Konzepten verdeutlichen (Schroeder et al., 2017). Verbundene Informationen werden in der *Concept Map* in räumlicher Nähe zueinander abgebildet. Die gebündelte Darstellung von Informationen in *Concept Maps* wird unter anderem durch die Vermeidung von Dopplungen ermöglicht. In Texten können sich zusammengehörige Informationen hingegen an verschiedenen Stellen befinden (Nesbit & Adesope, 2006; Adesope & Nesbit, 2013). Ferner wird angenommen, dass das Hervorheben der semantischen Beziehungen die Komplexität eines Inhalts reduzieren, indem das Arbeitsgedächtnis darin unterstützt wird eine mentale Repräsentation zu erstellen (Jonassen et al., 1997).

Da *Concept Maps* die Verdeutlichung des eigenen Wissensstands ermöglichen, können sie die Erlangung metakognitiven Wissens unterstützen. Metakognitives Wissen wird definiert als das Bewusstsein von Lernenden über ihre kognitiven Aktivitäten sowie deren Überwachung und Kontrolle (Flavell, 1976). Durch die Offenlegung eigener Wissenslücken, Fehlvorstellungen (*Misconceptions*) und formaler Denkfehler (Tergan, 1989), werden die Lernenden bei diesen metakognitiven Prozessen unterstützt. Dies ermöglicht anschließend die gezielte Auswahl von Lernstrategien. Mit Hilfe der ausgewählten Strategien gelingt es den Lernenden den eigenen Lernprozess zu steuern, zu überwachen und das Verstehen von Inhalten zu erleichtern (van Velzen, 2015). Eine solche Ausrichtung des eigenen Lernprozesses wird als metakognitives Lernen bezeichnet.

Bei der Konstruktion und Rezeption von *Concept Maps* sind die kognitiven und metakognitiven Effekte hinsichtlich ihrer Ausprägung unterschiedlich stark. Dies ist darauf zurückzuführen, dass es sich bei der Konstruktion von *Concept Maps* um einen aktiven Prozess handelt, während die Rezeption von *Concept Maps* vergleichsweise passiver Natur ist (Hardy & Stadelhofer, 2006). Bei der Konstruktion müssen basierend auf einem Lerntext²⁵ Schlüsselwörter aktiv identifiziert und in Beziehung gesetzt werden. Dieser

²⁴ Der Grad der Elementinteraktivität gibt Aufschluss über die Anzahl an Elementen, die gleichzeitig im Arbeitsgedächtnis aktiv verarbeitet werden müssen.

²⁵ Meist werden *Concept Maps* basierend auf einem (Lern)text konstruiert, wobei sie auch aus dem Gedächtnis heraus erstellt werden können (Blunt & Karpicke, 2014).

Prozess fordert und fördert ein hohes Maß an Elaboration, was ein wichtiger Bestandteil des bedeutungsvollen Lernens ist (Schroeder et al., 2017).

Bei der Rezeption versuchen die Lernenden, meist ohne zusätzliche Information in textlicher Form, anhand der vorliegenden *Concept Map*, Zusammenhänge zu identifizieren und zu verstehen. Zusätzlich zu den kognitiven, elaborativen Prozessen, werden von den Lernenden metakognitive Strategien eingesetzt, um fehlende Detailinformationen hinzuzufügen. Die Rezeption von *Concept Maps* fördert nach O'Donnell et al. (2002) ebenfalls das bedeutungsvolle Lernen, da die wichtigen Konzepte schnell erfasst und bereits bekannte Konzepte leicht identifiziert werden können. So wird die Integration neuer Informationen in bereits bestehende Wissensstrukturen erleichtert. Die Abbildung der eigenen kognitiven Wissensstruktur entfällt bei der Rezeption von *Concept Maps*.

Durch die Verwendung eines korrekten semantischen Netzwerks in Form von vorgefertigten (Experten-) *Concept Maps* können Lernende im Rahmen eines Vergleichs der Informationen den eigenen Wissensstand evaluieren und Denkfehler identifizieren.

Verglichen mit dem Lesen eines Textes, findet bei der Rezeption einer *Concept Map* ein höheres Maß an kognitiven und metakognitiven Prozessen statt (Schroeder et al., 2017). Ein noch höheres Maß an Verarbeitung des Inhalts wird für die Konstruktion von *Concept Maps* angenommen (ebd.). Daraus wird geschlussfolgert, dass die Konstruktion kognitiv anspruchsvoller ist und eine höhere kognitive Belastung verursacht, als die Rezeption (ebd.). Dennoch scheint sich in der Fachliteratur die Annahme durchzusetzen, dass für die Konstruktion von *Concept Maps* eine höhere Lernwirksamkeit erwartet werden kann, als für deren Rezeption (Schmid & Telaro, 1990).

Aktuelle empirische Befunde deuten ebenfalls darauf hin, dass die Konstruktion von *Concept Maps* lernförderlicher ist, als die Rezeption, wobei die Befundlage keine eindeutige Interpretation zulässt (siehe **Kapitel 2.3**).

2.3 Die empirische Befundlage zur Lernwirksamkeit von Concept Maps

Obwohl die Entwicklung von *Concept Maps* bereits mehrere Jahrzehnte zurück liegt und aufgrund intensiver Forschung bereits hunderte Primärstudien²⁶ zu diesen Themengebiet

²⁶ Aufgrund der Fülle an Primärstudien wird an dieser Stelle darauf verzichtet, über einzelne Studien im Detail zu berichten, stattdessen werden die Befunde einschlägiger Metaanalysen vorgestellt.

vorliegen, ist es nach wie vor relevant. Nesbit und Adesope (2006) beobachteten, dass die Menge an Publikationen im Zeitraum von 1985 bis 2006 stark angestiegen ist. Eine Suche im *Web of Science* für den Zeitraum von 2006 bis 2021 ergab, dass sich dieser Trend fortsetzt. Allein im Jahre 2021 wurden mehr als 200 Arbeiten zum Thema *Concept Maps* veröffentlicht²⁷. Obwohl sich der Einsatz von *Concept Maps* für alle Domänen eignet, beruht ein Großteil der bisherigen Forschung auf Untersuchungen in MINT-Fächern²⁸ (Schroeder et al., 2017). Der Beweggrund war die zentrale Bedeutung der Vermittlung von Zusammenhangswissen in diesen Schulfächern (Schmid and Telaro, 1990). In den meisten Primärstudien wurde eins der beiden *Concept Map*-Formate (Konstruktion oder Rezeption) mit einer anderen Lernmethode, etwa dem Lesen von Texten, dem Schreiben von Zusammenfassungen oder dem Diskutieren verglichen. Die Konstruktion von *Concept Maps* erwies sich dabei als lernförderlicher als das Lesen von Texten, das Schreiben von Zusammenfassungen und Notizen, das Diskutieren, das Markieren von Texten sowie das Besuchen einer Vorlesung (Horton et al., 1993; Nesbit & Adesope, 2006; Schroeder et al., 2017).

Die Effekte sind am größten, wenn die Konstruktion von *Concept Maps* mit passiveren Lernmethoden, wie dem Lesen von Texten, oder dem Besuch von Vorlesungen verglichen wurde. Der Vergleich mit ebenfalls aktiven Lernmethoden (z. B. dem Schreiben von Zusammenfassungen oder Notizen), ergab hingegen geringere Effekte.

Das Rezipieren von *Concept Maps* übertraf passive Lernmethoden, wie das Textstudium, das Lesen von Listen und den Besuch von Vorlesungen (Horton et al., 1993; Nesbit & Adesope, 2006; Schroeder et al., 2017). Auch bei einem Vergleich zu aktiveren Lernmethoden (z. B. das Führen von Gruppendiskussionen) erwies sich die Rezeption von *Concept Maps* als lernförderlicher (ebd.). Aus diesen Ergebnissen lässt sich ableiten, dass die Verwendung beider *Concept Map*-Formate das Lernen effektiv fördern kann (Hattie, 2008; Horton et al., 1993; Nesbit & Adesope, 2006, Schroeder et al., 2017). Die Frage, ob das Konstruieren oder das Rezipieren von *Concept Maps* lernförderlicher ist, ist in der Forschung nach wie vor umstritten. Horton et al. (1993) beobachteten einen größeren

²⁷ Hierfür wurden dieselben Suchbegriffe *Concept Map* OR Knowledge Map* OR Node-link Map** wie in der Metaanalyse von Schroeder et al. (2017) verwandt. Die Suche wurde auf Titel und Abstract begrenzt. Einbezogen wurden Primärartikel, Übersichtsarbeiten, *Early Access*-Artikel, Tagungsbandbeiträge und Buchkapitel. (Stand: März, 2022).

²⁸ Unter der Abkürzung MINT werden die Bereiche Mathematik, Informatik, Naturwissenschaften und Technik zusammengefasst. Hierunter fallen folgende Schulfächer: Mathematik, Informatik, Biologie, Chemie, Physik, Astronomie und Geographie (Steuer, 2014).

Effekt für die Rezeption ($d = .59$) verglichen mit der Konstruktion von *Concept Maps* ($d = .42$). Anzumerken ist hier, dass diese Analyse 18 Studien umfasste, wobei in 15 Studien die Konstruktion und lediglich in drei Studien die Rezeption untersucht wurde. Zu einem gegenteiligen Ergebnis kamen Nesbit und Adesope (2006) in einer größer angelegten Metaanalyse, welche 67 Effektgrößen umfasste. Sie beobachteten eine Überlegenheit der Konstruktion ($g = .82$)²⁹ gegenüber der Rezeption ($g = .37$). Schroeder et al. (2017) stützten diese Befunde und berichteten ebenfalls einen größeren Effekt für die Konstruktion ($g = .72$) als für die Rezeption ($g = .43$).

Eine Begründung für die gegensätzlichen Befunde könnten die bereits angemerkten Unterschiede in der Größe der Stichprobe sowie in der Gewichtung hinsichtlich der berücksichtigten *Concept Map*-Formate liefern. Schroeder et al. (2017) merken auch an, dass in keiner der einbezogenen Studien die Auswirkungen von Konstruktion und Rezeption im Rahmen einer experimentellen Studie direkt verglichen wurden. Da die Effekte somit sehr stark von den eingesetzten Vergleichsgruppen abhängen, sind die Vergleiche anhand der Effektstärken nur eingeschränkt interpretierbar. Dass die Lernenden in den verschiedenen Studien unterschiedlich viel Erfahrung mit der Nutzung von *Concept Maps* hatten, könnte ebenfalls zu einer Verzerrung der Ergebnisse führen (McCagg & Dansereau, 1991).

2.4 Der Einfluss der Vertrautheit mit Concept Maps auf die Lernwirksamkeit

Lernende, die mit der Nutzung von *Concept Maps* noch nicht vertraut sind (Noviz:innen³⁰), fühlen sich häufig beim Lernen mit *Concept Maps* überfordert (Blankenship & Dansereau, 2000). Zu dieser Erkenntnis kam man sowohl für die Konstruktion als auch für der Rezeption. Bei der Konstruktion von *Concept Maps* kann beispielsweise die Identifikation der wichtigsten Begriffe in einem Lerntext und die Planung des Konstruktionsprozesses ungeübten Lernenden schwerfallen (Schroeder et al., 2017). Bei der Rezeption von *Concept Maps* kann die unbekannte Darstellung der Informationen dazu führen, dass

²⁹ Bei Hedges g handelt es sich um eine häufig verwendete Effektstärke basierend auf der Mittelwertsdifferenz, welche insbesondere bei unterschiedlich großen Stichproben sinnvoll sein kann (Schroeder et al., 2017).

³⁰ In der vorliegenden Arbeit bezieht sich die Einteilung zwischen Expert:innen und Noviz:innen auf den Grad der Expertise im Umgang mit *Concept Maps* und nicht auf das Fachwissen.

Lernende einen sogenannten *Map Shock* erleben. Dieser Schock äußert sich laut Blankenship und Dansereau (2000) darin, dass Schüler:innen nicht wissen, wo sie bei der Betrachtung von *Concept Maps* beginnen und wie sie deren Struktur verstehen sollen. Die Freiheit in der Wahl der Leserichtung und die hiermit verbundenen Entscheidungsprozesse können insbesondere für Noviz:innen kognitiv anspruchsvoller sein als das Lesen von Texten, bei denen die Reihenfolge der Informationen vorgegeben ist (Adesope & Nesbit, 2013; Blankenship & Dansereau, 2000).

Die aus der hohen kognitiven Belastung resultierende Überforderung kann zur Verunsicherung und Demotivation der Lernenden führen (ebd). Dieser Umstand könnte bei Lehrenden zur Folge haben, dass sie der Integrierung von *Concept Maps* in ihrem Unterricht skeptisch gegenüberstehen. Diese Zurückhaltung macht es Schüler:innen wiederum unmöglich Vertrautheit mit der Methode zu erwerben und so die hohe anfängliche kognitive Belastung signifikant zu verringern.

Die Vertrautheit in der Nutzung von *Concept Maps* spielt nach McCagg und Dansereau (1991) eine zentrale Rolle für die Akzeptanz dieser Lernmethode sowie für ihre Effizienz in Bezug auf die Lernwirksamkeit. Auch Schroeder et al. (2017) konnten in ihrer Metaanalyse beobachten, dass mit steigender Dauer der Nutzung von *Concept Maps* auch eine Zunahme der Lernwirksamkeit zu verzeichnen war. Laut Jüngst und Schrittmacher (1995) erfordert der adäquate Gebrauch von *Concept Maps* neben faktischem Wissen über die Methode an sich (deklaratives Wissen) auch Wissen über ihre praktische Anwendung (prozedurales Wissen). Für die Vermittlung dieser Kenntnisse und Fähigkeiten werden vorgelagerte Einführungs- oder Trainingsmaßnahmen empfohlen (Becker et al., 2021a; Fürstenau, 2011; Großschedl & Tröbst, 2018; Hilbert & Renkl, 2008; McCagg & Dansereau, 1991; Mintzes et al., 2011; Renkl & Nückles, 2006). Diese sollen den Lernenden den Umgang mit *Concept Maps* vermitteln und einer anfänglichen Überforderung entgegenwirken. Hilbert und Renkl (2008) stellten bei Lernenden, die in der Konstruktion von *Concept Maps* bereits geübt waren, eine höhere Lernförderlichkeit im Vergleich zu Lernenden fest, die keine Erfahrung im Umgang mit *Concept Maps* hatten.

Lernende, die in der Konstruktion von *Concept Maps* bereits geübt waren, planten den Prozess der Konstruktion häufiger und erstellten daher kohärentere *Concept Maps*. Fehlende Vertrautheit hatte in einer Studie von Sumfleth et al. (2010) hingegen zur Folge, dass die *Concept Maps* geringere Qualität in Bezug auf Struktur und Inhalt aufwiesen. Welche negativen Ausmaße die Nutzung von *Concept Maps* annehmen kann, wenn sie nicht adäquat beherrscht wird, verdeutlicht eine Studie von Neuroth (2007). Hier wirkte sich die Konstruktion von *Concept Maps* bei Noviz:innen sogar negativ auf die Lernleistung aus. Bislang konzipierte Trainingsmaßnahmen zur Einführung von *Concept Maps* orientieren sich häufig an den Empfehlungen von (Novak & Cañas, 2006). Hier wird die Konstruktion von *Concept Maps* in vier aufeinanderfolgende Schritte eingeteilt (ebd.). Im ersten Schritt werden zentrale Begriffe aus einem Lerntext identifiziert, indem die enthaltenen Informationen auf den wesentlichen Inhalt reduziert werden. Um die Einführung von *Concept Maps* zu erleichtern, raten Novak und Gowin (1984) dazu mit einem Thema zu beginnen mit dem die Lernenden bereits vertraut sind. Zudem wird die Verwendung einer Fokusfrage (*Focus Question*), welche das Thema oder das Problem spezifiziert, empfohlen. Im zweiten Schritt werden die Konzepte gruppiert, organisiert und angeordnet (im Falle einer hierarchisch strukturierten *Concept Map* erfolgt die Anordnung nach Wichtigkeit). Der dritte Schritt besteht aus der Erstellung von Propositionen, also dem Verbinden der Konzepte untereinander. Im vierten und letzten Schritt wird geprüft, ob die erstellte *Concept Map* die Fokusfrage beantwortet und alle wichtigen Informationen beinhaltet.

Wenn diese Empfehlungen bei der Konzeption von Trainingsmaßnahmen für die Einführung von *Concept Maps* berücksichtigt werden, ergibt sich hinsichtlich Dauer, Häufigkeit, Einführungsmethode, Themenwahl sowie Grad der Unterstützung ein großer Interpretationsspielraum (Becker et al., 2021).

Die Dauer der Trainingsmaßnahmen variierte zwischen fünf Minuten (Blunt & Karpicke, 2014), zehn bis 30 Minuten (Hay et al. 2008; Klein et al., 2002), 60-90 Minuten (Becker et al., 2021; Gouli et al., 2003), und mehreren Wochen (Quinn et al., 2003).

Hier ist anzumerken, dass die angegebenen Trainingszeiten sich teilweise auf eine Sitzung beziehen (z. B. Blunt & Karpicke, 2014), aber auch mehrere Sitzungen umfassen können (z. B. Becker et al., 2021a; Quinn et al., 2003). Falls die Einführung von *Concept*

Maps in mehreren Sitzungen erfolgt, existieren verschiedene Ansätze bezüglich der zeitlichen Abstände zwischen den einzelnen Sitzungen.

Für eine längere Trainingsdauer spricht die Annahme, dass bei der Einführung einer neuen Lernstrategie Zeit für die Automatisierung eingeräumt werden muss (Hasselhorn & Gold, 2006). Erst wenn ein gewisser Automatisierungsgrad erreicht ist, kann mit einer Verbesserung der Lernleistung durch die Anwendung der neuen Strategie gerechnet werden (ebd.). In Bezug auf die Trainingsmethode kann im Wesentlichen zwischen zwei Einführungsformen unterschieden werden. Nach einer Methode soll ausschließlich deklaratives Wissen über *Concept Maps* vermittelt werden, nach der anderen soll sowohl deklaratives wie auch prozedurales Wissen abgedeckt werden. Auch in Bezug auf den Grad der Unterstützung gibt es für die Trainingsmaßnahmen verschiedene Ansätze. Lernende können sich die Konstruktion von *Concept Maps* beispielsweise anhand von Handouts selbst erschließen (Chiu, 2004), oder im Rahmen eines Trainingsprogramms angeleitet werden (z. B. Quinn et al., 2003).

Die erwähnten Interpretationsspielräume ergeben sich weiter aus dem Umstand, dass Trainingsmaßnahmen für unterschiedliche Ziel- und Altersgruppen konzipiert wurden. Es existieren Studien zu Grundschüler:innen (Chang et al., 2002), Schüler:innen der Mittel- und Oberstufen (Neuroth, 2007; Sumfleth et al., 2010), Studierenden (Quinn et al., 2003) und Berufstätigen (Daley et al., 2016). Erschwerend hinzu tritt, dass unterschieden werden muss, ob die Trainingsmaßnahme für eine PC-basierte Konstruktion von *Concept Maps* oder für eine papierbasierte (analoge) Anwendung konzipiert wurde. Computerbasierte Trainingsmaßnahmen (z. B. den Elzen-Rump & Leutner, 2007) setzen den Schwerpunkt oft auf den Umgang mit dem Gerät selbst und die verschiedenen Nutzungsmöglichkeiten der Software (meist *CmapTools*).

Aufgrund der vielen verschiedenen Trainings- und Einsatzmöglichkeiten fallen die Ergebnisse hinsichtlich der Lernwirksamkeit entsprechend heterogen aus.

Deshalb wird in der Forschung kontrovers diskutiert, ob eine vorgelagerte Trainingsmaßnahme notwendig und sinnvoll ist, und falls ja, wie diese aussehen soll (z. B. Correia et al., 2008; Ifenthaler, 2011; Karpicke & Blunt, 2011; Mintzes et al., 2011; Renkl & Nückles, 2006). Bei der Vielzahl an Stellschrauben, an denen im Rahmen einer *Concept Map*-Trainingsmaßnahme gedreht werden kann, ist es nicht verwunderlich, dass die bisherige Befundlage kein eindeutiges Bild ergibt.

3 Lernen mit Multimedia

3.1 Die Definition von Multimedia für den Lernkontext

In der wissenschaftlichen Literatur zum Thema Lernen mit Multimedia wird der Begriff *Multimedia* unterschiedlich definiert (Rey, 2009; Weidenmann, 2002). Nach Weidenmann (2002) umfasst die Definition drei Merkmale: Multimedialität, Multikodalität und Multimodalität³¹. Unter Multimedialität versteht man die Nutzung verschiedener technischer Geräte. Beispielweise wäre eine am Computer erstellte und im Anschluss mit Hilfe eines Beamer gehaltene Präsentation multimedial. Multikodalität bezieht sich auf die Nutzung verschiedener Symbolsysteme. Diese Symbolsysteme umfassen Informationen, die piktoral-bildhafter und textuell-verbaler Natur sind. Lernmaterial, das aus einer Bild-Text-Kombination besteht, ist somit multikodal. Nur für sich genommen werden einzelne Bilder oder Texte als monokodal klassifiziert³². Multimodalität liegt vor, wenn verschiedene Sinne angesprochen werden (hauptsächlich der Hör- und Sehsinn; Mayer, 2001). Eine Lernsoftware, die aus Bildern und Tönen besteht, wäre ein Beispiel für ein multimodales und multikodales Lernangebot (Schnotz, 2002).

Für das multimediale Lernen sind zwei Theorien maßgeblich: die *Cognitive Load Theory* von Sweller und Kollegen (z. B. Chandler & Sweller, 1991; Sweller, 2005) und die *Cognitive Theory of Multimedia Learning* von Mayer (2001)³³.

3.2 Die Grundannahmen der Cognitive Theory of Multimedia Learning

Unter multimedialem Lernen wird nach der *Cognitive Theory of Multimedia Learning* (CTML³⁴) von Mayer (2002) das Lernen aus einer Kombination von verbalen und non-verbalen Informationen verstanden. Verbale Informationen können in auditiver (z. B. ein

³¹ Interaktivität wird nicht berücksichtigt, da dieser Aspekt für die vorliegende Arbeit nicht von Relevanz ist. Weiterführende Literatur bietet beispielweise die Studie von Kirsh (1997).

³² Für weitere Informationen siehe Weidenmann (2009).

³³ Weitere für das Lernen mit Multimedia bedeutende Überlegungen wurden von Schnotz (2005) im Rahmen des *integrativen Modells des Text- und Bildverständnisses* zusammengefasst. Da sich dieses Modell nur in wenigen Punkten von der *Cognitive Theory of Multimedia Learning* (Mayer, 2001) unterscheidet, auf dem die vorliegende Arbeit thematisch basiert, wird auf eine separate Darstellung dieses Modells verzichtet. Auf Unterschiede wird an entsprechenden Stellen hingewiesen.

³⁴ Im Folgenden wird für die *Cognitive Theory of Multimedia Learning* die Abkürzung CTML verwendet.

Vortrag) oder in visueller Form (z. B. ein Lerntext) vorliegen. Die non-verbale Informationen können visuell, als statische (z. B. ein Foto) oder dynamische Abbildungen (z. B. ein Video) präsentiert werden (Mayer & Moreno, 2003). Multimediale Lernmaterialien sollen lernförderlicher sein als zum Beispiel rein textbasierte Lernmaterialien (sog. *Multimedia Effect*; Mayer, 2001). Beispiele für multimediale Lernmaterialien sind Textbücher, welche illustrierte Lerntexte umfassen, computerbasierte Lernumgebungen, welche Animationen und Audiosequenzen enthalten, oder folienbasierte Vorträge, welche graphische Elemente aufweisen (Mayer, 2014). Die *CTML* wurde entwickelt, um die kognitiven Prozesse beim Lernen multimedialer Informationen herzuleiten und versucht die Lernförderlichkeit theoretisch zu fundieren. Darüber hinaus stellt sie das Bindeglied zwischen kognitionspsychologischen Theorien und empirischer Forschung dar. Ziel dieser Theorie ist es Gestaltungs- und Handlungsempfehlungen für das Lernen mit Multimedia zu entwickeln.

Die Grundannahmen der *CTML* stützen sich neben der bereits vorgestellten *Cognitive Load Theory* (Chandler & Sweller, 1991; Sweller et al., 1998) auf weitere etablierte Theorien und Modelle. Diese umfassen das *Multi-store Model* nach Atkinson und Shiffrin (1968), Paivios *Dual Coding Theory* (1986), Baddeleys *Model of Working Memory* (1986) und *Generative Theory of Learning* (Wittrock, 1974)³⁵.

Die *Dual Coding Theory* von Paivio (1986) bildet die Basis vieler verbreiteter Theorien zur Wirkung von bildlichen und textlichen Darstellungen. Nach dieser Theorie gibt es zwei separate Kanäle (Systeme), die für die Verarbeitung verbaler und non-verbaler Informationen zuständig sind (siehe **Abbildung 5**). Die Kodalität der Informationen dient in diesem Modell somit als Differenzierungskriterium. Während das verbale System für das Lesen, Hören und Sprechen verantwortlich ist, werden im non-verbale System akustische, bildliche, haptische sowie olfaktorische Stimuli verarbeitet. Diese beiden Systeme sind miteinander verbunden und können nacheinander aber auch gleichzeitig aktiviert werden. Eine gleichzeitige Aktivierung beider Systeme führt zu einer Doppel-Kodierung (*Dual Coding*) der Information(en). Diese Form der Aktivierung wird vor allem bei der

³⁵ Das *Multi-store Model* nach Atkinson und Shiffrin (1968) sowie die *Cognitive Load Theory* (Sweller, 1998) wurden eingangs bereits vorgestellt (siehe **Kapitel 1.1**).

Verwendung von Bildern angenommen (Paivio, 1986), da Bildern automatisch verbale Informationen zugeordnet werden (*Picture-Superiority-Effect*). Informationen aus Texten werden hingegen nur dann bildliche Informationen zugeordnet, wenn diese konkrete und imaginierbare Inhalte betreffen (z. B. Baum oder Auto). Abstrakte Inhalte, welche schwer vorstellbar sind (z. B. Freiheit oder Gerechtigkeit) und somit keinem konkreten Bild zuzuordnen sind, werden hingegen nicht doppelt kodiert. Kritisiert wird dieses Modell unter anderem aufgrund seiner starken Vereinfachung (Kousta et al., 2011)³⁶.

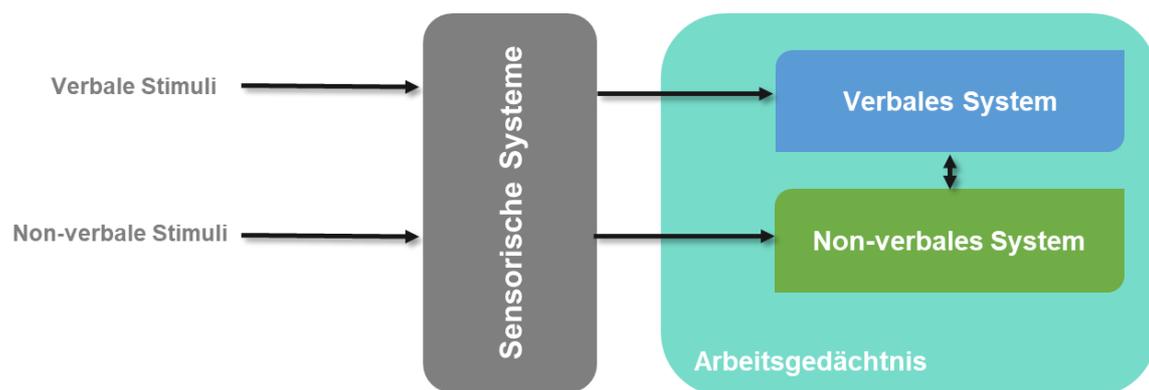


Abbildung 5. Die *Dual Coding Theory* nach Paivio (1986). Die Darstellung wurde angelehnt an Becker-Carus und Wendt (2017).

Basierend auf der *Generative Theory of Learning* (Wittrock, 1974) wird im Rahmen der *CTML* von einer aktiven Informationsverarbeitung durch die Lernenden ausgegangen. Kohärente mentale Repräsentationen von Informationen werden hiernach im Zuge eines aktiven Lernprozesses konstruiert (Mayer, 2001). Der Lernprozess wird somit maßgeblich durch die Lernenden mitbestimmt. Für die aktive Informationsverarbeitung sind drei kognitive Prozesse entscheidend: die Selektion von relevanten Informationen, die Organisation der ausgewählten Informationen und deren Integration unter Einbezug des Vorwissens.

3.3 Cognitive Theory of Multimedia Learning

Nach der *CTML* werden text- und bildbasierte Informationen eines multimedial gestalteten (Lern-)materials zunächst über die menschlichen Sinnesorgane aufgenommen (siehe **Abbildung 6 A**). Zu Beginn wird also betrachtet, in welcher Form (verbal oder

³⁶ Aufgrund dieser Kritik erweiterte Paivio seine Theorie (z. B. Paivio, 2006, 2014).

non-verbal) die eingehenden Informationen kodiert sind (Präsentationsmodus) und dann mit welchem Sinnesorgan (Ohr oder Auge) die Informationen aufgenommen werden (Modalitätsmodus). Die Aspekte der Modalität und der Kodalität, welche von Baddeley (1992) und Paivio (1986) separat betrachtet wurden, werden im Rahmen der *CTML* somit kombiniert berücksichtigt.

Bei der Aufnahme von Informationen in das Arbeitsgedächtnis findet ein Selektionsprozess statt. Lediglich Wörter und Bilder, die als relevant eingestuft werden, gelangen in das kapazitätsbegrenzte Arbeitsgedächtnis (Mayer & Moreno, 2003). Im Rahmen dieses Prozesses werden aus relevanten Wörtern verbale Repräsentationen und aus relevanten Bildern piktorale Repräsentationen. Eine Besonderheit bei der *CTML* ist die Möglichkeit, dass im Arbeitsgedächtnis Transformationsprozesse stattfinden können. Geschriebener Text, der zunächst über das non-verbale System aufgenommen und verarbeitet wurde, kann hier in das verbale System überführt werden (*Cross-channel Representation*). So kann ein geschriebener Text nicht nur in Form eines Bildes, sondern auch in Form einer gesprochenen Lautfolge verarbeitet werden (Mayer & Moreno, 2003). Auch aus einer auditiven Repräsentation, wie zum Beispiel aus dem gesprochenen Wort *Baum*, kann das mentale Bild eines Baumes erzeugt werden. Im nächsten Schritt folgt die Organisation und Umwandlung der selektierten Repräsentationen im Arbeitsgedächtnis zu einem kohärenten mentalen Modell (Mayer, 2001). Dabei wird aus auditiv-verbalen Repräsentationen ein verbales Modell und aus visuell-piktoralen Repräsentationen ein piktorales Modell erstellt, wobei an dieser Stelle angenommen wird, dass keine Informationen mehr zwischen den Modellen ausgetauscht werden³⁷. Abschließend kommt es zur Integration der Informationen, bei der Verknüpfungen zwischen den beiden mentalen Modellen hergestellt werden und durch den Einbezug des Vorwissens ein einheitliches mentales Modell erstellt wird (z. B. Clark & Mayer, 2008). Basierend auf der *CTML* wurden verschiedene Prinzipien für die Gestaltung³⁸ multimedialer Materialien entwickelt. Die Umsetzung

³⁷ Ein neben der *CTML* verbreitetes Modell im Bereich des Lernens mit Multimedia ist das *integrative Modell des Text- und Bildverständnisses* von Schnotz (2005). Da dieses Modell sich inhaltlich stark mit der *CTML* überschneidet, soll es in dieser Arbeit nicht gesondert betrachtet werden. Allerdings wird an dieser Stelle auf einen Unterschied zwischen beiden Modellen hingewiesen. Nach Schnotz (2005) wird angenommen, dass Transformationsprozesse nicht, wie nach der *CTML* postuliert, auf der ersten Repräsentationsebene stattfinden, sondern auf der darauffolgenden.

³⁸ Auf die für diese Arbeit relevanten Gestaltungsprinzipien wird in **Kapitel 4** eingegangen.

dieser Prinzipien soll die kognitive Belastung der Lernenden in einem optimalen Bereich halten und somit zur Schaffung bestmöglicher Lernbedingungen beitragen³⁹.

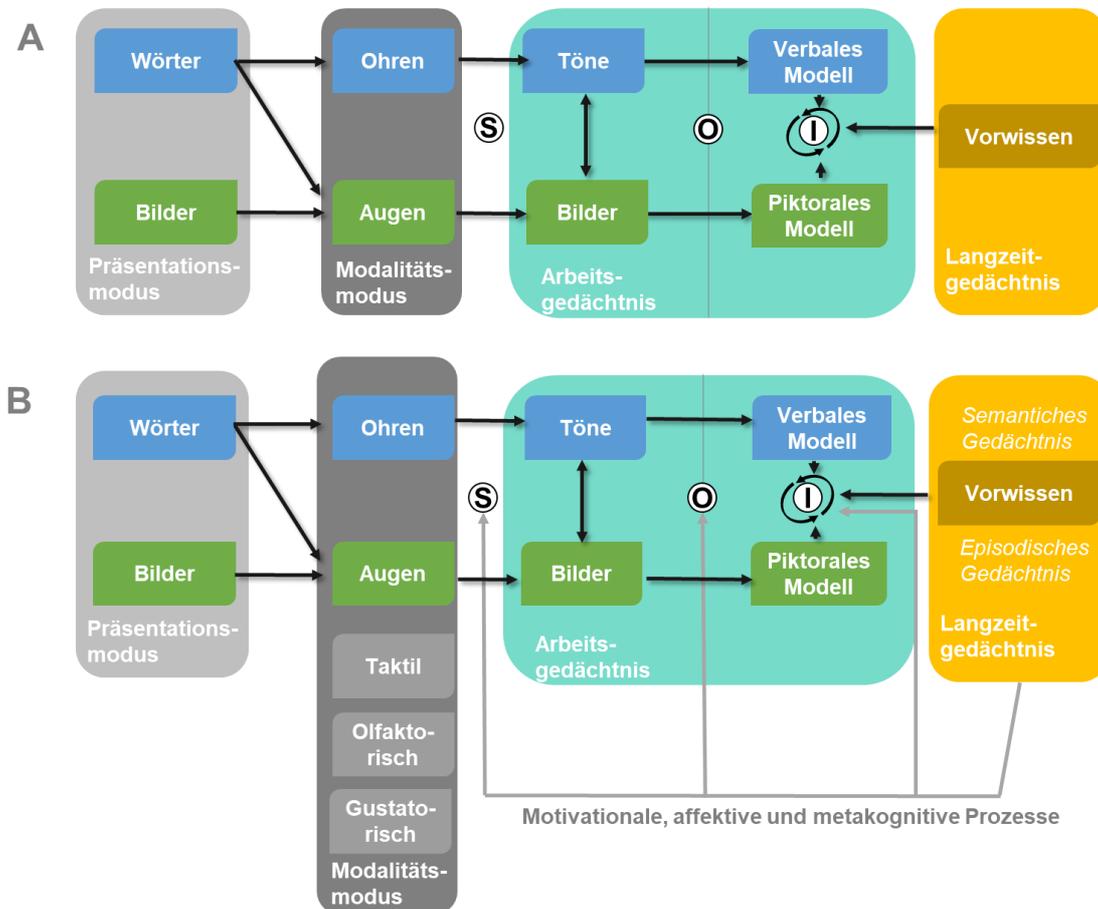


Abbildung 6. Zentrale Modelle für das Lernen mit Multimedia. **A)** Das Modell der *Cognitive Theory of Multimedia Learning* (Mayer, 2005). Diese abgewandelte Darstellung basiert auf Rey (2009, S. 52) und dem *Selektion-Organisation-Integrations-Modell (SOI-Modell)*; Clark & Mayer, 2008, S.35). **B)** Das Modell der *Cognitive-Affective Theory of Learning with Media* nach Moreno (2006), angelehnt an Oberfoell und Correia (2016).

3.4 Cognitive-Affective Theory of Learning with Media

Emotionale und motivationale Aspekte sowie kognitive Prozesse sind nach aktuellem Verständnis untrennbar miteinander verbunden (Plass & Kalyuga, 2019; siehe **Kapitel 4.3**). Da die *CTML* auf kognitive Prozesse fokussiert und diese Aspekte nicht berücksichtigt, wurde sie im Rahmen der *Cognitive-Affective Theory of Learning with Media*

³⁹ In den letzten Jahrzehnten wurden zahlreiche empirische Studien und Metaanalysen zur Untersuchung der *CTML* und von ihr abgeleiteten Prinzipien durchgeführt (z. B. Noetel et al., 2021). Auf diese wird in dieser Arbeit nicht explizit eingegangen.

(*CATLM*)⁴⁰ um motivationale und affektive Aspekte erweitert (Moreno, 2006; Moreno & Mayer, 2005)⁴¹. Diese Aspekte lassen sich den folgenden vier Annahmen (*Assumptions*) zuordnen. Nach der *Subdivision of Long-term Memory Assumption* wird eine Zweiteilung des deklarativen Langzeitgedächtnisses in das semantische und das episodische Gedächtnis angenommen⁴². Semantische Gedächtnisinhalte beinhalten Faktenwissen und episodische Gedächtnisinhalte umfassen autobiographisches Wissen (Buchner & Brandt, 2017). Die *Motivational Factors Assumption* schreibt motivationalen Faktoren zu das Lernen direkt oder indirekt beeinflussen zu können, was sich förderlich oder hinderlich auf Lernprozesse auswirken kann (Pintrich, 2003). Im Rahmen der *Metacognitive Factors Assumption* wird dem Wissen über und die Kontrolle der eigenen kognitiven Prozesse (metakognitiven Faktoren) ein mittelbarer Einfluss auf den Lernerfolg zugeschrieben (basierend auf z. B. McGuinness, 1990)⁴³. Hierbei wird angenommen, dass metakognitive Faktoren die kognitive Informationsverarbeitung, die Motivation sowie die Emotionen regulieren können. Ein Beispiel für einen metakognitiven Faktor ist das Wissen darüber, wie und wann man eine bestimmte Lernstrategie gewinnbringend einsetzen kann und wann ihr Einsatz nicht sinnvoll ist. Zusätzlich wird im Rahmen der *CATLM* auch die moderierende Funktion von Lerncharakteristika (z. B. das Vorwissen und bestimmte Fähigkeiten) thematisiert (*Learner Characteristics Assumption*). Es wird vermutet, dass diese Aspekte auf den Lernprozess einwirken, indem sie die Verfügbarkeit von kognitiven Ressourcen beeinflussen (Kalyuga et al., 2003). Neben der Berücksichtigung der genannten Annahmen werden in der *CATLM* zusätzlich zu den in der *CTML* aufgeführten Sinneseindrücke (visuell und auditiv) auch taktile, olfaktorische und gustatorische Eindrücke mit einbezogen und ein Einfluss auf den Lernprozess zugesprochen (siehe **Abbildung 6 B**).

⁴⁰ Im Folgenden wird für die *Cognitive-Affective Theory of Learning with Media* die Abkürzung *CATLM* verwendet.

⁴¹ Neben der *CATLM*, existieren auch noch weitere Modelle, welche affektive Prozesse beim Lernen mit Multimedia berücksichtigen. Etwa das *Integrated Cognitive Affective Model of Learning with Multimedia* (*ICALM*; Plass & Kaplan, 2016), in welchem die reziproke Wirkung von Lernen mit Multimedia und Emotionen mit aufgenommen wurde. Kürzlich wurde mit der *Cognitive-Affective-Social Theory of Learning in digital Environments* (*CASTLE*) eine Weiterentwicklung veröffentlicht, welche neben der affektiven Komponente auch soziale Aspekte des Lernens berücksichtigt (Schneider et al., 2021). Hier wird postuliert, dass die kognitive Verarbeitung von Informationen beim Lernen mit Multimedia durch soziale Prozesse mediiert wird. Diese Theorie bezieht sich insbesondere auf digitales Lernmaterial und wurde deshalb im Rahmen dieser Arbeit nicht berücksichtigt.

⁴² Das semantische und das episodische Gedächtnis sind Teil des deklarativen Langzeitgedächtnisses (siehe **Kapitel 1.1** für nähere Informationen zum deklarativen Gedächtnis).

⁴³ Die Rolle der Metakognition wird im Rahmen der *CATLM* nicht weiter ausgeführt. Mehr Informationen zu metakognitiven Aspekten und deren Einfluss auf den Lernerfolg bieten beispielsweise von der Linden (2008) und Wendt (2021).

4 Emotionen

4.1 Gefühl, Affekt, Emotion und Stimmung - Ein Abgrenzungsversuch

Gefühl, Affekt, Emotionen und Stimmung sind verschiedene Begriffe zur Beschreibung von Emotionalität (Scherer, 2005)⁴⁴. Die Bedeutung dieser Begriffe ist eng miteinander verwandt und weist Parallelen auf, was dazu führt, dass sie aus Ermangelung einer empirisch fundierten Differenzierung, synonym verwendet werden (Beedie et al., 2011). Dennoch gibt es Bestrebungen, diese Begriffe voneinander abzugrenzen. Im Folgenden werden die Abgrenzungsversuche für den deutschsprachigen Raum betrachtet. Hierfür werden vor allem die Eigenschaften Dauer und Intensität als Abgrenzungskriterien herangezogen (Ekman & Davidson, 1994, Morris & Schnurr, 1989).

Gefühle können als geistige Zustände oder Bewusstseinsphänomene verstanden werden (Demmerling, 2013). Anders als Gedanken, stellen Gefühle eine Art und Weise dar, sich gegenüber und in seiner Umwelt zu verhalten (ebd.)⁴⁵. Nach Mees (2006) lassen sich Gefühle in affektive und nicht-affektive Gefühle unterteilen. Der Begriff Affekt bezieht sich in diesem Kontext auf *die Repräsentation einer Bewertung* (z. B. gut oder schlecht bzw. angenehm oder unangenehm; Mees, 2006).

Abgesehen von dieser Verwendung des Begriffs Affekt, existiert in der psychologischen Emotionsforschung eine weitere gängige Definition, welche sich mit dem umgangssprachlichen Gebrauch dieses Begriffs deckt. Hier wird jeder emotionsgeladene Zustand, welcher durch äußere Reize verursacht wird, als Affekt bezeichnet (Davidson et al., 2009). Diese Zustände können mit einer starken Handlungstendenz oder dem Verlust der Handlungskontrolle (umgangssprachlich: *die Handlung im Affekt*) einhergehen (Otto, Euler & Mandl, 2000). Im angloamerikanischen Sprachgebrauch wird Affekt als Oberbegriff für Emotionen und Stimmungen verwendet. Da dieses Verständnis des Begriffs international anerkannt ist und auch in den für diese Arbeit relevanten Publikationen so

⁴⁴ Diese Arbeit erhebt nicht den Anspruch den aktuellen Stand der theoretischen und empirischen Fundierung der Definition und Abgrenzung von Emotionen in seiner Gänze wiederzugeben. An dieser Stelle wird auf einschlägige Publikationen verwiesen (Elster, 2009; Holland, 2007; Murphy et al., 2003; Watson et al., 1988).

⁴⁵ Eine andere Definition für Gefühle beschreibt diese als Wahrnehmungen des Geisteszustands, welche stattfinden, während Emotionen erlebt werden (Damasio, 2012).

vorgegangen wurde, ist die vorliegende Arbeit an diesen Sprachgebrauch angelehnt. Die Begriffe Emotionen und affektive Zustände werden daher synonym verwendet.

Nach Mees (2006) können affektive und nicht-affektive Gefühle situationsbezogen bzw. aktuell (situativ) oder veranlagt bzw. dauerhaft (dispositional) sein (siehe **Abbildung 7** für eine Übersicht). Situative, affektive Gefühle umfassen Emotionen, Stimmungen (z. B. Gereiztheit oder Niedergeschlagenheit) und Empfindungen bzw. Körpergefühle (z. B. Hunger oder Kälteempfinden). Unter den Aspekt der dispositionalen, affektiven Gefühle fallen Persönlichkeitseigenschaften bzw. das Temperament einer Person. Beispiele hierfür wären Jähzorn oder Ängstlichkeit. Zu den situativen nicht-affektiven Gefühlen werden Eindrücke bzw. Ahnungen gezählt, welche zum Beispiel Gewissheit oder Vertrautheit umfassen. Dispositionale nicht-affektive Gefühle können den Charaktereigenschaften (z. B. Verantwortungsgefühl oder Pflichtgefühl) oder den Fähigkeiten (z. B. Takt- oder Ballgefühl) einer Person zugeordnet werden.

Unter Emotionen werden situative Gefühle mit affektiven Komponenten, die von kurzer Dauer sind, verstanden (Mees, 2006). Allerdings handelt es sich bei Emotionen um ein „derartig vielschichtiges und schillerndes Konzept“, das bislang keine allgemein anerkannte Definition ermöglichte (Brandstätter et al., 2018, S. 164). Als konsensfähig erscheint jedoch die Betrachtung von Emotionen als „ein komplexes Muster körperlicher und mentaler Veränderungen [...] als Antwort auf eine Situation, die als persönlich bedeutsam wahrgenommen wurde“ (Zimbardo & Gerrig, 2008, S. 454). Nach Russell und Barrett (1999) handelt es sich um Emotionen, wenn Reaktionen auf einen Reiz kurzfristig, unwillkürlich, situativ und objektgerichtet ausfallen. Man hat beispielsweise Angst vor etwas oder jemandem und freut sich *über* etwas, *mit* jemandem oder *auf* jemanden.

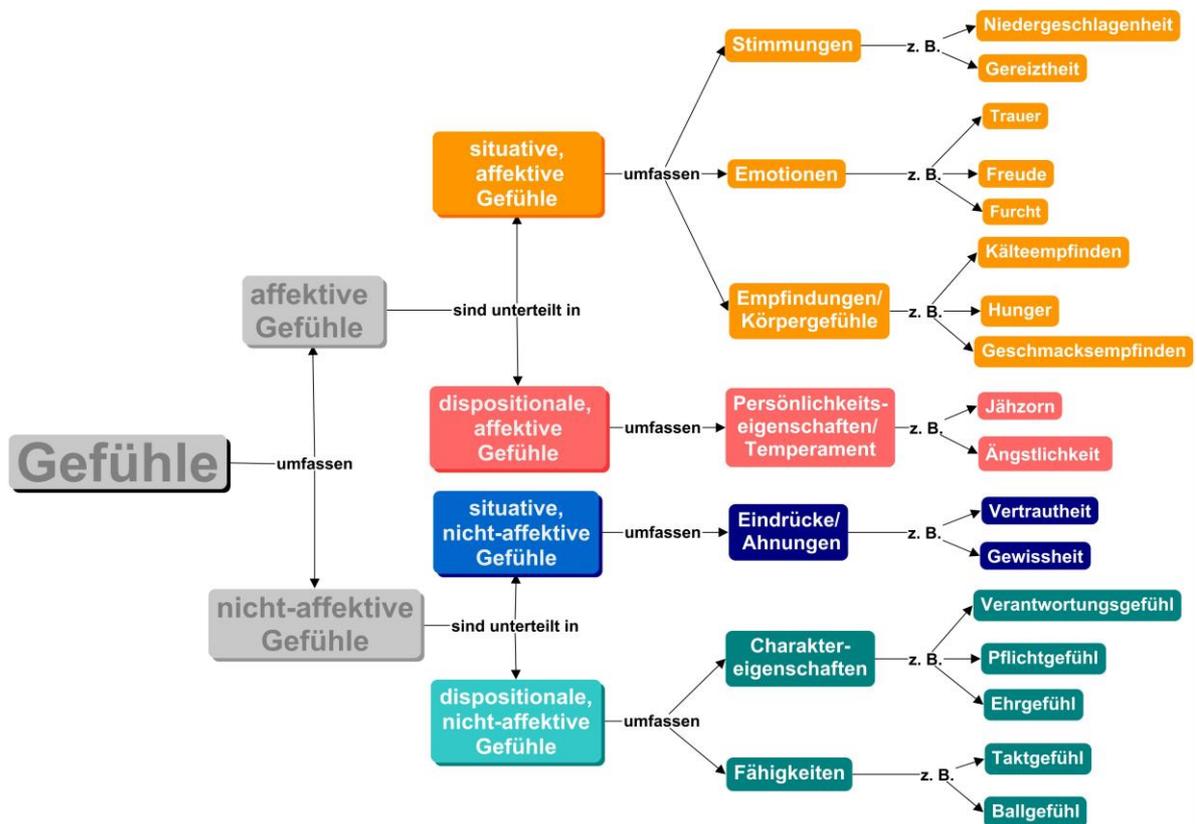


Abbildung 7. Concept Map zur Visualisierung affektiver und nicht-affektiver Gefühle (diese Abbildung stellt eine Modifikation der Darstellung von Mees [2006, S. 105] dar).

Unter Stimmung wird eine wenig intensive, eher diffuse Befindlichkeit verstanden, welche über einen größeren Zeitraum unverändert bleibt und nicht an Ereignisse oder Personen geknüpft ist (Schwarz, 1987). Diese Stimmung ist zeitlich weniger begrenzt als Emotionen. Sie bietet eine Art Bezugsrahmen, welcher den „Hintergrund [darstellt], vor dem sich Denkprozesse, aber auch Emotionen abspielen“ (Brandstätter et al., 2018, S. 164). Die Stimmung wird hauptsächlich von dem Grad der Bedürfnisbefriedigung (Wulf, 2017), aber auch durch den Grad der Übereinstimmung des aktuell wahrgenommenen Selbst im Vergleich zum Ideal- bzw. Soll-Selbst bestimmt (Brandstätter et al., 2018). Beispielsweise ist Niedergeschlagenheit eine Stimmung, die sich einstellt, wenn die Diskrepanz zwischen dem aktuell wahrgenommenen Selbst und dem Idealselbst hoch ist (ebd.).

4.2 Die Klassifizierung/Taxonomie von Emotionen

Es existieren verschiedene theoretische Perspektiven, nach denen Emotionen unterschiedlich klassifiziert werden können (Gross & Barrett, 2011). Zu den bekanntesten Theorien zählen die Basisemotionstheorie, die Bewertungstheorie, psychologisch konstruktivistische Theorien und soziale konstruktivistische Theorien (siehe z. B. Pekrun [2005] für detaillierte Informationen zu den einzelnen Theorien). Nach der Basisemotionstheorie wird angenommen, dass unterschiedliche Emotionen angeboren sind oder sich in der frühen Kindheit entwickeln und dass diese differenzierbar sind (Gross & Barrett, 2011). Uneinigkeit herrscht über die Anzahl an unterscheidbaren Basisemotionen. Ekman et al. (1971) unterscheiden zwischen sechs Emotionen (Ärger, Furcht, Freude, Überraschung, Traurigkeit und Ekel), während nach Ekman (1992) nur Ärger, Freude, Trauer und Furcht bzw. Angst zu den Basisemotionen gezählt werden. Im Rahmen der Bewertungstheorie werden Emotionen als Verhaltensweisen verstanden, welche auf einen bestimmten Reiz folgen (Scherer, 2005). Hierbei steht die Einschätzung der Relevanz eines Reizes für die Entstehung von Emotionen im Mittelpunkt.

Psychologisch konstruktivistische Theorien gehen davon aus, dass der Gemütszustand einer Person von vielen unterschiedlichen Faktoren bestimmt wird (Gross & Barrett, 2011). Der soziale Konstruktivismus betrachtet Emotionen als Produkt sozialer Interaktionen (Scherer, 2002). Laut dieser Annahme werden nicht durch bestimmte Situationen bestimmte Emotionen ausgelöst, vielmehr lösen Situationen, welche als bedeutsam für die eigenen Ziele oder das Wohlbefinden eingestuft werden, Emotionen aus (Frijda, 1986).

Basierend auf den verschiedenen Theorien, existieren unterschiedliche Forschungsansätze, in denen das jeweilige Emotionsverständnis berücksichtigt wird. Hierbei werden Emotionen unterschiedlich gruppiert bzw. klassifiziert, wobei zwischen kategorialen bzw. diskreten Modellen und dimensionalen Modellen unterschieden wird (Frenzel et al., 2009). Diskrete Emotionsmodelle basieren meist auf Emotionstheorien, welche sich in unterschiedliche Basisemotionen unterteilen (Plutchik, 1980). Anders als diskrete Modelle, wurden dimensionale Modelle nicht von Emotionstheorien abgeleitet. Vielmehr basieren sie auf der empirischen Forschung (Watson et al., 1988). Hier wird davon ausgegangen, dass sich alle Emotionen anhand weniger Dimensionen beschreiben lassen, wo

ihnen eine Position in einem Emotionsraum zugeteilt wird (ebd.). Dies ermöglicht eine konkrete Verortung und in Bezug auf die Anzahl eine uneingeschränkte Einordnung.

Ein bekanntes dimensionales Modell ist das *Circumplex Model of Affect*⁴⁶ von Russell (1980). Nach diesem Modell können Emotionen in einem zweidimensionalen System kategorisiert werden (siehe **Abbildung 8**). Dabei bildet die Valenz als Kontinuum von unangenehm bis angenehm eine Merkmalsdimension und der Grad der Erregung als Kontinuum von deaktiviert bis aktiviert die andere Dimension. Hieraus ergeben sich vier Felder für angenehm-deaktivierende, angenehm-aktivierende, unangenehm-aktivierende und unangenehm-deaktivierende Zustände (Pekrun, 2005, 2018; Pekrun et al., 2006). Die Emotionen werden diesen Feldern zugeordnet und in einem Kreis um einen neutralen Mittelpunkt platziert, wobei die Nähe der Emotionen zueinander die Ähnlichkeit der emotionalen Verfassung widerspiegelt. Beispiele für aktivierende Emotionen sind Fröhlichkeit oder Überraschung (angenehm) und Angst (unangenehm). Deaktivierende Emotionen sind beispielsweise Zufriedenheit (angenehm) und Traurigkeit (unangenehm).

⁴⁶ Das Modell von Russell (1980, 2003.) wird im Forschungskontext kritisiert, da subjektiv sehr unterschiedlich empfundene Emotionen (z. B. Zorn und Angst) auf ähnlichen Positionen innerhalb der beiden Dimensionen verortet werden (Schmidt-Atzert, 1996). Auch wird der Einbezug weiterer Dimensionen wie beispielsweise die erlebte Dominanz empfohlen (Frenzel et al., 2009). Im Rahmen dieser Arbeit wird sich dennoch an diesem Modell orientiert, da eine exakte Abbildung bzw. Einordnung von bestimmten Emotionen nicht den Gegenstand des Forschungsinteresses darstellt. Auch die für diese Arbeit relevanten Forschungsarbeiten und theoretischen Fundierungen nutzen Russells Modell oder darauf aufbauende Modelle.

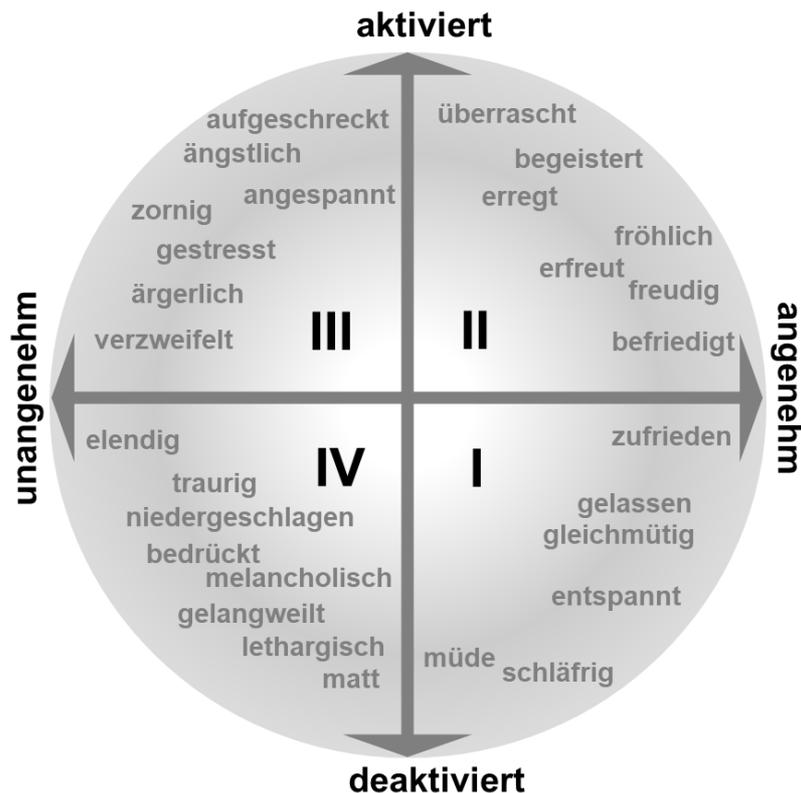


Abbildung 8. Dimensionale Kategorisierung affektiver Zustände. Dargestellt sind die Dimensionen Valenz (angenehm bis unangenehm) und Erregung (aktiviert bis deaktiviert). In den vier Feldern befinden sich angenehm-deaktivierende (I), angenehm-aktivierende (II), unangenehm-aktivierende (III) und unangenehm-deaktivierende (IV) Zustände. Die angegebenen Emotionen stellen, der Übersichtlichkeit geschuldet, lediglich eine Auswahl von emotionalen Zuständen dar. Bei der Abbildung handelt es sich um eine Modifikation des Kreismodells nach Feldman Barrett et al. [1999].

4.3 Emotionen, kognitive Prozesse und Lernen

Bezüglich der möglichen Auswirkungen von Emotionen auf das Lernen existieren zwei gegensätzliche Annahmen: die *Emotions-as-facilitator-of-learning-hypothesis* und die *Emotions-as-suppressor-of-learning-Hypothesis* (Um et al., 2012)⁴⁷.

Die *Emotions-as-facilitator-of-learning-hypothesis* besagt, dass das Erleben von Emotionen die Lernleistung verbessern kann. Nach Um et al. (2012), kann das Erleben positiver Emotionen während des Lernprozesses, entweder durch direkte Einwirkung oder durch vermittelnde Variablen, die Lernleistung verbessern. Eine vermittelnde Variable kann hierbei die Motivation darstellen. Für positive Emotionen und intrinsische Motivation

⁴⁷ Im Rahmen dieser Arbeit erfolgt nur eine stark verkürzte und vereinfachte Darstellung der möglichen Auswirkungen von Emotionen auf das Lernen. Differenziertere Ansätze zu diesem Thema bieten beispielsweise Goetz et al. (2003) und Pekrun (2014, 2018).

wird ein positiver Zusammenhang angenommen (Pekrun, 2006). Die entscheidende Rolle von Motivation im Rahmen von Lernprozessen wird zum Beispiel in der *Cognitive-Affective Theory of Learning with Media* betont (z. B. Moreno, 2006). Hier wird angenommen, dass die Motivation das Lernen beeinflusst, indem sie Lernende dazu bewegt ihr kognitives Engagement zu erhöhen oder zu reduzieren (vgl. Pekrun, 2006).

Bei emotionalen Reizen wird angenommen, dass das kognitive System über einen Wahrnehmungsmechanismus verfügt, der darauf spezialisiert ist, emotionale Reize von anderen Reizen zu unterscheiden und bevorzugt zu verarbeiten (Nummenmaa et al., 2006). Es wird vermutet, dass die Amygdala in der Lage ist auf emotionale Reize zu reagieren noch bevor einer Person diese Reize bewusst sind (Whalen et al., 1998). Dieser Effekt ist als *Emotion-related Attentional Bias* bekannt und wurde in zahlreichen Studien beobachtet (z. B. Harrison et al., 2010; Lea et al., 2018).

Darüber hinaus wird für emotionale Reize postuliert, dass sie die kognitiven Ressourcen erhöhen können (Isen, 1987), die für eine Aufgabe zur Verfügung gestellt werden und Lernende diese Ressourcen zudem effektiver nutzen können (Bless et al., 1996)⁴⁸. Gestützt wird diese Vermutung durch die Beobachtung, dass sich ein positiver Affekt dann einstellt, wenn die Bedürfnisse einer Person erfüllt sind. Dies ist ein Signal dafür, dass sie sich anderen Zielen zuwenden kann (Carver, 2003). Wenn die Grundbedürfnisse erfüllt sind, kann sich die Person zum Beispiel dem Lernen zuwenden (Fredrickson, 2001; Isen et al., 1987). Zudem wird angenommen, dass Emotionen Auswirkungen auf das Gedächtnis in Form der Abrufbarkeit von Informationen haben. Hier wird postuliert, dass positive Affekte als Abrufhinweise (*Retrieval Cues*) dienen und somit dazu beitragen, dass diese Informationen besser aus dem Langzeitgedächtnis abgerufen werden können (ebd.).

Im Gegensatz dazu wird im Rahmen der *Emotions-as-suppressor-of-learning-hypothesis* von einer lernbeeinträchtigenden Wirkung ausgegangen. Hier wird angenommen, dass das Erleben von Emotionen kognitive Ressourcen bindet, welche demnach nicht dem Lernprozess zur Verfügung stehen (Shuman & Scherer, 2014). Es besteht zudem die Annahme, dass durch Emotionen verstärkt lernirrelevante kognitive Prozesse hervorgerufen

⁴⁸ Diese Arbeit deckt nur Teilaspekte der komplexen Wirkung ab, welche für Emotionen auf kognitive Prozesse angenommen wird. In dieser Arbeit wird zudem der Fokus auf positive Emotionen gelegt. Plass und Kalyuga (2019) bieten einen detaillierten Überblick zu diesem Themenbereich.

werden (*Incidental Processing*; Mayer & Moreno, 2003). Ein Beispiel für *Incidental Processing* ist das Nachdenken über den eigenen emotionalen Zustand.

4.4 Die empirische Befundlage zu Emotionen und Lernerfolg

Verbreitet ist die Annahme, dass positive Emotionen typischerweise das Interesse und die Lernmotivation steigern, während negative Affekte hauptsächlich das Lernen beeinträchtigen (Gillet et al., 2013; Isen et al., 1987)⁴⁹. Aktuelle Forschung, welche sich mit der Wirkung von Emotionen auf Lernprozesse befasst, folgert jedoch, dass dichotome Einteilungen von Emotionen keine eindeutige Aussage bezüglich der Lernförderlichkeit von Emotionen zulassen (Pekrun, 2017; Shuman & Scherer, 2014). Hier wurde für die Dimension Valenz sowohl für positive als auch für einige negative Emotionen eine positive Wirkung auf das Lernen beobachtet (Münchow & Bannert, 2019). Für die positive Emotion Zufriedenheit konnte beispielsweise keine Lernförderlichkeit gefunden werden, wohingegen Neugier als lernförderlich eingestuft wurde (Pekrun, 2017). Für die negative Emotion Angst wurden sowohl positive als auch negative Effekte auf die Lernleistung gefunden, wohingegen Hoffnungslosigkeit meist mit einer lernhemmenden Wirkung assoziiert wurde (Pekrun et al., 2006; Pekrun & Stephens, 2012; Pekrun et al., 2017).

Auch die reine Betrachtung der Erregungsdimension (aktivierend, deaktivierend) ergibt kein eindeutiges Bild. Aktivierende Emotionen können lernförderlich sein, was zum Beispiel für Prüfungsangst, Frust über eigene Fehler und Freude am Lernen gezeigt werden konnte (z. B. Oser & Spychiger, 2005). Diese Emotionen können jedoch auch lernhinderlich wirken, indem sie vom Lernstoff ablenken (Pekrun, 2006). Angst kann zwar die Anstrengungsbereitschaft erhöhen (Gage & Berliner, 1996), es wurde jedoch auch beobachtet, dass Angst nicht nur aktivierend, sondern situations- und personenabhängig auch deaktivierend wirken kann (Heckhause & Heckhause, 2006).

Für die deaktivierende Emotion Erleichterung wurde eine hemmende Wirkung auf den sofortigen Lernerfolg beobachtet. Allerdings wird vermutet, dass Erleichterung für die

⁴⁹ Aufgrund der Fülle an Forschung, welche sich mit dem Zusammenhang zwischen Emotionen, kognitiven Prozessen und Lernen befassen, kann hier kein vollständiger Überblick gegeben werden. Anhand einer (subjektiv getroffenen) Auswahl werden stattdessen einige wichtige Ergebnisse des Forschungsfelds skizziert.

Aufrechterhaltung der Langzeitmotivation förderlich sein kann. Die deaktivierende Emotion Langeweile steht hingegen hauptsächlich im Zusammenhang mit einer Beeinträchtigung des Lernens (z. B. Aspinwall, 1998).

Basierend auf diesen Befunden wurde die Berücksichtigung beider Dimensionen (Valenz und Aktivierung), also der vier Zustände positiv-aktivierend, positiv-deaktivierend, negativ-aktivierend und negativ-deaktivierend, der Lernförderlichkeit von Emotionen vorgeschlagen (Pekrun, 2017). Diese Einbeziehung schafft ein etwas klareres Bild im Hinblick auf die empirische Befundlage. Positiv-aktivierende Emotionen sind demnach größtenteils lernförderlich, während negativ-deaktivierenden Emotionen eindeutig eine lernhinderliche Wirkung zugeschrieben werden kann (z. B. Pekrun, 2006; Pekrun et al., 2011; Pekrun, 2017). Die Wirkung positiv-deaktivierender und negativ-aktivierender Emotionen ist auch bei dieser Herangehensweise unklar, sie kann sowohl positiv als auch negativ ausfallen (ebd.).

Hierzu merkt Pekrun (2021) an, dass die Art der Wirkung (vorteilhaft oder hinderlich) verschiedener Emotionen beispielsweise von der Situation, der Art der Aufgabe und persönlichen Eigenschaften abhängt. Es konnte etwa gezeigt werden, dass positive Emotionen bei der Bewältigung von Aufgaben, welche kreatives und flexibles Denken erfordern, förderlich sein können, während negative Emotionen bei der sorgfältigen und systematischen Bearbeitung einer Aufgabe dienlich erscheinen (Bless & Fiedler, 1999; Pekrun & Stephens, 2012). Auch der Grad der Aktivierung scheint eine wichtige Rolle für den Lernerfolg zu spielen (Schneider et al., 2019). Um gute Leistungen zu erzielen, braucht es demnach einen gewissen Grad der Aktivierung. Wird dieser Grad überschritten, wird der Lernprozess behindert (ebd.). Der Aktivierungsgrad, der für ein optimales Lernen notwendig ist, kann von Person zu Person und abhängig vom Kontext variieren (ebd.). Erklärt wird dies dadurch, dass eine optimale Aktivierung ein hohes Maß an Aufmerksamkeit für die zu bewältigende Aufgabe ermöglicht, während eine zu geringe Aktivierung wenig kognitive Ressourcen mobilisiert. Bei zu hoher Aktivierung werden die mobilisierten Ressourcen allerdings auf das Befinden des eigenen Körpers gelenkt (z. B. auf den Atem oder den Herzschlag; ebd.).

Knörzer et al. (2016) vermuten, dass moderierende Merkmalsvariablen für die unterschiedlichen Befunde verantwortlich sind. Was die Gewichtung der emotionalen Erfahrungen während des Lernprozesses betrifft, nehmen Feller-Länzlinger et al. (2002) an,

dass dieser erfolgreicher ist, wenn die positiven Erfahrungen und Emotionen überwiegen. Negative Emotionen können jedoch in der richtigen Dosierung einen entscheidenden Beitrag leisten. Für das Erleben mehrerer Emotionen gleichzeitig (hier Freude, Wut, Verwirrung und Frustration) wurde beobachtet, dass dieser Zustand vom Lerngegenstand ablenken und den Lernprozess negativ beeinflussen kann (Dever et al., 2021). Die genannten Beispiele verdeutlichen die Komplexität des Zusammenhangs zwischen Emotionen und Lernen.

5 Emotional Design

Ein recht junges Forschungsfeld widmet sich der Entwicklung und Untersuchung affektiver Gestaltungsprinzipien⁵⁰ (*Emotional Design*), die potenziell den affektiv-motivationalen Zustand der Lernenden beeinflussen können (Plass & Kalyuga, 2019). Ziel der *Emotional Design*-Forschung ist es, durch die Beeinflussung von Emotionen, kognitive Prozesse zu verbessern und dadurch das Lernen zu fördern (Plass & Kaplan, 2016).

Im Rahmen der Gestaltungsprinzipien welche Um et al. (2012) erstmals für das *Emotional Design* beim Lernen mit Multimedia definierten, wird die Manipulation der Farbe und der Form empfohlen. Sie umfassen die Verwendung warmer, heller Farben sowie runder, anthropomorphisierter (vermenschlichter) Formelemente (siehe Beispiel in **Abbildung 9 A**)⁵¹. Die Illustration von einem Virus kann beispielsweise anthropomorphisiert werden, indem Merkmale wie Augen, Mund oder Gliedmaßen hinzugefügt werden (z. B. Mayer & Estrella, 2014). Die vorgestellten Gestaltungsprinzipien wurden aus theoretischen Überlegungen der *Cognitive-Affective Theory of Learning with Media* abgeleitet und orientieren sich an Befunden aus vorherigen Studien. In diesen Studien wurde in unterschiedlichen Domänen und Kontexten die Wirkung bestimmter Designmerkmale auf Emotionen untersucht. Zum Beispiel wurde im Rahmen der Forschung zu Multimedia-Designs herausgefunden, dass unterschiedliche Gestaltungen den affektiven Zustand der Betrachter:innen beeinflussen (Capota et al., 2007). Bestimmte Farben können einen

⁵⁰ Die Begriffe Design- und Gestaltungsempfehlungen werden in der vorliegenden Arbeit synonym verwendet.

⁵¹ In dieser Arbeit wird sich ausschließlich auf das emotionale Bild-Design bezogen, da sich der Großteil der bisherigen Forschung dieser Art der Gestaltung widmet und diese Art der Gestaltung auch Gegenstand der durchgeführten Studien ist. Neben dem emotionalen Bild-Design ist auch ein emotionales Text-Design möglich. Dieses umfasst eine emotionale Gestaltung relevanter Text-Informationen und wurde bislang noch wenig untersucht (Stark, 2018).

Einfluss auf Emotionen haben, wobei warme Farben eine stärkere emotionale Wirkung hervorrufen als kalte (Wolfson & Case, 2000). Helle, satte Farben (z. B. Gelb oder Orange) lösen eher positive Emotionen aus, als Farben wie Schwarz, Grau, Braun und Weiß (Palmer & Schloss, 2010).

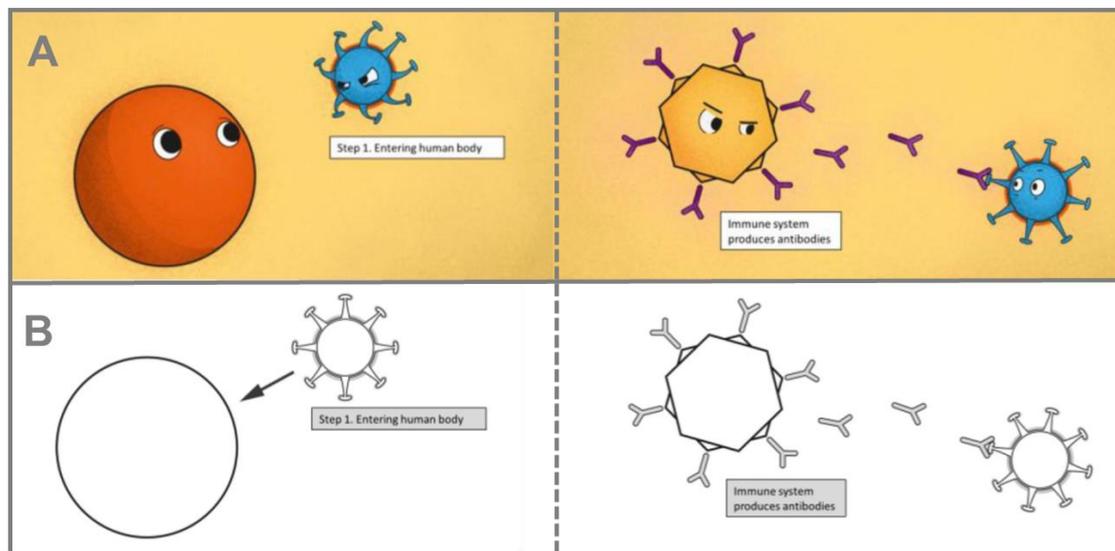


Abbildung 9. Lernmaterial zum Thema Immunologie (Ausschnitt entnommen aus Brom et al., 2018). Es sind zwei Schritte des Infektionsprozesses durch das Influenzavirus zu sehen. **A:** Lernmaterial im *Emotional Design*. Der Hintergrund ist in warmen, hellen Farben gestaltet und lernrelevante Elemente wurden anthropomorphisiert. **B:** Nicht-emotional gestaltetes Lernmaterial. Es wurde eine graustufige Färbung verwendet und lernrelevante Elemente wurden nicht anthropomorphisiert.

Ein weiteres Merkmal, welches positive Emotionen hervorruft, sind kindliche Gesichtszüge. Dieser sogenannte *Baby-face-bias* (Lorenz & Martin, 1971) umfasst Gesichtsm Merkmale wie eine runde Gesichtsform, große Augen, eine kleine Nase und ein kurzes Kinn. Mit diesen Merkmalen wird eine kindliche Persönlichkeit assoziiert und es wurde festgestellt, dass sie positive Emotionen induzieren und sowohl Aufmerksamkeit als auch den Fürsorgeinstinkt verstärken können (z. B. Glocker et al., 2009). Diese Befunde werden durch *fMRT*-Studien untermauert. Hier wurde festgestellt, dass Gesichter von Kleinkindern Hirnareale aktivieren können, die für eine erhöhte Aufmerksamkeit verantwortlich sind (Caria et al., 2012). Weitere Erkenntnisse, die die Verwendung von Anthropomorphismen unterstützen, sind zum Beispiel Studien, die zeigen, dass Gesichter schnell und automatisch verarbeitet werden, da dies unbewusst geschieht und wenig Aufmerksamkeit benötigt (Sreenivas et al., 2012). Dies scheint besonders für glückliche Gesichter zu gelten (ebd.).

5.1 Emotional Design, kognitive Prozesse und Lernen

Die folgenden Überlegungen zur Lernwirksamkeit von Materialien im *Emotional Design* basieren auf den eingangs vorgestellten, konkurrierenden Ansichten über die Wirkung von Emotionen auf den Lernprozess und Annahmen der Multimediaforschung.

Bezugnehmend auf die *Emotions-as-suppressor-of-learning-hypothesis* wird postuliert, dass das *Emotional Design* lernhinderlicher ist, als nicht-emotional gestaltetes, multimediales Lernmaterial, da es eine zusätzliche kognitive Belastung auslöst (siehe **Abbildung 10 A und B**). Die Verarbeitung der durch das *Emotional Design* ausgelösten Emotionen wird als nicht relevant für das Lernziel erachtet und als extrinsische kognitive Belastung eingestuft (Plass & Kalyuga, 2019). Eine weitere potenzielle Quelle extrinsischer kognitiver Belastung ist, dass dem Lernmaterial bestimmte Merkmale hinzugefügt werden müssen, um eine emotionale Gestaltung zu erreichen (Um et al., 2012). Diese Annahme stützt sich auf das Kohärenzprinzip (Mayer & Fiorella, 2014) welches besagt, dass Gestaltungselemente, die als interessant und unterhaltsam, aber als inhaltlich irrelevant beschrieben werden (*Seductive Details*), eine extrinsische kognitive Belastung verursachen können (Mayer & Gallini, 1990).

Basierend auf der *Emotions-as-facilitator-of-learning-hypothesis* wird die Lernförderlichkeit des *Emotional Designs* postuliert. Hier wird angenommen, dass das *Emotional Design* von *Seductive Details*, also von irrelevanten, dekorativen Elementen abzugrenzen ist. Beim *Emotional Design* werden lernrelevante Elemente mit emotionalen Eigenschaften ausgestattet (Um et al., 2012), wodurch eine unnötige kognitive Belastung vermieden werden soll. Es wird angenommen, dass das *Emotional Design* Lernende dazu motiviert, sich mit einem Lerninhalt zu befassen, da es Interesse weckt und die ansprechend gestalteten Lerninhalte Aufmerksamkeit auf sich ziehen (Park et al., 2015). Diese aufmerksamkeitslenkende Eigenschaft soll zu einer Reduzierung der kognitiven Belastung beitragen (Le et al., 2021, Park et al., 2015; Peng et al., 2021).

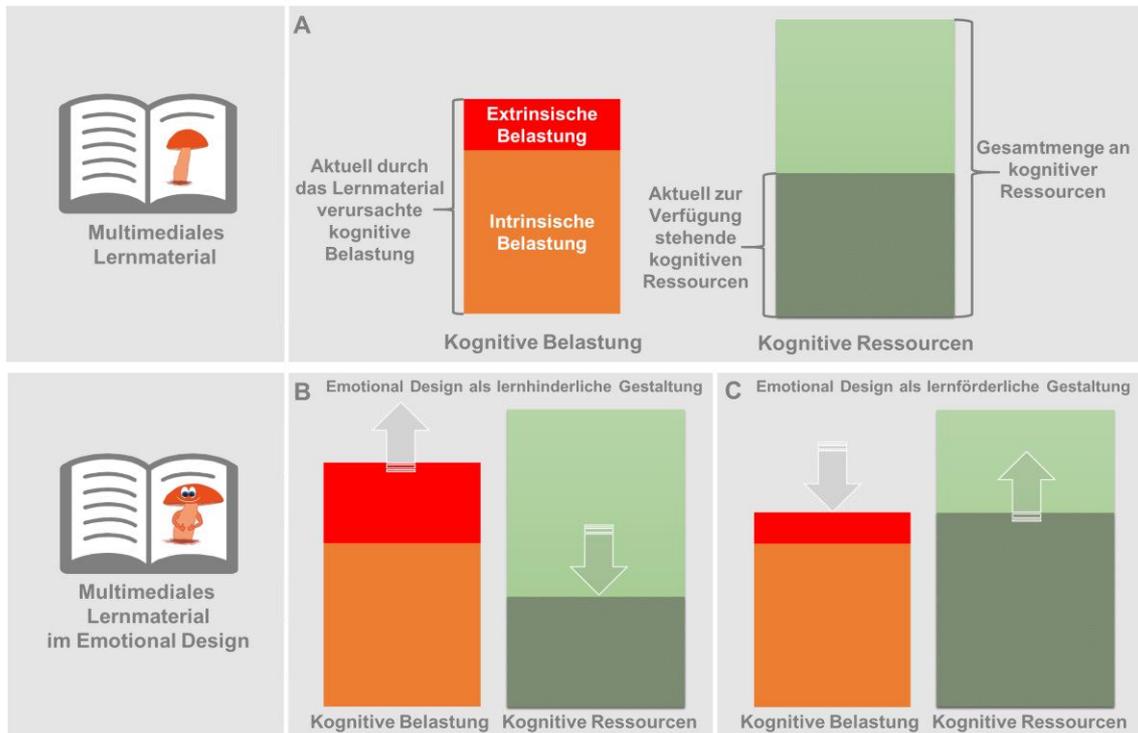


Abbildung 10. Mögliche Wirkung unterschiedlicher multimedialer Gestaltungen auf die kognitive Belastung und die kognitiven Ressourcen (die Abbildung ist angelehnt an die Graphik von Brom et al., [2018, S. 102] und wurde erweitert um Informationen aus Plass und Kalyuga, 2019). **A:** Multimedial gestaltetes Lernmaterial und seine mögliche Wirkung auf die kognitive Belastung und den Verbrauch kognitiver Ressourcen. **B:** Visualisierung einer lernhinderlichen Wirkung eines Lernmaterials im *Emotional Design*. Hier werden eine Erhöhung der extrinsischen kognitiven Belastung (*Extraneous Cognitive Load*) und eine Reduzierung der kognitiven Ressourcen angenommen. **C:** Visualisierung einer lernförderlichen Wirkung eines Lernmaterials im *Emotional Design*. Hier werden eine Reduzierung der extrinsischen kognitiven Belastung und eine Erhöhung der kognitiven Ressourcen vermutet.

5.2 Die empirische Befundlage zum Emotional Design

Mit bunten, comicartigen Zellen und anthropomorphisierten Antigenen untersuchten Um et al. (2012) als Erste die Wirkung von Lernmaterial im *Emotional Design* auf den affektiven Zustand, die Lernleistung und kognitive Prozesse⁵².

Die affektive Wirkung des emotionalen Designs wurde untersucht, indem diese mit der Wirkung externer emotionaler Stimuli (z. B. das Denken an einen glücklichen Moment)

⁵² Die *Emotional Design*-Manipulationen haben zum Ziel positive Emotionen hervorzurufen, weshalb sich nur wenige Studien dem Einfluss des *Emotional Designs* auf negative Emotionen widmen. Somit wird im Rahmen dieses Überblicks nicht auf den Einfluss auf negative Emotionen eingegangen (Informationen hierzu finden sich zum Beispiel in Kumar (2020) und Kumar et al. (2016).

verglichen wurden (*Velten-Technik*; Velten, 1968). Für externe emotionale Stimuli konnte gezeigt werden, dass sie, wenn sie vor dem Lernen angewendet werden, positive Emotionen hervorrufen und das Lernen verbessern können (Bryan et al., 1996). Mittels eines *Emotional Designs* konnten, ähnlich zur externen Emotionsstimulation, positive Emotionen ausgelöst werden (Um et al., 2012). Die Replikationsstudie von Plass et al. (2014) bestätigte diesen Befund. Im Gegensatz dazu kamen Park et al. (2015) zu dem Schluss, dass eine externe Stimmungsinduktion positive Emotionen hervorruft, während das *Emotional Design* keine Auswirkungen auf Emotionen hat.

Der Vergleich von Lernmaterialien, welche im *Emotional Design* gestaltet sind, mit Lernmaterialien welche nicht nach den *Emotional Design*-Prinzipien konzipiert wurden (Verwendung von Graustufen, ohne Anthropomorphisierungen und eckigen Formen; siehe Beispiel in **Abbildung 9 B**), ergab eine emotionsinduzierende Wirkung des *Emotional Designs* (Um et al., 2012). Während einige der nachfolgenden Studien diese Wirkung bestätigten (Bülbul et al., 2021, Mayer & Estrella, 2014; Tien et al., 2018; Um et al., 2012), wurden in anderen keine Emotionen durch das *Emotional Design* hervorgerufen (Münchow & Bannert, 2019; Park et al., 2015, Stárková et al., 2019). Park (2015) beobachtete sogar einen Anstieg der positiven Emotionen für das nicht-emotionale Design, während bei der Gruppe, die mit *Emotional Design* lernte, keine Zunahme positiver Emotionen festgestellt wurde.

Im Hinblick auf die Lernleistung wurde der Einfluss emotionaler Gestaltung auf unterschiedliche Aspekte des Lernens untersucht. Um et al. (2012) fanden heraus, dass das Lernen mit *Emotional Design* die Verstehens- und die Transferleistung verbessert. Auch Plass et al. (2014) und Mayer und Estrella (2014) konnten zeigen, dass *Emotional Design* das Verstehen des Lerninhaltes verbessern kann. Bei beiden Studien wurde allerdings keine Wirkung auf die Transferleistung gefunden. Dafür beobachteten Mayer und Estrella (2014), dass Schüler:innen, die mit emotionalem Design lernten, besser bei der Behaltensleistung abschnitten. Das Ergebnismuster, das aus einer Wirkung auf die sofortigen Lernergebnisse (Behalten), nicht aber auf verzögerte Lernergebnisse (Transfer) besteht, wurde in mehreren Studien beobachtet (Brom et al., 2016; Le et al., 2018; Uzun & Yıldırım, 2018). Während die vorgestellten Studien Effekte auf verschiedene Aspekte des Lernens fanden, konnten viele andere Studien keinen Lerneffekt durch die emotionale

Gestaltung feststellen (Heidig et al., 2015; Münchow et al., 2017; Navratil et al., 2018; Park et al., 2015; Stárková et al., 2019; Uzun & Yildirim, 2018).

Gemessen anhand der mentale Anstrengung (*Mental Effort*) konnte kein Einfluss des *Emotional Designs* auf den *Cognitive Load* gefunden werden (Um et al., 2012). Für externe Emotionsstimulation hingegen wurde eine Erhöhung der mentalen Anstrengung beobachtet (ebd.). Zudem konnte gezeigt werden, dass *Emotional Design* die wahrgenommene Aufgabenschwierigkeit verringern kann (ebd.). Mehrere andere Studien replizierten dieses Ergebnismuster (Le et al., 2018; Mayer & Estrella, 2014; Navratil et al., 2018; Plass et al., 2014). Im Gegensatz dazu berichteten Shangguan et al. (2020) eine Erhöhung der wahrgenommenen Aufgabenschwierigkeit sowie der mentalen Anstrengung für das *Emotional Design* verglichen mit nicht-emotionalem Design. In der Studie von Uzun und Yildirim (2018) investierten Schüler:innen, die mit einem bunten Design (ohne Anthropomorphismen) lernten, mehr mentale Anstrengung im Vergleich zu Schüler:innen, welche Lernmaterial im *Emotional Design* (bestehend aus farbenfrohen Anthropomorphismen und Geräuschen) erhielten.

Zusammenfassend zeigt sich bisher ein heterogenes Bild bezüglich der emotionalen Wirkung des *Emotional Designs*. Einige Ergebnisse unterstützen die Annahme, dass *Emotional Design* Emotionen hervorruft und das Lernen fördert (Brom et al., 2018; Bülbül et al., 2021, Mayer & Estrella, 2014; Navratil, et al., 2018; Um et al., 2012), während andere keine emotionale und lernförderliche Wirkung feststellen konnten (Münchow & Bannert, 2019; Park et al., 2015, Stárková et al., 2019). Auch die Metaanalysen von Brom et al. (2018) und Wong und Adesope (2021) ergaben, dass emotionales Design die Merkfähigkeit, das Verständnis und die Transferleistung verbessern kann. Darüber hinaus wurde eine verringerte wahrgenommene Aufgabenschwierigkeit, ein schwacher Effekt für positive Affekte aber keine signifikante Auswirkung auf die mentale Anstrengung gefunden (ebd.).

Bezüglich der Wirkung unterschiedlicher Gestaltungsmerkmale auf Emotionen zeigte sich, dass die Kombination aus runden, anthropomorphisierten Formen und Farben oder runden anthropomorphisierten Formen allein, effektiver sein kann, als die Verwendung warmer Farben allein (Plass et al., 2014). In der Metaanalyse von Wong und Adesope (2021) zeigte sich ein differenziertes Bild. Hier erwiesen sich verschiedene Gestaltungs-

merkmale für unterschiedliche Aspekte des Lernprozesses als förderlich. Anthropomorphisierte Elemente wirkten sich positiver auf die mentale Anstrengung und die intrinsische Motivation aus, als die Kombination aus Farben und Anthropomorphismen. In Bezug auf die Zufriedenheit schien die Farbe jedoch einen größeren Effekt zu haben, als die Kombination aus Farben und Anthropomorphismen.

6 Problemstellung

6.1 Die Relevanz von Concept Map-Trainingsmaßnahmen

Die Nutzung von *Concept Maps* kann aufgrund ihrer lernförderlichen Eigenschaften Lernende beim Erlangen von Zusammenhangswissen unterstützen und so bedeutungsvolles Lernen fördern (siehe **Kapitel 2.2** und **2.3**). Damit dieser lernförderliche Effekt jedoch eintreten kann, wird angenommen, dass Lernende zunächst Expertise und Erfahrung, sprich Vertrautheit, in der Nutzung von *Concept Maps* erlangen müssen (siehe **Kapitel 2.4**). Fehlende oder unzureichende Vertrautheit kann bei Lernenden zu einer Überforderung führen, welche auf eine zu hohe kognitive Belastung zurückzuführen ist.

Auf welche Art und Weise und in welchem zeitlichen Rahmen dieses Mindestmaß an Vertrautheit erreicht werden kann, ist jedoch ungeklärt. Die bislang verwendeten Einführungsformen und Trainingsprogramme für *Concept Maps* unterscheiden sich hinsichtlich vieler Parameter (z. B. Dauer, Themenwahl und Grad der Unterstützung), weshalb sich ein entsprechend heterogenes Bild in Bezug auf ihren Nutzen ergibt (siehe **Kapitel 2.4**). Bei der Vermittlung von reinem Faktenwissen über *Concept Maps* scheint die Lernförderlichkeit geringer auszufallen, als bei einer Kombination aus Faktenwissen und praktischer Übung im Umgang mit *Concept Maps* (prozedurales Wissen). Zudem sollte auch die Vermittlung von metakognitivem Wissen einbezogen werden, indem Lernende dazu befähigt werden, den Einsatz von *Concept Maps* lernförderlich zu gestalten (Leutner & Leopold, 2006). Laut Mandl und Friedrich (2006) sollte die Vermittlung von Lernstrategien sowohl systematisch als auch orientiert an kognitiver Trainingsforschung erfolgen. Bei bereits bestehenden Einführungen und Trainingsmaßnahmen sind die der Konzeption zugrundeliegenden theoretischen und methodischen Grundlagen jedoch nicht immer ersichtlich (z. B. Sumfleth et al., 2010; Quinn et al., 2003).

Aufgrund der Ungewissheit hinsichtlich der Gewichtung dieser beiden Aspekte (prozedurales Wissen und deklaratives Wissen) sowie welche Lerninhalte und praktische Übungen für eine Maximierung der Lernförderlichkeit enthalten sein sollten, weichen die aktuell verwendeten Einführungsformen bzw. Trainingsprogramme stark voneinander ab. Diese Unklarheiten münden in kontroversen Diskussionen über die Frage, ob eine Einführung in Form von Trainingsmaßnahmen grundsätzlich notwendig ist und wenn ja, wie

diese aussehen sollten. Hervorzuheben ist hier, dass die erwähnte Diskussion und die bereits entwickelten Trainingsmaßnahmen sich vorrangig mit der Konstruktion von *Concept Maps* beschäftigen, die Rezeption wird hier kaum näher betrachtet.

Es scheint davon ausgegangen zu werden, dass die Expertise in der Konstruktion von *Concept Maps* auch zur Nutzung des passiveren Formats befähigt. Diese Schlussfolgerung ist jedoch voreilig, da das Erstellen von *Concept Maps* andere kognitive Prozesse erfordert als das Rezipieren (siehe **Kapitel 2.2**). Daher kann nicht zwangsläufig davon ausgegangen werden, dass Lernende, die in der Konstruktion von *Concept Maps* geübt sind, die Struktur einer vorgefertigten *Concept Map* besser durchdringen können. Um diese Frage zu beantworten, müssen beide Formate im Hinblick auf die jeweils notwendige Vertrautheit differenziert betrachtet werden.

Die skizzierte unklare Forschungslage kann auch die praktische Umsetzung beeinflussen. In der Schulpraxis stellt die Vermittlung von Lernstrategien einen wichtigen Bestandteil des Lehrplans dar und hat somit auch einen großen Stellenwert im Unterricht (Mandl & Friedrich, 2006). Auch Friedrich und Mandl (1997) weisen darauf hin, dass die Nutzung verschiedener Lernstrategien schon früh in der schulischen Laufbahn gefördert werden sollte. Die Umsetzung der Lernstrategieförderung im Unterricht stellt momentan eher „ein Nebenprodukt schulischen Lernens“ dar (Mandl & Friedrich, 2006, S. 17) und muss aus der Sicht verschiedener Autor:innen ausgebaut werden (Artelt, 2000; Bannert & Schnotz, 2006). Zur Wissensvermittlung haben sich Lernstrategien wie das Erstellen von Notizen oder das Schreiben von Zusammenfassungen schon lange etabliert, und werden häufig verwendet (Hilbert & Renkl, 2008; Renkl & Nückles, 2006). Verschiedene Autor:innen betonen jedoch, dass es wichtig ist, Lernende im Laufe ihrer schulischen Ausbildung mit einem breiten Repertoire an Lernstrategien auszustatten (Artelt, 2000; Friedrich & Mandl, 1997).

Obwohl *Concept Maps* Lernende bei dem Erwerb von Zusammenhangswissen unterstützen können, scheint sich der Einsatz von *Concept Maps* im Unterricht noch nicht etabliert zu haben (Hilbert & Renkl, 2008). Dass *Concept Maps* hier wenig Beachtung finden ist vor allem vor dem Hintergrund nur schwer nachvollziehbar, dass der Erwerb von Zusammenhangswissen Lernenden zwar auf der einen Seite große Schwierigkeiten bereitet, aber auf der anderen Seite in vielen (naturwissenschaftlichen) Schulfächern, wie beispielsweise Biologie, unverzichtbar ist (Schmid & Telaro, 1990). In dem Beitrag mit dem Titel „If *Concept Mapping* is so helpful to learning biology, why aren't we all doing it?“ weist

Kinchin (2001, S. 1) auf die Diskrepanz zwischen dem evidenzbasierten Potenzial von *Concept Maps* in Bezug auf das Lernen und ihrer tatsächlichen Verwendung im Unterricht hin⁵³.

Über die Gründe, warum Lehrende *Concept Maps* nicht in ihrem Unterricht einsetzen, können aus Mangel an Forschung auf diesem Gebiet nur Vermutungen angestellt werden. Trotz der intensiven und jahrzehntelangen Forschung zu diesem Thema besteht die Möglichkeit, dass vielen Lehrenden die Vorteile von *Concept Maps* nicht bekannt sind. Ein weiterer Grund für die zurückhaltende Implementierung von *Concept Maps* könnte die fehlende Verfügbarkeit von Lehrmaterialien zu deren Einführung sein. Auch ist es möglich, dass Lehrende bereits versucht haben, *Concept Maps* im Unterricht einzusetzen, dieser Versuch jedoch aus Mangel an Handlungsempfehlungen missglückt ist. Ein weiterer Erklärungsansatz könnte sein, dass keine sofortige Lernförderlichkeit zu verzeichnen ist, sondern im Gegenteil die Verwendung von *Concept Maps* zunächst sogar zu Leistungsabfällen führen kann, bis die Lernstrategie ihr volles Potential entfaltet. Dies könnte Lernende demotivieren und die Beobachtung dessen könnte Lehrende wiederum davon abhalten *Concept Maps* (weiter) im Unterricht zu verwenden.

Basierend auf der skizzierten Sachlage ergeben sich folgende Teilziele für die Untersuchung von Concept Map-Trainingsmaßnahmen:

- a) Theoriegestützte Trainingsmaßnahmen⁵⁴ für *Concept Maps* mit dem Fokus auf die Rezeption oder die Konstruktion sollen entwickelt werden, welche in den Schulunterricht integrierbar sind.
- b) Die entwickelten Trainingsmaßnahmen sollen mit einer deklarativen Kurzeinführung hinsichtlich ihrer kognitiven und metakognitiven Wirkung verglichen werden.
- c) Es soll überprüft werden, ob Expertise in der Nutzung eines der beiden *Concept Map*-Formate (Konstruktion oder Rezeption), Lernende in der Verwendung des anderen Formats befähigt.

⁵³ Aktuelle Aussagen darüber in welchem Maße *Concept Maps* im Schulalltag eingesetzt werden und wieviel Erfahrung Schüler:innen mit dieser Methode haben, sind schwer aus der bestehenden Forschung abzuleiten.

⁵⁴ Theorien und Konzepte, nach denen die Erstellung der Trainingsmaßnahmen in dieser Arbeit erfolgte, werden in **Kapitel 8** dezidiert dargelegt.

- d) Um tiefere Erkenntnisse in Bezug auf expertisebedingte Unterschiede in der Informationsverarbeitung zu erlangen, sollen *Eye-Tracking*⁵⁵-Untersuchungen durchgeführt werden.

6.2 Die Wirkung von Emotional Design-Abbildungen

Alle Lernsituationen haben gemein, dass sie mit einer kognitiven Belastung einhergehen (Chandler & Sweller, 1991; Sweller, 1998). Bei der Rezeption von *Concept Maps* kann eine hohe kognitive Belastung zu einem sogenannten *Map Shock* führen, der sich nachteilig auf die Motivation und den Affekt der Lernenden auswirken kann (siehe **Kapitel 2.4**). Im Rahmen der Forschung zu multimedialem Lernen wurde beobachtet, dass die simultane Darstellung von bildlichen und textlichen Elementen die kognitive Belastung reduzieren kann (siehe **Kapitel 3.3**).

Obwohl häufig argumentiert wird, dass *Concept Maps* bereits beide Elemente in Form von Begriffen bzw. Relationen (textliche Elemente) und Verbindungspfeilen (bildliche Elemente) beinhalten, könnten größere Effekte erzielt werden, wenn Abbildungen (z. B. Illustrationen) in *Concept Maps* integriert werden (Alpert und Grueneberg, 2001).

Im Rahmen der Forschung zum sogenannten *Emotional Design* wird postuliert, dass bestimmte Gestaltungsmerkmale den affektiv-motivationalen Zustand der Lernenden beeinflussen können und dadurch das Lernen fördern können (**Kapitel 5**). Zwar wurde in vorherigen Studien der emotionale Einfluss des *Emotional Designs* untersucht, eine differenzierte Betrachtung der Emotionsdimensionen (Valenz oder Aktivierung) existiert bislang jedoch nicht. Da insbesondere für die Dimension der Aktivierung ein Einfluss auf die Lernleistung vermutet wird (Schneider et al., 2019), wäre eine differenzierte Betrachtung sinnvoll.

In vorherigen Studien wurde beobachtet, dass das *Emotional Design* die empfundene Aufgabenschwierigkeit reduzieren kann (Brom et al., 2018; Uzun & Yildirim, 2018). Diesbezüglich wird angenommen, dass das *Emotional Design* die extrinsische kognitive

⁵⁵ Hierbei handelt es sich um eine Methode mit der Blickbewegungsdaten aufgezeichnet und analysiert werden können. Laut der *Eye-Mind*-Hypothese ermöglicht dies eine nicht-invasive und objektive Untersuchung kognitiver Prozesse (Just & Carpenter, 1980).

Belastung eines Lernmaterials reduzieren kann, indem es bei der Aufmerksamkeitslenkung unterstützt. Diese Eigenschaft des *Emotional Designs* ist derzeit noch nicht erwiesen (Le et al., 2021, Park et al., 2015; Peng et al., 2021). Um dezieltere Erkenntnisse über die emotionale Wirkung des *Emotional Designs* und dessen Einfluss in Bezug auf die Aufmerksamkeit zu erlangen, empfehlen mehrere Autor:innen die Verwendung psychophysiologischer Messinstrumente⁵⁶ (z. B. Brom et al., 2018; Bülbül & Abdullah, 2021; Peng et al., 2021).

Die emotionale Gestaltung von Lernmaterialien ist für das schulische Lernen besonders interessant, da ihre Wirkung auf den affektiven Zustand und das Lernen in vor allem jüngeren Altersgruppen hoch sein soll (Schneider et al., 2019; Wong & Adesope, 2021). Die meisten Studien auf dem Gebiet des *Emotional Designs* wurden bisher allerdings an College-Studierenden durchgeführt (Brom et al., 2018). Die Beantwortung der Frage, ob emotionale Gestaltungsmerkmale die Lerneffizienz von *Concept Maps* erhöhen und die kognitive Belastung reduzieren können, würde nicht nur einen Erkenntnisfortschritt in Bezug auf die Gestaltung von *Concept Maps* liefern, auch das noch junge Forschungsfeld des *Emotional Designs* könnte von dieser noch wenig erforschten Übertragung auf *Concept Maps* profitieren⁵⁷.

Basierend auf der skizzierten Sachlage ergeben sich folgende Teilziele für die Untersuchung von Emotional Design-Abbildungen:

- a) Die Gestaltungsprinzipien des *Emotional Designs* sollen auf *Concept Maps* übertragen werden.
- b) Die kognitive und affektive Wirkung soll zwischen *Concept Maps*, die im *Emotional Design* gestaltet wurden und *Concept Maps*, die nicht nach diesen Prinzipien erstellt wurden, verglichen werden.

⁵⁶ Durch psychophysiologische Messungen lassen sich emotionale und kognitive Prozesse anhand physiologischer Daten untersuchen (Fahr & Hofer, 2013). In dieser Studie wurden *Eye-Tracking*-Messungen und Messungen der Hautleitfähigkeit (elektrodermale Aktivität) vorgenommen.

⁵⁷ Tien et al. (2018) führten eine Studie an College-Studierenden zu emotional gestalteten, multidimensionalen *Concept Maps* (digitale *Concept Maps*, bei denen Videos, Bilder oder Sounds mit den Konzepten verknüpft sind) durch. Sie fanden heraus, dass die emotionale Gestaltung zu besseren Lernergebnissen führen kann, als nicht-emotional gestaltete multidimensionale *Concept Maps* (Tien et al., 2018). Es sollte jedoch beachtet werden, dass Tien et al. (2018) das Emotionale Design anders interpretierten als es die Gestaltungsprinzipien nach Um et al. (2012) vorgeben. Anstatt Anthropomorphismen zu verwenden, wandten sie Werbung für das tägliche Leben, Musik, Animationen und Bilder an, um Emotionen hervorzurufen.

- c) Um tiefere Erkenntnisse in Bezug auf die emotionalen und die potenziell aufmerksamkeitslenkenden Eigenschaften des *Emotional Designs* zu gewinnen, sollen psychophysiologische Untersuchungen durchgeführt werden.

7 Projektübersicht

Mit dem Ziel das Lernen mit *Concept Maps* zu erleichtern und perspektivisch die Diskrepanz zwischen theoretischer Wirksamkeit und praktischer Nutzung von *Concept Maps* zu reduzieren, wurden im Rahmen dieser Arbeit zwei Projekte durchgeführt (siehe **Abbildung 11**). Diese widmeten sich *Concept Maps* im Rahmen des Biologieunterrichts und bestanden aus jeweils einer Haupt- und einer Begleitstudie. Bei den Hauptstudien handelte es sich um (quasi-)experimentelle Interventionsstudien, welche durch psychophysiologische Messungen im Rahmen von Begleitstudien ergänzt wurden. Während sich **Projekt 1** der Entwicklung und Untersuchung von *Concept Map*-Trainingsmaßnahmen widmete, wurden im Rahmen von **Projekt 2** verschiedene Designs und deren Integration in *Concept Maps* untersucht⁵⁸.

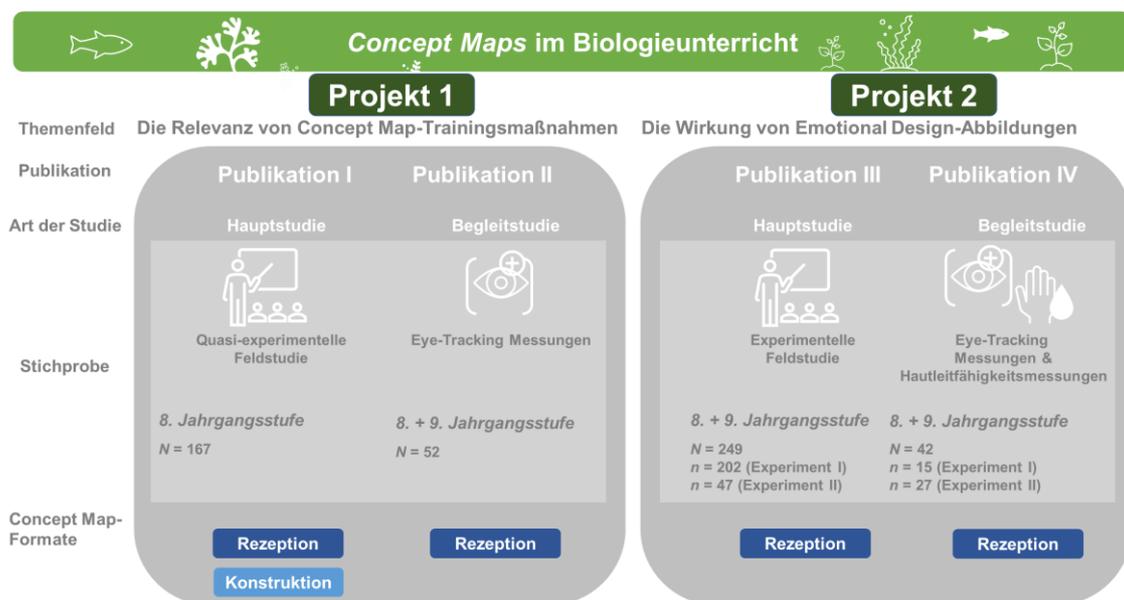


Abbildung 11. Übersicht über die im Rahmen dieser Arbeit durchgeführten Projekte.

⁵⁸ Eine detaillierte Vorstellung der einzelnen Studien folgt in **Kapitel 8 bis 11**.

Eigene empirische Studien

In diesem Kapitel werden die Publikationen, die dieser Dissertation zugrunde liegen vorgestellt⁵⁹. Für jede Publikation werden die korrespondierenden Studien kurz zusammengefasst. Zudem werden die behandelten Fragestellungen, der methodische Zugang und der Eigenanteil an der Studie dargelegt.

8 Die Relevanz von Concept Map-Trainingsmaßnahmen

Die in **Publikationen I** und **II** veröffentlichten Studien beschäftigen sich mit der Relevanz von *Concept Map*-Trainingsmaßnahmen. Die theoretischen Grundlagen dieser Studien sind in **Kapitel 2** zu finden.

8.1 Comparing construction and study of concept maps – An intervention study on cognitive, metacognitive and emotional effects of training & learning (Publikation I)

8.1.1 Zusammenfassung, Fragestellungen und methodischer Zugang

Viele Wissenschaftler: innen auf dem Gebiet der *Concept Map*-Forschung unterstützen die Annahme, dass die Lerneffektivität von *Concept Maps* von der Vertrautheit mit dieser Methode abhängt (Aguiar & Correia, 2017; Großschedl & Tröbst, 2018; Holley & Dansereau, 1984; Renkl & Nückles, 2006). Dennoch ist umstritten, ob und in welcher Intensität diese Methode verwendet bzw. geübt werden muss, damit sie ihre volle Lernförderlichkeit entfalten kann (Mintzes et al., 2011). Diese Debatte befasst sich hauptsächlich mit dem Prozess der Konstruktion und vernachlässigt die Rezeption. Im Rahmen dieser Studie wurden zwei *Concept Map*-Trainingsmaßnahmen und ein Kontrolltraining entwickelt. Die *Concept Map*-Trainingsmaßnahmen widmeten sich entweder dem Prozess der Rezeption oder dem der Konstruktion. Die Lernwirkung der Trainingsmaßnahmen von beiden Prozessen wurde ebenfalls getrennt voneinander untersucht. An der

⁵⁹ Die Originalmanuskripte wurden für die Aufnahme in diese Arbeit in Bezug auf die Schriftgröße, die Schriftfarbe und die verwendeten Hervorhebungen einheitlich formatiert. Alle weiteren Formatierungen wurden nicht verändert. Aus diesem Grund weichen einige Formatierungsweisen sowie die Umsetzung der geschlechtergerechten Sprache von dem Rahmen der Dissertation ab.

durchgeführten quasiexperimentellen Feldstudie im 3x2-faktoriellen Design nahmen insgesamt $N = 167$ Schüler:innen teil.

Die konkreten Fragestellungen, welche im Rahmen dieser Studie beantwortet werden sollen, lauten:

F₁*: Können durch *Concept Map*-Trainingsmaßnahmen die lernhinderlichen Formen der kognitiven Belastung reduziert werden?

F₂*: Verbessert sich die Lernleistung und die Qualität konstruierter *Concept Maps* sowie die Genauigkeit der Selbsteinschätzung durch das Absolvieren einer *Concept Map*-Trainingsmaßnahme?

F₃*: Haben Schüler:innen mehr Freude am Lernen mit *Concept Maps*, wenn sie deren Nutzung zuvor trainiert haben?

F₄: Lässt sich die Expertise in der Nutzung eines *Concept Map*-Formats auf das jeweils andere Format übertragen?

*Hier ist ebenfalls der Unterschied zwischen den beiden *Concept Map*-Trainingsmaßnahmen von Interesse.

8.1.2 Eigenanteile

Die Autorin der vorliegenden Dissertation war maßgeblich an der Konzeption der Studie, welche **Publikation I** zugrunde liegt, beteiligt. Sie hat die Studie geplant und durchgeführt. Bei der Durchführung der Studie und der Dateneingabe wurde sie von Abschlussarbeitskandidat:innen⁶⁰ unterstützt. Die Datenanalyse, die Verfassung des Manuskripts sowie Einreichung des Manuskripts erfolgte gemeinsam mit Frau Stefanie Elsner. Das Manuskript wurde von Herrn Prof. Dr. Großschedl überarbeitet, welcher auch die Supervision übernahm.

⁶⁰ Judith Keusch, Sophie Telaak, Johanna Backes, Natascha Schomacher und Joahanna Oesterle.

8.1.3 Publikation I: Originalmanuskript

Comparing construction and study of concept maps – An intervention study on cognitive, metacognitive and emotional effects of training & learning

Sina Lenski, Stefanie Elsner, Jörg Großschedl

Abstract

Concept maps are graphical tools for organizing and representing knowledge. They are recommended for biology learning to support conceptual thinking. In this study, we compare concept map construction (CM-c, i.e., creating concept maps) and concept map study (CM-s, i.e., observing concept maps). Existing theories and indirect empirical evidence suggest distinct effects of both formats on cognitive, metacognitive and emotional aspects of learning. We developed a CM-c training, a CM-s training, and a brief introduction to concept maps (control training) for junior high school students. We investigated effects on *learning performance*, *concept map quality*, *cognitive load* (cognitive effects), accuracy of *self-evaluation* (metacognitive effects) and *enjoyment* (emotional effects) of these trainings in a subsequent learning phase (CM-c learning vs. CM-s learning) in a quasi-experimental 3x2 factorial-designed study ($N = 167$), involving the factors training type and learning type. Results reveal that CM-c training increased *learning performance* and *concept map quality*. Effects of CM-c training on *learning performance* transferred onto learning with CM-s. *Self-evaluation* was slightly more accurate after CM-c training than CM-s training. Students reported moderate, and highly varying *enjoyment* during CM-c and CM-s learning. The superiority of CM-c over CM-s in *learning performance* and *concept map quality* probably lies in its characteristic of being an active learning strategy. We recommend practitioners to favor CM-c training over CM-s training, and foster students' active engagement and *enjoyment*.

Keywords: Biology education, concept maps, metacognition, learning performance, training

Introduction

Natural sciences deal with the description, explanation and prediction of natural phenomena. Inherent to understanding the natural sciences is conceptual thinking. Conceptual thinking involves organisation of new knowledge and the integration of it into already existing knowledge. Modern biology lessons aim to provide opportunities for students to develop skills in conceptual thinking and educate students to apply these skills to become solution-focused problem solvers. While conceptual thinking can be challenging for students (Ekinici & Şen, 2020; OECD, 2016), it can be encouraged through many different learning strategies. Working with Concept maps provides such a learning strategy (e.g., Tseng, 2020). Concept maps (CMs) are network-like diagrams for organising and representing knowledge. They summarise and visualise the most important concepts of a topic and the relationships between these concepts. Concepts are linked with labelled arrows whereas the direction of the arrowheads specify the reading direction. Concept map construction (CM-c) is the process of creating a concept map (mostly) based on textual material by self-organising concepts and arrows. Concept map study (CM-s), on the other hand, is the process of viewing a previously (expert-) designed concept map without additional textual material.

Concept maps have been intensively examined and further developed since their introduction in the 1970s by Joseph Novak. Many recommendations were given for their use (see e.g., Schroeder et al., 2017 for a recent overview). Heterogeneous results regarding the learning effectiveness of concept maps are often explained by the notion that the learners had different expertise in the use of concept maps. Up to now, it is controversially discussed whether concept map training is necessary in order to use concept maps successfully and how this training should be structured. This study presents and examines two concept map trainings, focusing on concept map construction on the one hand and concept map study on the other. The aim of this study was to 1) develop a training structure based on theoretical foundation and empirical evidence, 2) examine the effects of familiarity with concept maps on the learning process, and 3) investigate to what extent expertise with one learning format (e.g., concept map study) is conducive to the use of the other format (here: concept map construction). We specifically aim at deriving implications for practitioners and future research from our study.

Theoretical Framework

Learning effectiveness of the construction and study of concept maps

CM-c and CM-s are regularly used in classrooms and empirical comparison of their effects on learning seems valuable. Learning with concept maps can yield improved learning outcome (Visible learning Meta^X research base ®; 2021). This is especially prevalent when CM-c and CM-s are compared with other learning strategies. Learners who constructed concept maps outperformed learners who took notes (Reader & Hammond, 1994), created summaries, discussed with fellow students (Chularut & DeBacker, 2004), marked texts (Amer, 1994), and read texts or attended a lecture (Hwang et al., 2021; Nesbit & Adesope, 2006; Woldeamanuel et al., 2020). Learners who studied (animated) concept maps outperformed others who studied texts (Nesbit & Adesope, 2011, O'Donnell et al. 2002, Patterson et al., 1992, Rewey et al., 1989), lists (Lambiotte, Skaggs, & Dansereau, 1993), or outlines (Salata, 1999). Meta-analyses report mixed findings when comparing CM-c and CM-s based on effect sizes. Horton et al. (1993) observed greater benefits for CM-s than for CM-c. In contrast, Schroeder et al. (2017) and Adesope and Nesbit (2013) observed greater benefits for CM-c than CM-s. The more recent meta-analysis including more studies and larger sample sizes, provide evidence for superiority of CM-c over CM-s in *learning performance*. We are not aware of empirical studies that directly compared the effects of CM-c and CM-s on learning outcome. Comparing CM-c and CM-s will offer insight into the robustness of theory-driven cognitive mechanisms of learning with concept maps. Findings might also provide guidance for practitioners to make decisions about learning strategy use.

Cognitive effectiveness of the construction and study of concept maps

Based on Ausubel's theory on learning (Ausubel et al., 1978), it is argued that concept maps promote meaningful learning (Novak & Cañas, 2008; Schroeder et al., 2017). Meaningful learning is taking place when new knowledge is created or assimilated into existing interconnected knowledge structures through cognitive elaboration (Novak & Cañas, 2008). Meaningful learning involves well-organized, relevant knowledge structure and emotional commitment to integrate new knowledge with existing knowledge (Novak & Cañas, 2008). Potential cognitive effects of learning with concept maps are proposed (Nesbit & Adesope, 2006; Schroeder et al., 2017). They include: 1.) Dual coding through visual and verbal information in concept maps supports effective retrieval,

2.) Cognitive load is reduced and overloading of the memory system is prevented, 3.) Centralization of the key concept allows for better semantic integration, 4.) Semantic structure is marked more clearly compared to text formats, 5.) Simple syntax allows for easy access to learners with yet poor reading and writing abilities, 6.) Greater elaborative thinking is promoted through decision making processes, 7.) Greater elaborative thinking is promoted through higher degree of concision and summarization.

With respect to these proposed cognitive effects, a distinction must be made between different concept map formats. CM-c and CM-s differ particularly in their degree of elaborative thinking and cognitive load (mechanisms 2, 6 and 7). CM-c is presumed to promote learners' active engagement with the interconnections of the content (Hardy & Stadelhofer, 2006; Freeman et al., 2014); it is more cognitively demanding, supports deeper engagement, and fosters a higher level of elaborative thinking than CM-s (Schroeder et al., 2017). Taken together, enhanced *learning performance* through CM-c than CM-s can be assumed. The impact on other relevant learning variables is likely to differ between CM-c and CM-s, too.

Construction and study of concept maps –Training, cognitive load, and transfer

Despite a small number of studies concluding that a short introduction to concept maps is sufficient or that learning with concept maps does not need to be practiced at all (Ifenthaler, 2011; Karpicke & Blunt, 2011; Ruiz-Primo, 2004), research predominantly recommends concept map practice. Most scholars in the field support the notion that the learning effectiveness of concept maps depends on the degree of familiarity (Aguiar & Correia, 2017; Großschedl & Tröbst, 2018; Correia et al., 2008; Holley & Dansereau, 1984; Mintzes et al., 2011; Renkl & Nückles, 2006). Trainings (i.e., extended periods of practice) increase familiarity and hence support learning effectiveness. It was shown that CM-c trainings improve the ability to construct concept maps (Becker et al., 2021; den Elzen-Rump & Leutner, 2007; Jin & Wong, 2010; Leopold & Leutner, 2015; Sumfleth et al., 2010). In line with this, it was observed that expertise in the use of knowledge maps (Chmielewski & Dansereau, 1998) and concept maps (Chang et al., 2002) improves knowledge structuring and information encoding when summarizing texts. CM-s training increased level of expertise measured through eye movement (Lenski & Großschedl, in press). For untrained students, on the other side, CM-c yielded negative effects on *learning performance* (Neuroth, 2007).

These negative effects are probably due to excessive cognitive load. Learners' working memory may get overloaded when processing two types of information simultaneously: strategy-related information about concept mapping and learning-related information about learning contents. Learners might experience a so-called *map shock* when studying concept maps. This is characterized by "bewilderment of not knowing where to start or how to penetrate the topography of the map" (Blankenship and Dansereau, 2000; p. 294).

Theoretically, memory resources can be occupied by three types of *cognitive load*: *intrinsic*, *germane*, and *extraneous load* (Sweller, 2010). *Intrinsic load* arises from the difficulty and complexity of the task. It depends on the number of interacting elements (element interactivity) and learners' prior knowledge. *Intrinsic load* can be manipulated by activating the learners' prior knowledge or simplifying the learning content (Klepsch & Seufert, 2020).

Intrinsic load cannot be altered directly by the design of learning material. On the other side, *extraneous load* is caused by suboptimal design of learning material (e.g., plain, text-based learning materials; e.g., Orru & Longo, 2018; Poppenk et al., 2010). A reduction in *extraneous load* could free resources to be available for acquiring and automating schemes in long-term memory (*germane load*). *Germane load* refers to the learning-related load and comprises resources that are available for acquiring and automating schemes in long-term memory.

Increasing the familiarity with concept maps through training could result in a reduction of *intrinsic* and *extraneous load*; and prevent a *map shock*. Greater familiarity with the task could reduce the amount of new strategy-related information, simplify the learning process and reduce the perceived difficulty (intrinsic load, Young et al., 2014). As a consequence, more cognitive resources for content-related processes (*germane load*) will be available (Mayer & Moreno, 2003).

We presume *intrinsic* (H1.3a) and *extraneous cognitive load* (H1.3b) to be reduced and *germane load* (H1.3c) to be increased through both, CM-c training and CM-s training. We expect this effect to be evident compared to a control training.

Furthermore, we assume that learners who are trained in the use of CM-c or CM-s, show improved skills in constructing concept maps (*concept map quality*) (H1.2) and increased *learning performance* compared to untrained learners (H1.1a).

We additionally aim at understanding whether skills acquired through training in one specific format of working with concept maps impact working with another format. Although both learning formats are somewhat similar, it needs to be assumed that different skills are needed for each type of learning e.g., CM-c learning requires learners to (re-)structure, CM-s learning requires learners to recognize information and compare new knowledge with already existing knowledge. We address the question whether CM-c training is conducive to CM-s and vice versa. If such a transfer effect exists, we might see similar results in *learning performance* when learning with CM-c and CM-s after CM-c training. We assume that CM-c training has higher transfer potential on CM-s learning than CM-s training has on CM-c learning, because concept mapping skills are probably transferred from the (more) active type of use to the (more) passive type of use (H1.1b). Taken together, an advantage of CM-c training on cognitive measurements is expected.

Metacognition in concept map trainings: Accuracy of self-evaluation

In educational research, the term metacognition refers to the ability to plan, to monitor and evaluate one's own learning processes (Flavell, 1979; Schraw, 1998); it conceptually overlaps with 'self-regulated learning' (Panadero, 2017). According to Flavell, metacognitive processes involve metacognitive knowledge about person, task and strategy characteristics and metacognitive experience applying metacognitive strategies. When practicing a new learning strategy, accurate *self-evaluation* seems pivotal, because it might determine appropriate adjustment of learning efforts towards a learning goal. Following Zimmerman's (2000) idea of a circular learning process, accurate *self-evaluation* leads to adapted planning behavior. If the evaluation of one's own skills is accurate, appropriate planning behavior will be applied. Attainment of the learning goal will become more likely for students.

It was shown that some students overestimate, and others underestimate their abilities in various skills (Kruger & Dunning, 1999). The Kruger-Dunning effect was shown to be less evident after improving these skills (Kruger & Dunning, 1999). We assume that the Kruger-Dunning effect, which probably occurs in working with concept maps as well, can be overcome by CM training. Through CM trainings, students acquire necessary declarative and procedural skills. Hence, student's ability to accurately *self-evaluate* their own skills is likely to improve. While we assume that both trainings (CM-c and CM-s) improve student's *self-evaluation*, we expect higher accuracy following a CM-c training

(H2). We expect this because of a higher degree of procedural concept map experience in CM-c training.

Emotion in concept trainings: Enjoyment

According to Ausubel et al. (1978), emotional commitment is an inherent part of meaningful learning. Emotional commitment to a learning task is reflected in the construct of *enjoyment*. *Enjoyment* can be defined as an activity related affective state Pekrun et al. (2006). It is experienced when the activity or the learning material is positively valued and perceived as controllable by the learner (Pekrun et al., 2006). Experiencing *enjoyment* increases task engagement and supports persistent use of a learning strategy beyond training or a formal research study. A few studies report insights into the perception of *enjoyment* during concept map tasks. Romero et al. (2017) observed that students largely enjoy working with concept maps. A percentage of 77.8% and 88.2% of two groups of 13 to 14 year old students stated to 'like working on the subject through concept mapping experience'. A study with university students indicates that *enjoyment* differs between learning formats (Blunt & Karpicke, 2014). Students gave higher reports of *enjoyment* for constructing concept maps after reading a text compared to summarizing the same text in a paragraph (while the text is still present). In this study moderate *enjoyment* was reported (29 to 51 on a scale from 0 = 'not at all' to 100 = 'totally').

CM-trainings have the potential to increase *enjoyment*. Negative affective states which accompany (potential) excessive cognitive demands might be reduced as a consequence of familiarity with concept maps. Learners will be more likely to perceive the task as controllable. We assume that CM-c and CM-s trainings increase familiarity with concept maps, reduce cognitive demands and therefore increase *enjoyment* with working with concept maps (H3). Potential differences between the learning formats (CM-c learning, CM-s learning) are of equal interest in this study.

Overview of the study

We investigate the effects of concept map trainings (CM-c training, CM-s training, control training) and concept map learning type (CM-c learning, CM-s learning) on cognitive (*learning performance, concept map quality, cognitive load*), metacognitive (accuracy of *self-evaluation*) and emotional aspects (*enjoyment*) through a direct comparison.

Methods

This study was conducted at non-academic track schools during regular school days and term. The study was carried out by two instructors. We followed the respective local school law agreements (North Rhine-Westphalian Ministry of Education Science and Research, 2005) and the ethical principles and guidelines for the protection of human subjects of research (Department of Health, 2014).

Design and procedure

Schools were approached and classes invited to take part in the quasi-experimental intervention study. The study covered a period of about three weeks and was carried out in regular biology or natural science lessons (see **Figure 1**). The study involved three main phases: A pretesting phase, a training phase, and a combined learning & testing phase. The study comprised six lessons of 45 min each with visiting times of two lessons each week. Pretesting phase, in which demographical data were gathered, took place in the first school lesson. It was identical for all participants. All Classes were randomly assigned to one of the trainings by drawing lots. Students of all classes underwent a training phase with a CM-c training, a CM-s training or a control training. Training phase lasted for three lessons. After the training phase, students underwent a learning & testing phase, in which their ability to develop knowledge through CM learning was tested. Students were randomly assigned to either one of two types of learning: CM-c learning or CM-s learning. Learning & testing phase lasted for two lessons.

Participants

A total of 201 eighth graders from nine classes (between 12 and 35 students per class) at non-academic track schools in North Rhine-Westphalia, Germany participated in this study. **Supplementary material SM 1** gives an overview of participant allocation, exclusion criteria and the variables analyzed. We excluded thirty-four students from data analyses because crucial parts of the study were missed. The remaining $N = 167$ participants were on average $M = 14.05$, $SD = 0.82$ years old. Of all participants, 47.3% were female and 44.9% were male (7.8% did not provide an answer). A percentage of 52.1% were German native speakers and 25.7% stated another language than German as their first language (22.2% did not provide an answer). Reading fluency was lower (80.42 ± 13.62) than in norm samples (100 ± 15) as assessed by Salzburger Lesescreening (Auer et

al., 2005). The average biology grade was 2.68 (grading scale from 1 = ‘very good’ to 6 = ‘insufficient’). Students were informed that this study will not affect their academic reports.

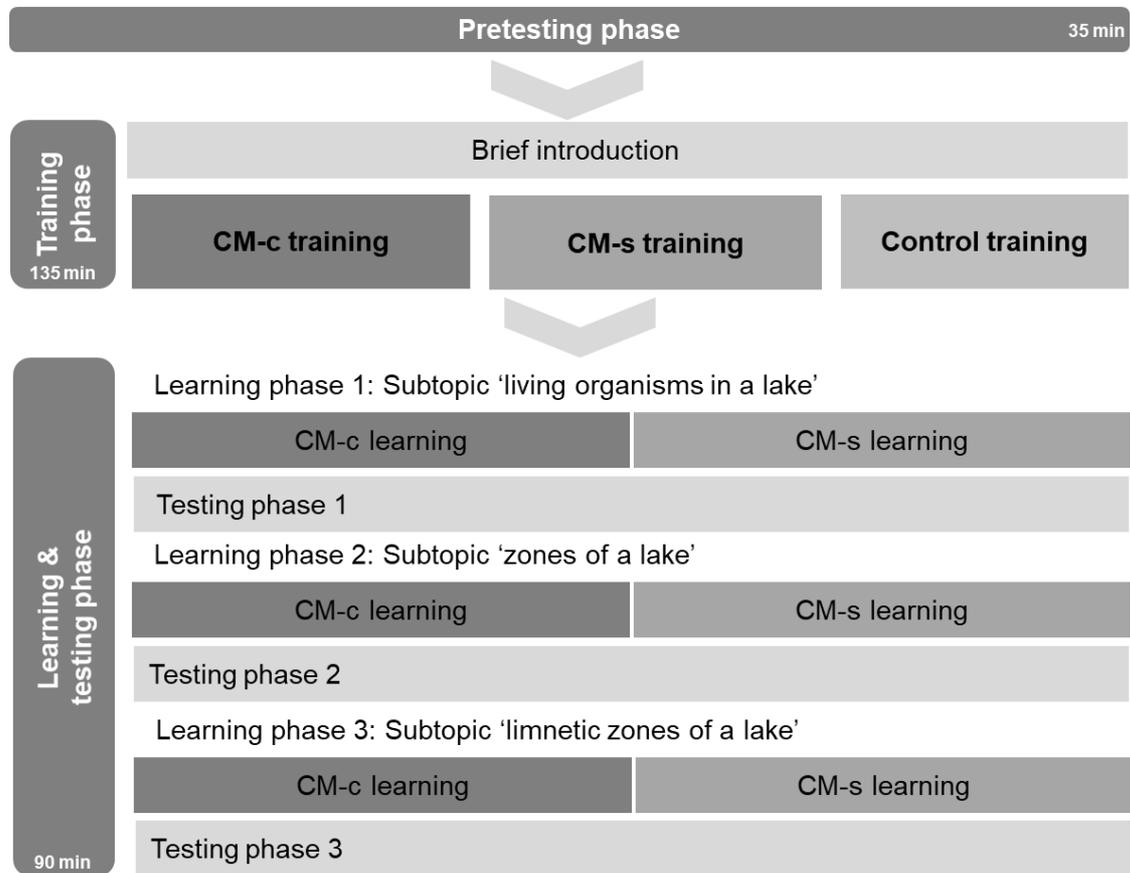


Figure 1. Study design and procedure. Participant allocation can be obtained from **supplementary material SM 1**.

Pretesting phase

During pretesting phase, we gathered students' demographical information, reading fluency and prior knowledge about ecosystems to account for individual differences potentially influencing *learning performance*. Reading fluency was assessed through the Salzburger Lesescreening 5-8 with reported reliability of $r_{tt} = .89$ (SLS 5-8; Auer et al., 2005). Prior knowledge about ecosystems in general and the ecosystem lake was evaluated in a written test including single and multiple-choice questions (see **supplementary material SM 2**). The questionnaire consisted of six self-developed questions and two modified questions obtained from Keusch and Telaak (2017). Additionally, three questions were obtained from the third International Mathematics and Science Study TIMSS (Harmon et al., 1997; Baumert et al., 1998), as the items were validated for grade eighth and cover the

topic ecosystems (see **supplementary material SM 2**, items taken from the TIMSS study are marked accordingly). An item on general knowledge about ecosystems includes, for example, the task of filling in an incomplete food chain (**supplementary material SM 2**, p. 3, item 4). An item focusing on the lake ecosystem covers, for example, the limnetic zone of a lake (**supplementary material SM 2**, p. 5, item 6). Test scores were transformed into a percentage value with 100% indicating solely correct answers. We report a Cronbach's α of .36. We presume this low internal consistency to be a result of high guess probability as shown by other authors (e.g., Bergman et al., 2015).

Training phase

All trainings were based on cognitive theories as recommended by Klauer (1988), Renkl (2010) & Collins et al. (1988). The *theory of adaptive control of thought* (ACT; Anderson, 1983) recommends teaching declarative knowledge (e.g., facts, ideas, and rules) followed by procedural knowledge (knowledge of how an activity is performed) to acquire competence in a certain process. Based on this, all trainings began with a 25-minute introduction to concept mapping. This introduction included declarative knowledge about concept maps, the general idea of concept maps and the use of this new learning method. In CM-c and CM-s training, procedural knowledge about CM-c and CM-s was conveyed. The *cognitive apprenticeship theory* (CAT; Collins et al., 1988) is a constructivist approach to instruction. Cognitive and metacognitive processes which take place during the execution of complex tasks are made visible. This is done by an instructor who verbalizes these processes while the task is performed and provides support and feedback for the learners when performing the task on their own.

Based on this, students underwent four phases (*modelling, scaffolding, fading, and coaching*). The modelling phase was administered for declarative introduction (instructor constructs a sample concept map on the blackboard) whereas the remaining three phases were only carried out in the CM-c and the CM-s trainings but not for the control training. Students in the control training did not receive any further instruction or in-depth information on concept maps beyond the 25-min introduction to concept maps. Instead, students took part in a non-academic social training (team building activity) which did not include a learning activity (see **supplementary material SM 3** for detailed description of the trainings and their theoretical foundation). In (Lenski & Großschedl, 2021), the complete

teaching concept for the construction training in German including all necessary materials is available.

Learning & testing phase

In the learning & testing phase, we examined students' ability to develop knowledge through CM-c learning and CM-s learning. Students studied the topic 'ecosystem lake' in three subtopics ('living organism in a lake', 'zones of a lake', 'limnetic zones of a lake') through either CM-c learning or CM-s learning. The three subtopics were studied consecutively with a learning period of 20 min each. Students' knowledge was tested after studying each subtopic.

During CM-c learning, students constructed concept maps based on learning texts. Stickers with concepts were provided to promote and simplify the construction of concept maps (for a similar approach see Gehl, 2013). During CM-s learning, students were asked to study expert designed concept maps. These concept maps had been designed based on the same textual material as used in CM-c learning. Validity was secured through three independent raters with content equivalence of o Fleiss' $\kappa = 0.96$ for concept map 1 ('living organisms in a lake'), of Fleiss' $\kappa = 1$ for concept map 2 ('zones of a lake'), and of Fleiss' $\kappa = 0.82$ for concept map 3 ('limnetic zones of a lake'). After each learning period, we measured *learning performance*, *concept map quality* (only for CM-c learning, not CM-s learning), *cognitive load*, *self-evaluation*, and *enjoyment*.

Instruments

Learning performance

We assessed learning performance on the topic ecosystem lake by a paper-based questionnaire with open-ended and single choice questions. The questionnaire can be obtained from **supplementary material SM 4**. This questionnaire comprised five self-developed questions, two questions from the TIMSS study (Martin et al., 1997) and 16 modified questions based on Keusch and Telaak (2017). Test scores were transformed into a percentage value with 100% indicating solely correct answers. We report internal consistency of Cronbach's $\alpha = .75$.

Concept map quality

We assessed *concept map quality* through a scoring system as suggested by Clausen and Christian (2012). It allows evaluation of concept map structure and content. Students in CM-c learning condition constructed three concept maps on three subtopics of the ‘ecosystem lake’. Numbers between one and five were assigned for each proposition accounting for the type of relation, labels and connecting structures; 0 = two linked concepts without substantial relation, 1 = two linked concepts, arrow without label but with substantial relation, 2 = two linked concepts with labelled arrow and descriptive relation, 3 = two linked concepts with hierarchical relation, 4 = cause-effect relation without labelled arrow, 5 = cause-effect relation with labelled arrow. Numbers were added to a sum-score. Two rating teams evaluated ten percent of all maps while one rating team rated the entire material. We report an interrater reliability of Cohen’s $\kappa = 0.75$ for concept map 1 (‘living organisms in a lake’), of Cohen’s $\kappa = 0.94$ for concept map 2 (‘zones of a lake’), and of Cohen’s $\kappa = 0.94$ for concept map 3 (‘limnetic zones of a lake’). One overall mean value of all three concept map-sum-scores was calculated for each student.

Cognitive load

We assessed *cognitive load* via the seven-item version of a self-reporting questionnaire designed by Klepsch et al. (2017). We measured *extraneous load (ECL)*, *intrinsic load (ICL)* and *germane load (GCL)*. Questionnaire statements were modified only by the replacement of ‘the task’ with ‘the concept map’ (e.g., ‘When looking at concept maps, many things needed to be kept in mind simultaneously.’). Students rated statements on a 7-point Likert scale ranging from ‘*I fully disagree*’ to ‘*I fully agree*’. Mean values for the subscales over all three times of assessments were computed. We report the following internal consistencies: *extraneous load (ECL)*, Cronbach’s $\alpha = .68 - .78$, *intrinsic load (ICL)*, Cronbach’s $\alpha = .55 - .75$, *germane load (GCL)*, Cronbach’s $\alpha = .75 - .78$.

Self-evaluation

Self-evaluation on students’ concept map skills was measured with five statements; ‘I read the text thoroughly’, ‘I used all the concept stickers’, ‘I paid attention to the direction of the arrows.’, ‘I labelled all the arrows.’ and ‘I understood connections between concepts’. Students rated their agreement on a three-stepped emoticon-based scale (joyful,

indifferent, sad smiley) according to den Elzen-Rump and Leutner (2007). We report internal consistencies for *self-evaluation* for each subtopic (concept map 1: Cronbach's $\alpha = .68$, concept map 2: Cronbach's $\alpha = .77$, concept map 3: Cronbach's $\alpha = .76$).

Enjoyment

Enjoyment was measured with a single question in reference to Blunt and Karpicke (2014). We asked students to answer the question 'How much did you enjoy this task?' on a scale from 0% to 100% in increments of 10%.

Results

Preliminary tests and statistical analyses

Preliminary analysis at an α -level of .10 indicated that there were no differences between training groups in reading fluency, $F(2, 130) = 2.04, p = .135$, prior knowledge about ecosystems, $F(2, 152) = 0.76, p = .471$ or gender proportions, $\chi^2(2) = 1.34, p = .513$ but in age, $F(2, 152) = 2.98, p = .054$ (for descriptive data see **supplementary material SM 5**). As we perceive reading fluency and prior knowledge as greater predictors of *learning performance* than age, we did not regard the age difference between groups as substantial. For most variables, analyses on standard distribution and outliers ($> 3x$ interquartile range) did not yield unusual data distribution. Alternative tests were used in the case of a violation of assumptions. Throughout the results section we use the terms 'TRAINING' and 'LEARNING' for the two independent variables. 'TRAINING' relates to the type of training, which students took part in: CM-c training, CM-s training, control training. 'LEARNING' relates to the type of learning phase, which students underwent subsequently to training. Students studied a topic either through CM-c or CM-s. For all further analyses, α -level was set at $p = .05$. If not provided by IBM SPSS Statistics (version 24.0), effect sizes were calculated according to Lenhard and Lenhard (2016). Because of missing data in the control group and potential distorting statistical results, we interpret statistical results for *cognitive load*, *self-evaluation* and *enjoyment* in both training groups but not in the control group.

Learning performance

To investigate whether training type (CM-c training, CM-s training, control training) and type of learning (CM-c learning, CM-s learning) influenced *learning performance*, we

ran a two-way analysis of variance on *learning performance*. **Table 1** and **Figure 2** show means and standard deviations of *learning performance*.

Table 1: Means and standard deviations for learning performance, concept map quality, cognitive load, self-evaluation and enjoyment separate for training type and learning & testing phase

Training type	CM-c training			CM-s training			Control training												
	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD	n	M	SD	n	
Learning & testing phase																			
Learning performance in %	64.78	21.77	31	60.05	19.07	27	54.53	17.08	29	50.83	18.74	30	57.35	21.09	20	52.99	17.12	30	
Concept map quality	29.88	15.58	31	n.a.	n.a.	n.a.	20.03	13.90	29	n.a.	n.a.	n.a.	15.77	14.09	20	n.a.	n.a.	n.a.	
ECL	3.55	1.35	31	3.36	1.37	27	3.92	1.53	28	3.54	1.26	29	3.50	0.97	7	3.13	1.08	29	
GCL	4.24	1.53	31	3.68	1.23	27	4.17	1.50	28	4.31	1.42	29	5.40	0.97	7	3.67	1.48	29	
ICL	3.79	1.29	31	4.11	1.17	27	3.82	1.27	28	4.45	1.57	29	5.26	1.05	8	4.18	1.11	29	
Self-evaluation	2.54	0.49	30	2.45	0.52	27	2.44	0.42	28	2.56	0.47	29	2.65	0.26	7	2.50	0.43	29	
Enjoyment	45.16	35.42	31	26.85	24.95	27	37.70	25.79	29	37.70	25.79	29	67.38	30.15	7	36.67	24.88	29	

Note. CM-c = Concept map construction, CM-s = Concept map study, *cognitive load* was measured on a 7-point Likert scale ranging from (1) = low cognitive load to (7) = high cognitive load, self-evaluation- was measured on a three-stepped pictorial scale, enjoyment was measured on a scale from 0% to 100%.

We observed that *learning performance* was higher after CM-c training (62.58% ± 20.52%) compared to CM-s training (52.65% ± 17.89%); $F_{TRAINING}(2, 161) = 4.03, p = .020, \eta^2_p = .05$ with post-hoc analyses (Bonferroni) resulting in $p = .017, d = .52$ (partially support for **H1.1a**). We observed no differences between the control training (54.74% ± 18.73%) and both CM trainings ($p_{CM-c\ training} = .105; p_{CM-s\ training}$

= 1.00). We did not find that the type of learning impacted *learning performance* (CM-c learning: 59.21% \pm 20.28%, CM-s learning: 54.44% \pm 18.50%); $F_{LEARNING}$ (1, 161) = 2.03, $p = .157$. We did not observe an interaction of training type with type of learning; $F_{TRAINING \times LEARNING}$ (2, 161) = 0.01, $p = .989$ (support for **H1.1b**).

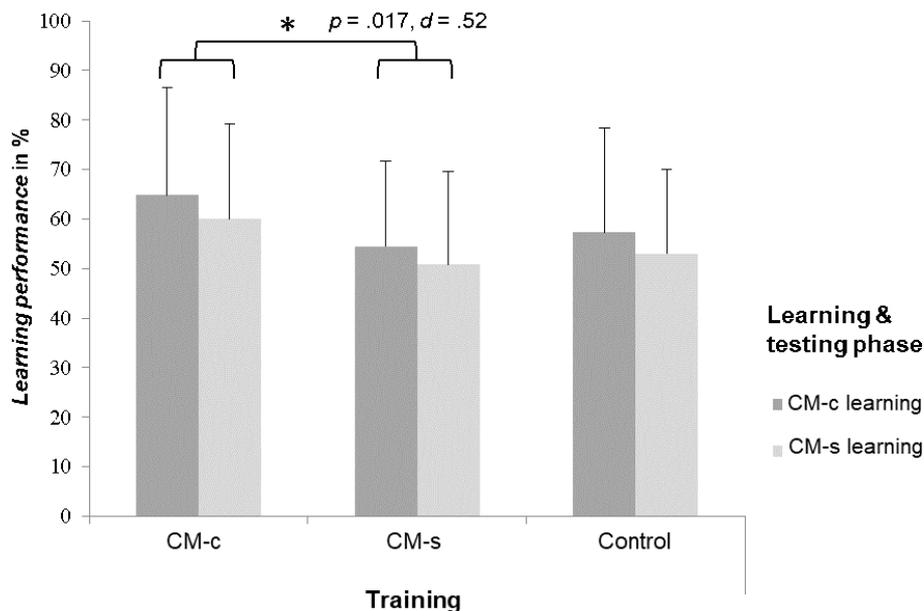


Figure 2. Learning performance. CM-c = Concept map construction, CM-s = Concept map study.

5.3 Concept map quality

To examine differences in *concept map quality* between training groups (CM-c training, CM-s training, control training) during CM learning, we ran a one-way analysis of variance. **Table 1** and **Figure 3** show means and standard deviations for *concept map quality*. Results showed that concept map quality was higher following CM-c training (29.88 ± 15.58) compared to CM-s training (20.03 ± 13.90), $F(2, 77) = 6.47$, $p = .003$, $\eta^2_p = .14$ with post-hoc analyses (Bonferroni) of $p = .033$, $d = .67$. *Concept map quality* was also higher following CM-c training compared to the control training (15.77 ± 14.09 ; $p = .004$, $d = .95$) (partial support of **H1.2**). There was no difference between CM-s training and the control training ($p = .956$).

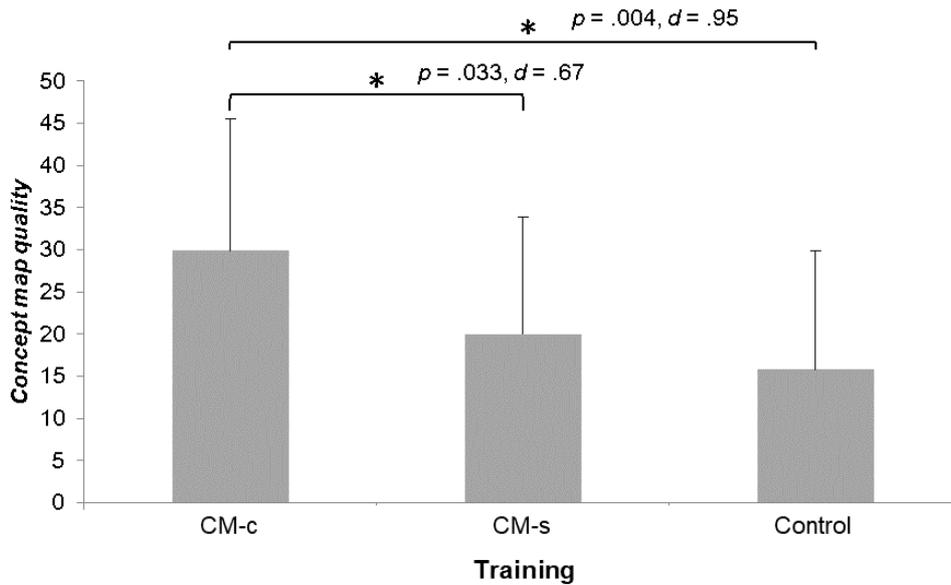


Figure 3. Concept map quality. CM-c = Concept map construction, CM-s = Concept map study.

Cognitive load

To investigate whether training type (CM-c training, CM-s training) and type of learning (CM-c learning, CM-s learning) influenced *cognitive load*, we ran a two-way multivariate analysis of variance on *cognitive load* including *extraneous (ECL)*, *intrinsic (ICL)* and *germane load (GCL)*). **Table 1** shows means and standard deviations. Results of the multivariate analysis revealed no difference in cognitive load between training groups $F_{\text{TRAINING}}(3, 109) = 0.45, p = .715, \text{Wilks' } \Lambda = 0.99, \eta^2_p = .12$, but a difference between type of learning phase $F_{\text{LEARNING}}(3, 109) = 5.25, p = .002, \text{Wilks' } \Lambda = 0.87, \eta^2_p = .13$. This effect did not reach statistical significance after post-hoc testing ($F_{\text{ICL}}(1, 111) = 3.63, p = .059, \eta^2_p = .032$; $F_{\text{GCL}}(1, 111) = 0.61, p = .437, \eta^2_p = .005$; $F_{\text{ECL}}(1, 111) = 1.20, p = .277, \eta^2_p = .011$).

No interaction of training type with type of learning phase was evident $F_{\text{TRAINING} \times \text{LEARNING}}(3, 109) = 1.55, p = .205$. Taken together, training type (CM-c training, CM-s training) and type of learning did not differ in their impact on student's cognitive load (lack of support of **H1.3a-c**).

Self-evaluation

We investigated whether CM trainings influenced accuracy of students' *self-evaluation*.

In our study, accuracy of *self-evaluation* is reflected in the congruency of students' *self-evaluation* (evaluation of concept map skills) and objective assessment (*concept map quality*). As a measurement of congruency, we ran Spearman correlations for ordinal data with *self-evaluation* and *concept map quality* for each training group. High correlations indicate high accuracy of *self-evaluation*. Correlations reveal highest accuracy after CM-c training ($r_s = .66, p < .001, n = 30$), followed by CM-s training ($r_s = .52, p = .004, n = 28$) and the control training ($r_s = .60, p < .159, n = 7$; partially support for H2). **Table 1** shows means and standard deviations for *self-evaluation* and *concept map quality*. We observed that only a small number of participants in the control training provided answers to *self-evaluation* questions. Only a comparison between correlations after CM-c training and CM-s training is legitimate.

Enjoyment

To investigate whether training type (CM-c training, CM-s training) and type of learning (CM-c learning, CM-s learning) influenced emotional commitment to learning with CMs, we ran a two-way analysis of variance on *enjoyment*. *Enjoyment* was analyzed with Box-Cox transformed data because of a violation of homogeneity of error variances. **Table 1** shows untransformed means and standard deviations for *enjoyment*. We observed moderate *enjoyment* and high variability across students ($38.35\% \pm 30.09\%$) with a range of 0% to 100% in *enjoyment*. Students reported average enjoyment following the CM-c ($36.64\% \pm 32.09\%$) and CM-s training ($37.40\% \pm 29.33\%$) with high variability during learning phase. Training type did not influence *enjoyment*; $F_{Training}(1, 111) = 0.40, p = .530$ (lack of support for H3). We observed no effect of type of learning; $F_{Learning}(1, 111) = 2.12 \cdot 10^4, p = .988$. Training type and type of learning did not interact; $F_{TRAINING \times LEARNING}(1, 111) = 3.26, p = .074$. It needs to be noted that analyses revealed a violation of the assumption of homogeneity of error variances. Box-Cox transformation reduced heterogeneity but did not entirely stabilize data as assessed by Levene's test, $p = .036$. The unusually dispersed data might have obscured potential effects. Results need to be observed and interpreted with caution.

Discussion

Learning performance & concept map quality

As expected, results show higher *learning performance* for students who took part in CM-c training instead of CM-s training (partially support for **H1.1b**). As we observed that CM-c training improved *concept map quality* (partially support for **H1.2**), it is likely that the increased *learning performance* is a result of improved concept mapping skills.

In line with other findings (Hilbert & Renkl, 2008; Jin & Wong, 2010; Sumfleth et al., 2010), we assume that CM-s training and the control training are not sufficient to enable students to construct concept maps. A specific training in the construction of concept maps is needed to improve students' ability to construct *concept map* as suggested by other authors (e.g., Großschedl & Tröbst, 2018; den Elzen-Rump & Leutner, 2007; Sumfleth et al., 2010).

Students were able to apply these skills and to engage more deeply with the learning content. This finding supports the assumption that CM-c promotes elaborative thinking. Elaborative thinking probably takes place to a greater extent in CM-c than in CM-s. We ascribe this superiority of CM-c training in *learning performance* to its active nature. Active learning tasks are generally associated with increased *learning performance* (Chang et al., 2002; Freeman et al., 2014; McCagg & Dansereau, 1991).

However, contrary to our hypothesis we did not observe a difference in *learning performance* between CM trainings and the control training. We assume that students who took part in the control training probably did not acquire the necessary skills to effectively apply CM-c or CM-s during learning. Instead of applying concept mapping skills, students probably used other learning strategies that appeared to be beneficial for them in the past (e. g. repeated reading) (see Wild, 2001 for more information on individual learning strategy use). This is supported by the observation of lower concept quality after the control training. Increase in *learning performance* following the control training cannot be explained by an increase in concept mapping skills.

In conclusion, in contrast to CM-s training, CM-c training enabled students to apply concept mapping skills to a degree that allowed them to learn effectively with concept maps. Students improved their ability to construct concept maps and they were able to use this learning strategy to acquire similar knowledge as the use of other naïve strategies would.

To be able to use concept maps as a more effective way of learning, we suggest practice of more than three lessons. The maximum potential of concept maps as a learning strategy might only be exploited by a prolonged training.

Transfer effect

We addressed the questions whether CM-c training impacts CM-s learning and vice versa. Our results show CM-c training increased *learning performance* irrespective of whether students constructed or studied concept maps in a subsequent learning task (support for **H1.1b**). Here, the absence of a statistically significant interaction effect suggests the existence of a transfer effect. An evident interaction effect (i.e., higher *learning performance* after CM-c training for those students who constructed concept maps during learning & testing phase but not for those students who studied concept maps) would have suggested that skills learned through CM-c training are only applied in CM-c learning but not in CM-s learning. We did not observe such an interaction effect and conclude that skills learned through CM-c training are also applied in CM-s learning. The CM-c training most likely altered student's overall information processing strategies, enabling them to implicitly interact with a different CM learning format. This is in line with previous studies suggesting that the familiarity with particular formats can positively influence *learning performance* in similar formats (e.g., Royer, 1979; Royer & Cable, 1976). Our results could be explained by the nature of the tasks (passive vs. active learning task). The familiarity in an active learning task (here CM-c) has higher transfer potential compared to the passive learning task. We conclude that CM-c training benefits *learning performance* regardless of which learning format (CM-c or CM-s) is applied after training.

Cognitive load

We expected *intrinsic* (**H1.3a**) and *extraneous cognitive load* (**H1.3b**) to be reduced and *germane load* (**H1.3c**) to be increased through both, CM-c training and CM-s training compared to the control training. Statistical results showed that CM-c training and CM-s training did not differ in their impact on *cognitive load*. We observed no difference between types of learning.

That cognitive load seemed uninfluenced by training in our study, reflects methodological limitation instead of providing an answer to our research question. We surmise that the used instrument did not differentiate between sources of *ECL* and *ICL* as mentioned by

Klepsch and Seufert (2020), which was published after the conduction of this study. For settings where *ICL* and *EGL* may be intertwined, Klepsch and Seufert (2020) recommend using complex instruments to uncover the underlying processes. We also suspect methodological issues with measuring *GCL* and agree with the authors of the instrument that the ‘wording of the current items was ambiguous, so learners understood them differently’ (Klepsch et al., 2017, p. 9). Therefore, our findings should be treated with caution. Further research is needed to find measurements that reliably assess *cognitive load* during learning activities. We emphasize that simple and clear language that is comprehensible also for low-achieving students should be used.

Self-evaluation

We assumed that CM trainings increase accuracy of *self-evaluation* while we expected that CM-c training has higher influence than CM-s training. Our data only allow a comparison of CM-c and CM-s because of a low number of participants in the control group. Based on effect sizes, results show that accuracy of *self-evaluation* is improved through CM-c training to a greater extent than CM-s training (partially support for **H2**). We assume that this outcome is due to higher amount of procedural knowledge acquired through CM-c training.

Beyond this, we would like to address the question whether accurate *self-evaluation* is a premise or a consequence of successful skill acquisition. The answer to this question has relevant implications for practitioners. If accurate *self-evaluation* is a premise, teachers should include teaching methods that support *self-evaluation* such as providing opportunities for students to reflect on their current level of task skills. If accurate *self-evaluation* is a consequence of successful skill acquisition, teachers should focus on students’ skill practice while *self-evaluation* ‘automatically’ improves. We believe that *self-evaluation* and skill acquisition could be improved at the same time through specific feedback on task skills.

We suggest that specific feedback on task skills should be given when working with any concept map format including CM-c and CM-s. Based on our data, we cannot conclude whether the Kruger-Dunning effect (Kruger & Dunning, 1999) was overcome by training. Nor can we state whether a Kruger-Dunning effect is evident in working with concept maps.

Enjoyment

We hypothesized that CM-c and CM-s trainings increase *enjoyment* during learning with concept maps compared to a control training. Because of missing data, we are unable to answer this research question. Nevertheless, a comparison of CM-c and CM-s learning is legitimate. CM-c and CM-s did not differ in their degree of *enjoyment*. In contrast to Romero et al. (2017), but in line with Karpicke & Blunt (2014), we observed merely moderate *enjoyment* for working with concept maps, while Karpicke & Blunt carried out their study with university students and not school students. We observed in our study higher variability in *enjoyment* than Romero et al (2017), who carried out their study with medium to high achieving students. Moderate *enjoyment* and high variability in our study, lead us to conclude that concept maps should be applied with the aim to enhance *enjoyment*, especially for those students with yet low to medium academic skills as seen in our study.

Interactive concept maps might provide such an opportunity. Results from meta-analysis have already shown promising effects on learning performance (Schroeder et al., 2017), but the small number of studies does not allow a reliable conclusion. Emotional commitment measured as *enjoyment* is an integral part of meaningful learning. Based on our findings, we recommend to take high variability in *enjoyment* into account and support *enjoyment* for students with the aim to enhance meaningful learning.

Limitations

As common for empirical studies, our results need to be viewed in the context of some limitations. We examined *learning performance* immediately after training, as most past findings on trainings on graphic strategies did (Moorf & Readence, 1984). However, delayed learning tests are more sensitive to effects of learning compared to immediate tests (Dunlosky et al., 2013). Future studies might consider analyzing long term effects following concept map trainings to unveil potentially delayed learning effects and we also strongly suggest including motivational measurements as control variables. As most instruments were not designed for the application with junior high school students test validity for this age group has to be confirmed. Furthermore, in one class, only a small number of students gave evaluable answers to the questions regarding *cognitive load*, *self-evaluation* and *enjoyment*. We note that the instruction was disregarded by the students. Moreover, we observed high variability in student's answers e.g., *enjoyment*,

which reflects ‘real life’ situations but limits options for inferential statistical analyses. Potential effects might be obscured.

Conclusion and practical implications

Acknowledging the limitations of our study, the direct comparison of CM-c and CM-s allows us to contribute to recent meta-analytical findings (Schroeder et al., 2017). We conclude that CM-c and CM-s differ in their cognitive and metacognitive effects on learning, but not in emotional effects. CM-c training has a greater impact on *learning performance*, *concept map quality* and accurate *self-evaluation* than CM-s training. Concept mapping skills acquired through CM-c training transfer onto learning with CM-s. Enjoyment does not differ between CM-c and CM-s, but is highly dispersed across students with an overall moderate manifestation.

Based on this study, we aim to provide advice for the use of concept maps in the classroom. We recommend teachers to use CM-c training instead of CM-s training if their goal is to improve *learning performance*, *concept map quality* and students’ ability to *self-evaluate* their own work with concept maps accurately. If a teacher decides to apply CM-s, he or she could use a preceding CM-c training instead of a CM-s training to take advantage of the observed *transfer effect*. We emphasize interindividual differences in students’ *enjoyment* and advise teachers to promote enjoyment to enhance long-term commitment. From our analyses we cannot draw conclusions concerning students’ *cognitive load*. Instead, we advise teachers to seek students’ direct feedback so as to prevent cognitive overload.

Acknowledgments

The authors would like to thank our bachelor’s and master’s candidates (Judith Keusch, Sophie Telaak, Johanna Backes, Natascha Schomacher, and Joahanna Oesterle) for their participation in material development, data collection, and data acquisition.

Author Contributions:

Conceptualization, S.L. and J.G.; methodology, S.L. and J.G.; formal analysis, S.L. and S.E.; investigation, S.L.; resources, J.G.; writing—original draft preparation, S.L. and S.E.; writing—review and editing, J.G.; visualization, S.L. and S.E.; supervision, J.G.; All authors have read and agreed to the published version of the manuscript.

Declarations

Data Availability Statement

Data are openly available (DOI10.17605/OSF.IO/MW356).

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian / next of kin.

Conflicts of Interest

The authors declare no conflict of interest. The authors have no relevant financial or non-financial interests to disclose.

Funding details

This study is part of the project 'Learning Biology through Concept Mapping: Importance of a learning strategy training for cognitive load, cognitive processes and learning performance', which was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft; DFG), Grant (GR 4763/2).

References

- Adesope, O. O., & Nesbit, J. C. (2013). Animated and static concept maps enhance learning from spoken narration. *Learning and Instruction, 27*, 1-10.
- Aguiar, J. G., & Correia, P. R. M. (2017). From representing to modelling knowledge: Proposing a two-step training for excellence in concept mapping. *Knowledge Management & E-Learning: An International Journal, 9*(3), 366-379.
- Amer, A. A. (1994). The effect of knowledge-map and underlining training on the reading comprehension of scientific texts. *English for Specific Purposes, 13*(1), 35-45.
- Anderson, J. R. (1983). *The architecture of cognition*. Harvard University Press.
- Auer, M., Gruber, G., Mayringer, H., & Wimmer, H. (2005). Salzburger Lese-Screening für die Klassenstufen 5-8 (SLS 5-8) [Salzburg Reading Screening for Grades 5–8]. Huber.
- Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational psychology - A cognitive view*. Holt, Rinehart and Winston.
- Baumert, J., Lehmann, R., Lehrke, M., Clausen, M., Hosenfeld, I., Neubrand, J., ... & Günther, W. (1998). Testaufgaben Naturwissenschaften TIMSS 7./8. Klasse (Population 2) [Test questions natural sciences TIMSS 7./8. Class (population 2)]. Max-Planck-Institut für Bildungsforschung.
- Becker, L. B., Welter, V. D. E., Aschermann, E., & Großschedl, J. (2021). Comprehension-Oriented Learning of Cell Biology: Do Different Training Conditions Affect Students' Learning Success Differentially? *Education Sciences, 11*(8), 438. <https://doi.org/10.3390/educsci11080438>
- Bergman, E. M., de Bruin, A. B., Vorstenbosch, M. A., Kooloos, J. G., Puts, G. C., Leppink, J., Scherpbier, A. J., & van der Vleuten, C. P. (2015). Effects of learning content in context on knowledge acquisition and recall: A pretest-posttest control group design. *BMC Medical Education, 15*(1), 1-11.
- Blankenship, J., & Dansereau, D. F. (2000). The effect of animated node-link displays on information recall. *The Journal of Experimental Education, 68*(4), 293-308.
- Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology, 106*(3), 849-858. <https://doi.org/https://psycnet.apa.org/doi/10.1037/a0035934>
- Chang, K.-E., Sung, Y.-T., & Chen, I.-D. (2002). The Effect of Concept Mapping to Enhance Text Comprehension and Summarization. *The Journal of Experimental Education, 71*(1), 5-23.
- Chmielewski, T. C., & Dansereau, D. F. (1998). Enhancing the recall of text: Knowledge mapping training promotes implicit transfer. *Journal of Educational Psychology, 90*(3), 407.
- Chularut, P., & DeBacker, T. K. (2004). The influence of concept mapping on achievement, self-regulation, and self-efficacy in students of English as a second language. *Contemporary Educational Psychology, 29*(3), 248-263.
- Clausen, S., & Christian, A. (2012). Concept Mapping als Messverfahren für den außerschulischen Bereich. [Concept Mapping for Measurement in a non scholar context]. *Journal für Didaktik der Biowissenschaft, 3*, 18-31.

- Collins, A. M., Brown, J. S., & Newman, S. E. (1988). Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics. *Thinking: The Journal of Philosophy for Children*, 8(1), 2-10. <https://doi.org/https://doi.org/10.5840/thinking19888129>
- Correia, P. R. M., Infante-Malachias, M. E., & Godoy, C. E. C. (2008). From theory to practice: the foundations for training students to make collaborative concept maps. In A. J. Cañas, J. D. Novak, P. Reiska, & M. K. Ahlberg (Eds.), *Proceedings of the third international conference on concept mapping* (Vol. 2, pp. 414-421). Vali Press.
- den Elzen-Rump, V., & Leutner, D. (2007). Naturwissenschaftliche Sachtexte verstehen - Ein computerbasiertes Trainingsprogramm für Schüler der 10. Jahrgangsstufe zum selbstregulierten Lernen mit einer Mapping-Strategie [Understanding scientific factual texts - A computer-based training program for 10th grade students for self-regulated learning with a mapping strategy]. In M. Landmann & B. Schmitz (Eds.), *Selbstregulation erfolgreich fördern [Successfully promoting self-regulation]* (pp. 251-268). Kohlhammer.
- Department of Health, E. (2014). The Belmont Report. Ethical principles and guidelines for the protection of human subjects of research. *The Journal of the American College of Dentists*, 81(3), 4.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4-58.
- Ekinci, S., & Şen, A. İ. (2020). Investigating grade-12 students' cognitive structures about the atomic structure: A content analysis of student concept maps. *International Journal of Science Education*, 42(6), 977-996.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415.
- Gehl, D. (2013). Vom Betrachten zum Verstehen [About viewing to understanding]. Springer.
- Großschedl, J., & Tröbst, S. (2018). Biologie lernen durch Concept Mapping: Bedeutung eines Lernstrategietrainings für kognitive Belastung, kognitive Prozesse und Lernleistung – Kurzdarstellung des DFG-Projekts [Learning biology by concept mapping: the importance of learning strategy training for cognitive load, cognitive processes and learning performance - brief description of the DFG project]. *Zeitschrift für Didaktik der Biologie (ZDB) - Biologie Lehren und Lernen*, 22(1), 20-30. <https://doi.org/10.4119/zdb-1630>
- Hardy, I., & Stadelhofer, B. (2006). Concept Maps wirkungsvoll als Strukturierungshilfen einsetzen: Welche Rolle spielt die Selbstkonstruktion? *Zeitschrift für Pädagogische Psychologie*, 20(3), 175-187.
- Harmon, M., Smith, T. A., Martin, M. O., Kelly, D. L., Beaton, A. E., Mullis, I. V., et al. (1997). Performance Assessment: IEA's Third International Mathematics and Science Study (TIMSS). Amsterdam: International Association for the Evaluation of Educational Achievement.
- Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.

- Hernik, J., & Jaworska, E. (2018). The effect of enjoyment on learning. *Proceedings of INTED2018 Conference, 1*, 508-514.
- Hilbert, T. S., Nückles, M., Renkl, A., Minarik, C., Reich, A., & Ruhe, K. (2008). Concept Mapping zum Lernen aus Texten: Können Prompts den Wissens- und Strategieverwerb fördern? [Concept mapping for learning from texts: Can prompts promote knowledge and strategy acquisition?]. *Zeitschrift für pädagogische Psychologie [Journal of Educational Psychology]*, 22(2), 119-125.
- Hilbert, T. S., & Renkl, A. (2008). Concept mapping as a follow-up strategy to learning from texts: What characterizes good and poor mappers? *Instructional Science*, 36(1), 53-73.
- Holley, C. D., & Dansereau, D. F. (1984). Networking: The technique and the empirical evidence. In D. F. Dansereau & C. D. Holley (Eds.), *Spatial Learning Strategies* (pp. 81-108). Elsevier.
- Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, 77(1), 95-111. <https://doi.org/10.1002/sce.3730770107>
- Hwang, G. J., Chang, S. C., Song, Y., & Hsieh, M. C. (2021). Powering up flipped learning: An online learning environment with a concept map-guided problem-posing strategy. *Journal of Computer Assisted Learning*, 37(2), 429-445.
- Ifenthaler, D. (2011). Identifying cross-domain distinguishing features of cognitive structure. *Educational Technology Research and Development*, 59(6), 817-840.
- Jin, H., & Wong, K. (2010). Training on concept mapping skills in geometry. *Journal of Mathematics Education*, 3(1), 104-119.
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772-775.
- Keusch, J., & Telaak, S. (2017). Einfluss der Darstellungsform (Concept Maps vs. Fließtext) auf die Rezeptionsleistung von Schülerinnen und Schülern im Themenbereich „Ökosystem See“ [Influence of the form of representation (concept maps vs. continuous text) on the reception performance of schoolchildren in the subject area "lake ecosystem"] [unpublished master's thesis, University of Cologne].
- Klauer, K. J. (1988). Teaching for learning-to-learn: A critical appraisal with some proposals. *Instructional Science*, 17(4), 351-367.
- Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, 8, 1997.
- Klepsch, M., & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science* 48(1), 45-77. <https://doi.org/10.1007/s11251-020-09502-9>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121.
- Lambiotte, J. G., Skaggs, L. P., & Dansereau, D. F. (1993). Learning from lectures: Effects of knowledge maps and cooperative review strategies. *Applied Cognitive Psychology*, 7(6), 483-497.

- Lenhard, W., & Lenhard, A. (2016). *Berechnung von Effektstärken [Calculation of effect sizes]*. Dettelbach, Germany. Psychometrica. <https://www.psychometrica.de/effektstaerke.html>. (Accessed March 11, 2021).
- Leopold, C., & Leutner, D. (2015, August). Improving students' science text comprehension through metacognitive self-regulation when applying learning strategies. *Metacognition and Learning*, 10(313-346), 23-27.
- Lenski, S., & Großschedl, J. (in press). Biologie lernen mit Concept Maps: Lässt sich die Expertise im Umgang mit Concept Maps von den Augen ablesen? [Learning biology with concept maps: Can the expertise in dealing with concept maps be read from the eyes?]. In P. Klein, M. Schindler, N. Graulich & J. Kuhn (Eds.), *Eye Tracking als Methode in der Mathematik- und Naturwissenschaftsdidaktik: Forschung und Praxis*. Springer.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38(1), 43-52.
- McCagg, E. C., & Dansereau, D. F. (1991). A convergent paradigm for examining knowledge mapping as a learning strategy. *The Journal of Educational Research*, 84(6), 317-324.
- Mintzes, J. J., Canas, A., Coffey, J., Gorman, J., Gurley, L., Hoffman, R., Y., M. S., Miller, N., Moon, B., Trifone, J., & Wandersee, J. H. (2011). Comment on "Retrieval practice produces more learning than elaborative studying with concept mapping" [Technical Comment]. *Science*, 334, 453-453.
- Moorf, D. W., & Readence, J. F. (1984). A quantitative and qualitative review of graphic organizer research. *The Journal of Educational Research*, 78(1), 11-17.
- Nesbit, J. C., & Adesope, O. O. (2011). Learning from animated concept maps with concurrent audio narration. *The Journal of Experimental Education*, 79(2), 209-230.
- Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, 76(3), 413-448. <https://doi.org/10.3102%2F00346543076003413>
- Neuroth, J. (2007). Concept-mapping als Lernstrategie: Eine Interventionsstudie zum Chemielernen aus Texten [Concept mapping as a learning strategy: An intervention study on chemistry learning from texts]. Logos.
- North Rhine-Westphalian Ministry of Education Science and Research. (2005). *Schulgesetz für das Land Nordrhein-Westfalen [School law for the state of North Rhine-Westphalia]*. https://recht.nrw.de/lmi/owa/br_vbl_detail_text?anw_nr=6&vd_id=3928&vd_back=N102&sg=&menu=1
- Novak, J. D., and Cañas, A. J. (2008). *The Theory Underlying Concept Maps And How To Construct Them*. Technical Report IHMC CmapTools 2006-01 Rev 01-2008. Florida Institute for Human and Machine Cognition, Pensacola, Florida, USA. Available online at: <http://www.ssu.ac.ir/fileadmin/templates/fa/Moavenatha/Moavenate-Amozeshi/edicupload/olymp-3.pdf> (accessed March 11, 2021).
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press.
- O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review*, 14(1), 71-86.
- OECD. (2016). PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy. OECD publishing.

- Orru, G., & Longo, L. (2018). The evolution of cognitive load theory and the measurement of its intrinsic, extraneous and germane loads: A review. In L. Longo & M. C. Leva (Eds.), *International Symposium on Human Mental Workload: Models and Applications* (Vol. 1012, pp. 23-48). Springer.
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422.
- Patterson, M. E., Dansereau, D. F., and Newbern, D. (1992). Effects of communication aids and strategies on cooperative teaching. *J. Educ. Psychol.* 84, 453–461.
- Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology*, 98(3), 583. <https://doi.org/10.1037/0022-0663.98.3.583>
- Poppenk, J., Köhler, S., & Moscovitch, M. (2010). Revisiting the novelty effect: When familiarity, not novelty, enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(5), 1321.
- Reader, W., & Hammond, N. (1994). Computer-based tools to support learning from hypertext: Concept mapping tools and beyond. In *Computer Assisted Learning: Selected Contributions from the CAL'93 Symposium* (pp. 99-106). Pergamon.
- Renkl, A. (2010). *Lehren und Lernen [Teaching and Learning]*. In Handbuch Bildungsforschung (pp. 737-751). VS Verlag für Sozialwissenschaften.
- Renkl, A., & Nückles, M. (2006). Lernstrategien der externen Visualisierung [External visualization learning strategies]. *Handbuch Lernstrategien*, 135-147.
- Rewey, K. L., Dansereau, D. F., Skaggs, L. P., and Hall, R. H. (1989). Effects of scripted cooperation and knowledge maps on the processing of technical material. *J. Educ. Psychol.* 81, 604–609.
- Romero, M. d. C., Cazorla, M., & Buzón García, O. (2017). Meaningful learning using concept maps as a learning strategy. *Journal of Technology and Science Education*. 7(3), 313-332.
- Roth, W.-M., & Roychoudhury, A. (1993). The concept map as a tool for the collaborative construction of knowledge: A microanalysis of high school physics students. *Journal of Research in Science Teaching*, 30(5), 503-534.
- Royer, J. M. (1979). Theories of the transfer of learning. *Educational Psychologist*, 14(1), 53-69.
- Royer, J. M., & Cable, G. W. (1976). Illustrations, analogies, and facilitative transfer in prose learning. *Journal of Educational Psychology*, 68(2), 205.
- Ruiz-Primo, M. A. (2004). Examining concept maps as an assessment tool. Proceedings of the first international Conference on Concept Mapping, Pamplona, Spain.
- Salata, M. W. A. (1999). *Concept maps as organizers in an introductory university level biology course*. University of Virginia.
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26(1), 113-125.
- Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review*, 30(2), 431-455. <https://doi.org/10.1007/s10648-017-9403-9>

- Sumfleth, E., Neuroth, J., & Leutner, D. (2010). Concept Mapping – eine Lernstrategie muss man lernen. [Concept mapping – Learning strategy is something you must learn]. *Chemkon*, 17(2), 66-70.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123-138.
- Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology* 4(1), 25-29.
- Tseng, S.-S. (2020). Using concept mapping activities to enhance students' critical thinking skills at a high school in taiwan. *The Asia-Pacific Education Researcher*, 29(3), 249-256.
- Visible learning Meta^X research base ® (2021). Global research data base. CORWIN. Version 1.1. updated August 2021. <https://www.visiblelearningmetax.com/influences> (accessed December 21, 2021).
- Wang, Z., Adesope, O., Sundararajan, N., & Buckley, P. (2021). Effects of different concept map activities on chemistry learning. *Educational Psychology*, 41(2), 245-260.
- Wild, K.-P. (2001). Lernstrategien und Lernstile [learning strategies and learning styles]. In D. H. Rost (Eds.), *Handwörterbuch Pädagogische Psychologie* [Handbook of educational psychology] (pp. 309-312). Weinheim: Beltz, Psychologie VerlagsUnion.
- Woldeamanuel, Y. W., Abate, N. T., & Berhane, D. E. (2020). Effectiveness of concept mapping based teaching methods on grade eight students' conceptual understanding of photosynthesis at ewket fana primary school, Bahir Dar, Ethiopia. *EURASIA Journal of Mathematics, Science and Technology Education*, 16(12).
- Young, J. Q., van Merriënboer, J., Durning, S., & Ten Cate, O. (2014). Cognitive load theory: implications for medical education: AMEE Guide No. 86. *Medical Teacher*, 36(5), 371-384.
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego: Academic Press.

8.2 Biologie lernen mit Concept Maps: Lässt sich die Expertise im Umgang mit Concept Maps von den Augen ablesen? (Publikation II)

8.2.1 Zusammenfassung, Fragestellungen und methodischer Zugang

Während subjektive Fragebögen nur eingeschränkte Einblicke in Prozesse der Informationsverarbeitung erlauben, können Blickbewegungsmessungen mittels *Eye-Tracking*-Geräten tiefere Erkenntnisse zu diesen Prozessen ermöglichen (Gegenfurtner et al., 2011). Laut Funke und Spering (2006) eignen sich diese Messinstrumente zudem um Expertiseunterschiede sichtbar zu machen. Hierzu wurden in vorherigen Studien für Expert:innen typische Blickbewegungscharakteristika identifiziert. In der vorangegangenen Studie (**Publikation I**) wurde beobachtet, dass Schüler:innen welche an einem Konstruktionsstraining teilnahmen, bessere *Concept Maps* erstellten, als Schüler:innen die an einem Rezeptionstraining oder einer Kurzeinführung teilnahmen. Um zu dieser Beobachtung tiefergehende Erkenntnisse zu gewinnen, wurden in der vorliegenden Studie Blickbewegungscharakteristika von *Concept Map*-Expert:innen ($n = 14$) und Noviz:innen ($n = 38$) untersucht, während sie Informationen aus *Concept Maps* entnehmen sollten. Als Expert:innen wurden Schüler:innen betrachtet, welche an einem im Rahmen von **Publikation I** konzipierten *Concept Map*-Konstruktionsstraining teilnahmen. Schüler:innen, die eine Kurzeinführung erhielten wurden hingegen als Noviz:innen eingestuft.

Die konkreten Fragestellungen, welche im Rahmen dieser Studie beantwortet werden sollen, lauten:

F₁: Führt ein *Concept Map*-Konstruktionsstraining zu einer höheren Expertise, welche sich in typischen Blickbewegungsmustern äußert?

F₂: Sind Expertiseunterschiede in einer hierarchisch strukturierten *Concept Map* ausgeprägter als bei netzartig strukturierten *Concept Maps*?

8.2.2 Eigenanteile

Die Autorin der vorliegenden Dissertation hat die Studie, welche **Publikation II** zugrunde liegt konzipiert und durchgeführt. Bei der Durchführung der Studie wurde sie von

Abschlussarbeitskandidat:innen⁶¹ unterstützt. Die Autorin führte die Datenanalyse durch, verfasste das Manuskript und reichte es ein. Das Manuskript wurde von Herrn Prof. Dr. Großschedl überarbeitet, welcher auch die Supervision übernahm.

⁶¹ Benjamin Bayer, Sven Nam Karsten und Regina Drevs.

8.2.3 Publikation II: Originalmanuskript

Biologie lernen mit Concept Maps: Lässt sich die Expertise im Umgang mit Concept Maps von den Augen ablesen?

Sina Lenski & Jörg Großschedl

Abstract

Die Struktur und Funktion lebender und belebter Systeme stellen den Unterrichtsgegenstand im Fach Biologie dar. Die Komplexität dieser Systeme bedingt häufig Verständnisprobleme seitens der Lernenden und resultiert u. a. aus der mitunter großen Vielfalt von Systemelementen, die in Wechselwirkung zueinander stehen. Die Visualisierung solcher Wechselwirkungen durch Concept Maps kann Lernende in ihrem Lernprozess unterstützen. Günstig erweist sich dabei die sehr einfache syntaktische Struktur von Concept Maps. Trotz dieser einfachen Struktur können auch Concept Maps Lernende überfordern, was in einem sogenannten *Map Shock* zum Ausdruck kommen kann. Bislang besteht in der Lehr- und Lernforschung Uneinigkeit darüber, ob die Rezeption von Informationen aus Concept Maps trainiert werden muss. In der vorliegenden Studie nahm eine Gruppe von Lernenden an einem intensiven Concept Map-Training teil (Expert*innen; $n = 14$), während eine zweite Gruppe nur einer Kurzeinführung unterzogen wurde (Noviz*innen; $n = 38$). Anschließend wurden die Blickbewegungscharakteristika von Expert*innen und Noviz*innen während der Rezeption von Informationen aus Concept Maps erfasst. Die Datenanalyse zeigt, dass die Expert*innen einige für einen hohen Expertisegrad typische Blickbewegungscharakteristika (z. B. kürzere Fixationsdauern und höhere Fixationszahlen) aufwiesen. Expert*innen schienen sich außerdem intensiver mit den bereitgestellten Concept Maps auseinandergesetzt zu haben als Noviz*innen, was in einer längeren Gesamtbetrachtungsdauer zum Ausdruck kam. Die erwarteten Expertiseunterschiede traten unabhängig von der Struktur der Concept Maps, nämlich sowohl bei hierarchisch als auch bei netzartig strukturierten Concept Maps, auf. Die Befunde deuten darauf hin, dass die durch ein Concept Map-Training erhöhte Vertrautheit mit dieser Methode die Entnahme von Informationen aus vorgefertigten Concept Maps fördert. Gemeinsam mit einer vorangegangenen Studie stärken die Ergebnisse die Annahme, dass für die Rezeption von Concept Maps Übung notwendig ist, um Informationen aus Concept Maps zu rezipieren. Die vorliegende Studie stellt somit einen wichtigen Erkenntnisfortschritt im Bereich der

Expertiseforschung zu Concept Maps dar und bietet neue Impulse für zukünftige Studien und Implementationen von Trainingsmaßnahmen.

Schlüsselwörter: Concept Maps, Concept Map-Training, Expertise

Einleitung

Empirische Studien legen nahe, dass Lernende bei der Beschäftigung mit biologischen Inhalten besonderer Unterstützung bedürfen (Wendt et al., 2016). Als geeignetes Unterstützungsinstrument werden Concept Maps betrachtet (u. a. Fischer & Mandl, 2000). Sie bestehen aus Begriffen (Konzepten), die über beschriftete Pfeile (Relationen) in eine sinnvolle Beziehung gebracht werden, wobei die Pfeilrichtung die Leserichtung anzeigt (Novak & Gowin, 1984). Concept Maps können entweder als vorgefertigte Informationsquelle (Rezeption von Informationen) bereitgestellt werden oder von den Lernenden aus den verfügbaren fachlichen Informationen selbst konstruiert werden. Sie können dabei unterschiedlich strukturiert sein (z. B. netzartig oder hierarchisch; vgl. **Kapitel „Concept Maps, Eye-Tracking und Expertise“**).

Unerfahrene Lernende scheinen mit der Konstruktion und Rezeption von Concept Maps jedoch häufig überfordert zu sein (u. a. Mintzes et al., 2011). Dies bringt auch der von Blankenship und Dansereau (2000) geprägte Begriff des *Map Shocks* zum Ausdruck. Dieser bezeichnet Schwierigkeiten von Lernenden, Informationen aus vorgefertigten Maps zu entnehmen und ihre Topographie sinnvoll zu entschlüsseln. Ob die Konstruktion und Rezeption von Concept Maps besonderer Trainingsmaßnahmen bedürfen, ist jedoch bislang weitgehend ungeklärt (Großschedl & Tröbst, 2018).

Die vorliegende Studie geht aus einer Follow-up-Erhebung einer größer angelegten Trainingsstudie (Lenski, Elsner & Großschedl, in Vorbereitung) hervor, in der untersucht wurde, ob sich ein Concept Map-Training positiv auf eine nachfolgende Lerngelegenheit mit Concept Maps auswirkt. Dabei bildeten die Lernleistung und die kognitive Belastung der Schüler*innen das Forschungsinteresse. Diese Variablen informieren jedoch nur eingeschränkt über die zugrunde liegenden, internal ablaufenden Prozesse der Informationsverarbeitung. Eye-Tracking gewährt hierbei tiefere Einblicke, da sich diese internalen Prozesse der Informationsverarbeitung aus den Blickbewegungscharakteristika von Personen ableiten lassen (Gegenfurtner, Lehtinen & Säljö, 2011). Expertise in

einer Domäne geht mit Blickbewegungscharakteristika einher, welche eine Unterscheidung zwischen Expert*innen und Noviz*innen ermöglichen (Funke & Spering, 2006; vgl. Kapitel „*Eye-Tracking und Expertise*“).

Dementsprechend wurde in der vorliegenden Studie untersucht, ob Schüler*innen, die ein Concept Map-Training erhielten, im Vergleich zu Schüler*innen, die an einem Kontrolltraining teilnahmen, Blickbewegungscharakteristika aufweisen, wie sie für Experten typisch sind. Neben einer objektiven Evaluation von Expertiseunterschieden stellt Eye-Tracking damit Informationen bereit, wie Experten Concept Maps „lesen“. Dies wiederum kann hilfreich sein, um Noviz*innen im richtigen Umgang mit Concept Maps zu trainieren.

Theoretischer Hintergrund

Eye-Tracking und Expertise

Das *Novizen-Experten-Paradigma* (Funke & Spering, 2006) postuliert, dass Expert*innen andere Herangehensweisen zur Lösung von Aufgaben verfolgen als Noviz*innen. Blickbewegungsanalysen mittels Eye-Tracking Systemen werden seit einigen Jahrzehnten in den verschiedensten Anwendungsbereichen zur Aufklärung von Lösungsstrategien und Expertiseunterschieden eingesetzt (Gegenfurtner et al., 2011). Schachexpert*innen erkennen relevante Bereiche für den nächsten Spielzug beispielsweise schneller als Noviz*innen (Sheridan & Reingold, 2014) und erfahrene Radiologen identifizieren Krebs in einer Mammographie im Gegensatz zu ihren jungen Berufskolleg*innen bereits innerhalb von Sekundenbruchteilen (Kundel, Nodine, Conant & Weinstein, 2007).

Gegenfurtner et al. (2011) untersuchten in ihrer Meta-Analyse den Stand der Expertiseforschung im Bereich des Betrachtens von Visualisierungen (u. a. Diagramme). Dabei konnten sie zeigen, dass Expert*innen im Vergleich zu Noviz*innen unter anderem aufgrund kürzerer Fixationsdauern ($r = 0.27$), höherer Fixationszahlen ($r = 0.53$) sowie längerer Sakkaden ($r = 0.30$) Visualisierungen schneller erfassen können. Bezüglich der Länge von Scanpaths wurde gezeigt, dass steigende Expertise mit kürzeren Scanpaths einhergeht (Krupinski, 2000).

Diese Befunde lassen sich nach Gegenfurtner et al. (2011) durch verschiedene theoretische Annahmen erklären: (1) Auf Grundlage der *Eye-Mind*-Hypothese wird davon ausgegangen, dass Blickbewegungscharakteristika kognitive Prozesse abbilden (Just

& Carpenter, 1984). (2) Die Theorie des *Long-Term Working Memory* (Ericsson & Kintsch, 1995) geht davon aus, dass Vorwissen die Kapazitäten für die Informationsverarbeitung aufgrund des Erwerbs von Abrufstrukturen erhöht. Dies ermöglicht Expert*innen neue Informationen schneller und effizienter im Langzeitgedächtnis zu enkodieren und schließlich gezielt abzurufen. Bezogen auf Eye-Tracking Messungen sollten Expert*innen sich daher durch eine schnellere Informationsverarbeitung mit kürzeren Fixationsdauern, höheren Fixationszahlen und kürzeren Scanpaths auszeichnen. Zudem ist bekannt, dass Expert*innen Probleme gründlicher analysieren als Noviz*innen (Patel & Groen, 1986), was auf die höheren Kapazitäten für die Informationsverarbeitung zurückgeführt werden könnte. Basierend hierauf wäre zu erwarten, dass sich Expert*innen ausführlicher mit den Stimuli beschäftigen und somit die Gesamtbetrachtungsdauer länger ausfällt als bei Noviz*innen. (3) Das *Holistic Model of Image Perception* (Kundel et al., 2007) beschreibt, dass Expert*innen Informationen zunächst global analysieren und sich dann der feineren Analyse widmen. Dies sollte durch längere Sakkaden für relevante Bereiche zum Ausdruck kommen. (4) Im Rahmen der *Information-Reduction Hypothesis* (Haider & Frensch, 1999) wird zudem postuliert, dass Expert*innen die Menge der verarbeiteten Informationen reduzieren, indem sie selektiv aufgabenrelevante Informationen priorisieren und irrelevante Informationen ausblenden. Übertragen auf die Blickbewegungen sollten Expert*innen demnach geringere Fixationszahlen und kürzere Fixationsdauern in aufgabenirrelevanten Bereichen (z. B. White Spaces; Moser, Lohmeyer, Meboldt, Distler & Becker, 2020) und dafür höhere Fixationszahlen sowie längere Fixationsdauern in aufgabenrelevanten Bereichen aufweisen.

Concept Maps, Eye-Tracking und Expertise

Bislang existieren nur wenige Eye-Tracking Studien zur Untersuchung von Expertiseunterschieden bei der Rezeption und Konstruktion von Concept Maps (Amadiou & Salmerón, 2014; Amadiou, van Gog, Paas, Tricot & Mariné, 2009; Dogusoy-Taylan & Cagiltay, 2014). Anders als in der vorliegenden Studie, wird der Expertisegrad in diesen Studien durch das vorhandene Fachwissen (Vorwissen) und nicht durch die Expertise im Lernen mit Concept Maps definiert. Während sich Dogusoy-Taylan und Cagiltay (2014) sowie Amadiou und Salmerón (2014) der Konstruktion von Concept Maps widmeten, stand bei Amadiou et al. (2009) die Rezeption von hierarchisch und netzartig strukturierter Concept Maps im Vordergrund.

Hypothesen

Basierend auf dem skizzierten Stand der Forschung ergeben sich die folgenden Hypothesen:

Im Vergleich zu Noviz*innen zeichnen sich Expert*innen bei der Rezeption von Concept Maps im Mittel durch eine längere Gesamtbetrachtungsdauer (**H1.1**), kürzere Fixationsdauern (**H1.2**), höhere Fixationszahlen (**H1.3**), kürzere Scanpaths (**H1.4**) sowie längere Sakkaden (**H1.5**) aus.

Expert*innen weisen im Mittel höhere Fixationsdauern auf relevanten Bereichen (Konzepte und Relationen; **H2.1**) bzw. geringere Fixationsdauern auf irrelevanten Bereichen von Concept Maps (White Spaces; **H2.2**) auf als Noviz*innen.

Expert*innen weisen im Mittel sowohl für die hierarchisch strukturierte als auch für die netzartig strukturierte Concept Map eine kürzere Fixationsdauer auf als Noviz*innen (**H3.1**). Aufgrund der höheren Strukturiertheit, die hierarchisch strukturierte Concept Maps den Lernenden bieten, wird ein *Aptitude-Treatment-Interaction Effect* erwartet: So sollten sich die mittleren Fixationsdauern von Expert*innen und Noviz*innen bei der Rezeption von Informationen aus hierarchisch strukturierten Concept Maps weniger stark voneinander unterscheiden als bei der Rezeption von Informationen aus netzartig strukturierten Concept Maps (**H3.2**).

Material und Methode

Stichprobe

Es wurden 62 Schüler*innen der Sekundarstufe I (achte und neunte Jahrgangsstufe) rekrutiert. Aufgrund von Kalibrierfehlern und aus schulorganisatorischen Gründen (Nichterscheinen zum Messtermin oder vorzeitiger Abbruch der Messung) mussten zehn Schüler*innen ausgeschlossen werden. Die finale Stichprobe bestand aus $N = 52$ Schüler*innen (69% weiblich; $M_{\text{Alter}} = 13.9$ Jahre, $SD = 0.52$). Die Gruppe der Expert*innen bestand aus $n = 14$ Schüler*innen, während die Gruppe der Noviz*innen $n = 38$ Schüler*innen umfasste.

Instrument und Ablauf

In einer vorausgehenden Trainingsstudie erhielt eine Gruppe von Schüler*innen ein intensives Concept Map-Training (Expert*innen), während eine zweite Gruppe lediglich an einer Kurzeinführung über Concept Maps teilnahm (Noviz*innen; Kontrollgruppe). Im Anschluss an die Trainingsstudie wurden die Schüler*innen eingeladen an der vorliegenden Eye-Tracking Studie⁶² teilzunehmen.

Nach einer kurzen Begrüßung und Aufklärung über den Ablauf erhielten die Schüler*innen Instruktionen zur Eye-Tracking Messung. Anschließend wurde das Infrarotlicht-basierte Eye-Tracking System *SMI RED250 mobile* (*SensoMotoric Instruments GmbH*, Teltow, Deutschland) auf einem 17-Zoll-Bildschirm (Auflösung 1920 x 1080 Pixel) für jede*n individuelle*n Schüler*in durch 5-Punkt-Kalibrierung kalibriert. Aufgrund einer *Accuracy* von 0,4° mussten die Kopfbewegungen laut Hersteller nicht eingeschränkt werden. Die Daten wurden bei 60 Hz aufgezeichnet und der Betrachtungsabstand betrug ca. 40–50 cm. Die Schüler*innen wurden zufällig einem von zwei Settings (A und B) zugeordnet. Beide Settings waren inhaltlich identisch und bestanden aus zwei Concept Maps. Bevor diese den Schüler*innen gezeigt wurden, wurden sie durch einen Einstieg/Problemstellung (vgl. **Abbildung 2**) in einen sinnvollen Kontext eingebettet. Unmittelbar nach der Beschäftigung mit jeder Concept Map folgte ein Kurztest, in dem die Lernenden Fragen über die Inhalte der jeweiligen Concept Map beantworten mussten (siehe **Abbildung 2**). Die Kurztests wurden jeweils vor der Rezeption der Concept Maps angekündigt und sollten eine konzentrierte/ernsthafte Auseinandersetzung der Schüler*innen mit den Maps unterstützen. Sie wurden nicht weiter in die Auswertung miteinbezogen und stellen kein Lernerfolgsmaß dar. Die Settings A und B unterschieden sich in der Reihenfolge der Concept Maps. In Setting A wurde die netzartig strukturierte Concept Map (zum Thema „Viren“) zuerst und nachfolgend die hierarchisch strukturierte Concept Map (zum Thema „Zellen“) präsentiert. In Setting B erfolgte die Präsentation in umgekehrter Reihenfolge.

⁶² Bei der Umsetzung wurden die ethischen Standards der Deklaration von Helsinki beachtet (u. a. Information der Schüler*innen, Schulleiter*innen und Erziehungsberechtigten; Arbeit mit anonymisierten Daten; risikofreier Charakter der Studie [Taupitz, 2001]).

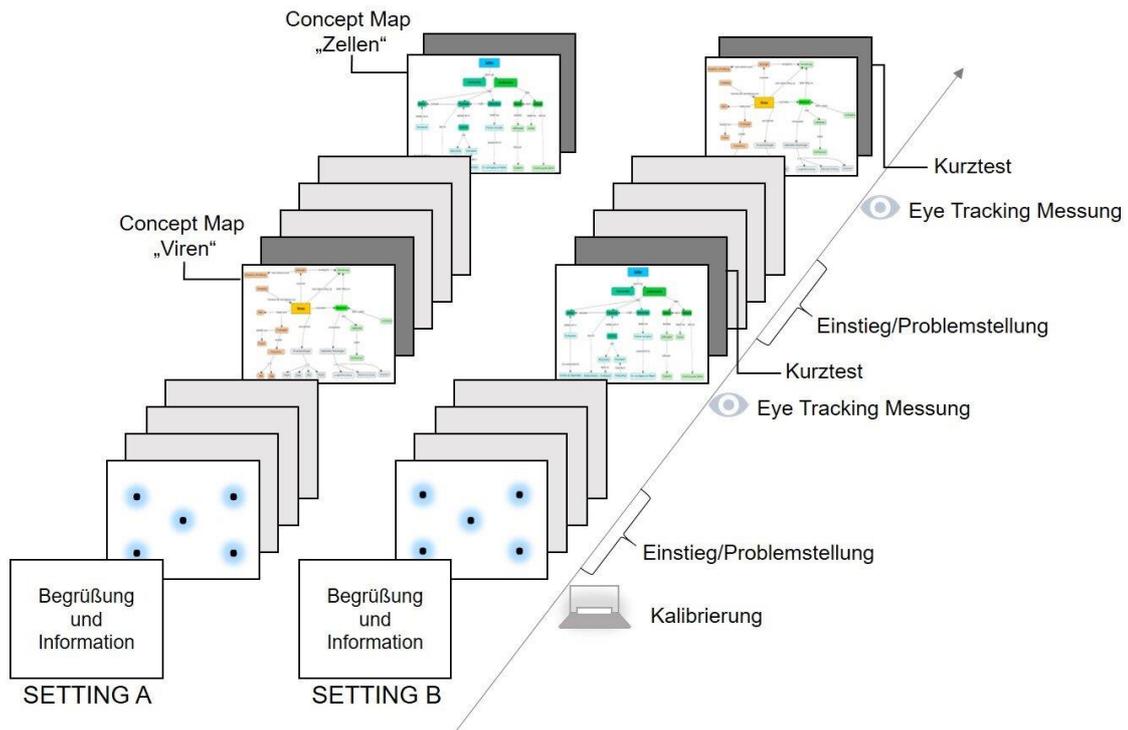


Abbildung 2. Design und Ablauf der Eye-Tracking Messung. Die Schüler*innen wurden einem von zwei Settings (A und B) zugeordnet. Die netzartig und die hierarchisch strukturierten Concept Maps zum Thema „Viren“ bzw. „Zellen“ wurden in den beiden Settings in unterschiedlicher Reihenfolge präsentiert.

Dieses Design wurde gewählt, um Reihenfolgeeffekte zu vermeiden (Causse et al., 2019). Nachdem die Messung von der/dem Testleiter*in gestartet wurde, konnten die Schüler*innen die ausgearbeiteten Concept Maps entsprechend ihrer eigenen Logik und ihres eigenen Zeitbedarfs ohne weitere strukturierende Vorgaben selbstgesteuert rezipieren und den Beginn des Kurztests bestimmen. Die Eye-Tracking Daten wurden lediglich für die Rezeption der Concept Maps, nicht aber für den Einstieg / die Problemstellung und die Kurztests (siehe **Abbildung 2**) erhoben (minimale Messdauer: 0.39 min, maximal Messdauer: 4.54 min; siehe Gesamtbetrachtungsdauer, **Tabelle 1**). Die gesamte Untersuchung nahm etwa 15–20 Minuten in Anspruch. Um identische Versuchsbedingungen während der Eye-Tracking Messung zu gewährleisten, wurden der Ablauf der Untersuchung und der Versuchsaufbau standardisiert (z. B. Anpassung der Lichtverhältnisse).

Stimulusmaterial

Das Stimulusmaterial bestand aus zwei Concept Maps, welche basierend auf den allgemeinen Prinzipien von Novak und Gowin (1984) mit dem Softwarepaket *Cmap-Tools*⁶³ konstruiert wurden. Die Concept Map „Zellen“ wurde hierarchisch (21 Konzepte und 19 Relationen) und die Concept Map „Viren“ netzartig strukturiert (24 Konzepte und 18 Relationen; siehe **Abbildung 1**).

Concept Map-Training und Kontrolltraining

Das Concept Map-Training und das Kontrolltraining umfassten insgesamt 135 Minuten und bestanden aus zwei Teilen. Eine 25-minütige Kurzeinführung stellte den ersten Teil dar (für mehr Informationen siehe Lenski et al., in Vorbereitung). In diesem Teil wurde deklaratives Strategiewissen über Concept Maps vermittelt (u. a. Aufbau und Nutzen von Concept Maps). Im zweiten Teil wurden die Expert*innen in der Nutzung von Concept Maps intensiv trainiert, wobei prozedurales Wissen über die Nutzung von Concept Maps vermittelt wurde. Das verwendete Trainingsmaterial deckt Themen verschiedener nicht-biologischer Inhaltsbereiche ab (Lenski & Großschedl, 2021). Die Noviz*innen nahmen stattdessen an einem Kontrolltraining⁶⁴ teil, das neben der Kurzeinführung keine weitere Lerngelegenheit für das Lernen mit Concept Maps vorsah.

Datenanalyse

Zur Datenaufbereitung und -analyse wurden die Messungen von Setting A und Setting B mittels *BeGaze* in einen Datensatz zusammengeführt. In Anlehnung an Lykins, Meana und Kambe (2006) wurde nur die Position des rechten Auges analysiert. *IViewX* wurde zur Datenerfassung verwendet. Fixierungen wurden unter Verwendung des *SMI*-Standardalgorithmus auf Dispersionsbasis (100 Pixel und Mindestdauer von 80 ms) erkannt.

⁶³ <https://cmap.ihmc.us>

⁶⁴ Im Kontrolltraining wurde das Spiel „Eierfall“ gespielt. Dazu wurden die Schüler*innen in Gruppen eingeteilt und entwickelten gemeinsam und unter Zeitdruck einen Plan, um ein rohes Ei für einen Wurf aus vier Metern vorzubereiten. Ziel war es, kooperatives Arbeiten zu fördern. Gemeinsam sollten Ideen und Lösungen entwickelt werden, um das Ei mit den vorgegebenen Materialien vor dem Sturz zu schützen. Diese Beschäftigung sollte sicherstellen, dass das Kontrolltraining den gleichen zeitlichen Umfang einnahm wie das Concept Map-Training.

Zur Testung der Hypothesen **H1.1–H1.5** wurde die Gesamtbetrachtungsdauer, die Fixationsdauer, die Fixationszahl, die Länge der Scanpaths und die Sakkadenlänge für beide Concept Maps gemittelt.

Um **H2.1** zu testen, wurden die Fixationsdauern auf einem zuvor definierten AOI-Set, das ausschließlich relevante Bereiche (Relationen und Konzepte) markierte, analysiert. Hierbei wurden Überlappungen von AOIs vermieden und Abstände zwischen den verschiedenen AOIs gelassen (Holmqvist et al., 2011). Die Bereiche, welche nicht in den AOIs enthalten waren (White Spaces), wurden für die Berechnung der Fixationsdauern auf irrelevanten Bereichen (**H2.2**) verwendet.

Zur Beantwortung von Hypothese **H3.1** und **H3.2** wurden die Fixationsdauern auf der netzartigen und der hierarchischen Concept Map getrennt voneinander betrachtet.

Deskriptiv- und inferenzstatistische Auswertung

Obwohl Kolmogorov-Smirnov-Tests nur bei einigen Variablen auf eine Verletzung der Normalverteilung hindeuteten, legten weitere Einschränkungen der Daten (z. B. Verletzung der Varianzhomogenität der Gruppen, stark variierende Gruppengrößen) eine Verwendung nichtparametrischer Verfahren für die Hypothesentestung nahe (Bortz & Lienert, 2008). So wurden die unabhängigen Stichproben der Expert*innen und Noviz*innen mittels Mediantest verglichen. Der Mediantest ist dem Mann-Whitney-U-Test bei unterschiedlich großen Stichproben vorzuziehen, da der Mann-Whitney-U-Test an Schärfe verliert, wenn die zu vergleichenden Stichproben unterschiedlich groß sind (Bortz & Lienert, 2008). Zur Vermeidung einer Alpha-Fehler-Kumulierung durch mehrfaches Testen innerhalb der gleichen Stichprobe wurde eine Bonferroni-Holm-Korrektur vorgenommen (Hemmerich, 2021). Dies betrifft in der vorliegenden Studie all jene Vergleiche, die sich auf Gruppenunterschiede in der Fixationsdauer beziehen (**H1.2**, **H2.1**, **H2.2**, **H3.1** und **H3.2**) und somit der Untersuchung übergreifender Hypothesen mittels verwandter Variablen dienen. Zur Beantwortung von **H3.2** wurde im Rahmen eines Solomon-2-Gruppen-Plans ein Mediantest für Paardifferenzen mit der unabhängigen Variable „Gruppe“ (Concept Map-Training vs. Kontrolltraining) durchgeführt. Um die Fixationsdauern auf hierarchisch und netzartig strukturierten Concept Maps vergleichbar zu machen, wurden die zugehörigen Variablen zunächst z-standardisiert (Bortz & Schuster, 2011). Anschließend wurde die abhängige Variable berechnet, die sich aus der Differenz

der z-standardisierten Variablen ergab. Effektstärken wurden nach Cohen (1984) berechnet und interpretiert.

Ergebnisse

H1.1 Vergleicht man die Gesamtbetrachtungsdauer von Expert*innen und Noviz*innen fällt auf, dass Expert*innen die Concept Maps insgesamt länger betrachteten (siehe **Tabelle 1**) als Noviz*innen. Dieser Unterschied erwies sich hypothesenkonform als statistisch bedeutsam, wobei der Effekt als groß zu interpretieren ist, $\chi^2(1) = 8.17$, $p = .004$, $d = 1.63$.

H1.2 Expert*innen wiesen bei der Rezeption der Concept Maps eine kürzere Fixationsdauer auf als Noviz*innen (siehe **Tabelle 1**). Dieser Unterschied erwies sich jedoch als nicht signifikant ($\chi^2[1] = 4.79$, $p = .087$ [nach Bonferroni-Holm korrigiert]), sodass diese Hypothese verworfen werden muss.

H1.3 Hypothesenkonform wiesen Expert*innen eine höhere Fixationszahl (siehe **Tabelle 1**) auf als Noviz*innen, $\chi^2(1) = 7.92$, $p = .005$, $d = 0.95$. Der gefundene Effekt ist als groß zu bewerten.

H1.4 Konträr zu dieser Hypothese zeigten Expert*innen signifikant längere Scanpaths in Pixeln (px; siehe **Tabelle 1**) als Noviz*innen, $\chi^2(1) = 7.92$, $p = .005$, $d = 0.87$). Auch hier liegt ein großer Effekt vor.

H1.5 Bezüglich der Sakkadendauer (siehe **Tabelle 1**) traten keine signifikanten Unterschiede zwischen Expert*innen und Noviz*innen auf ($\chi^2[1] = 0.23$, $p = .630$), sodass die Hypothese an dieser Stelle wiederum verworfen werden muss.

H2.1 Wider Erwarten unterschieden sich Expert*innen und Noviz*innen nicht hinsichtlich ihrer mittleren Fixationsdauern auf den relevanten Bereichen der Concept Maps (siehe **Tabelle 1**; $\chi^2[1] = 4.79$, $p = .087$ [nach Bonferroni-Holm korrigiert]).

H2.2 Ebenfalls erwartungswidrig unterschieden sich Expert*innen und Noviz*innen nicht hinsichtlich ihrer mittleren Fixationsdauern auf den irrelevanten Bereichen (White Spaces) der Concept Maps ($\chi^2[1] = 4.10$, $p = .755$ [nach Bonferroni-Holm korrigiert]); siehe **Tabelle 1**).

Tabelle 1. Deskriptive Daten von Expert*innen und Noviz*innen

Maß	Expert*innen (n = 14)		Noviz*innen (n = 38)	
	M	SD	M	SD
Gesamtbetrachtungsdauer (ms)	138222.16	9277.33	15072.53	44515.99
Fixationsdauer (ms)	373.43	59.88	456.04	163.15
Fixationsdauer (ms) auf relevanten Bereichen ^A	379.75	66.45	479.44	172.80
Fixationsdauer (ms) auf irrelevanten Bereichen ^B	359.16	128.08	366.59	186.44
Fixationsdauer (ms) auf hierarchischen Concept Maps	394.37	76.16	493.48	212.26
Fixationsdauer (ms) auf netzartigen Concept Maps	352.48	48.85	418.61	145.23
Fixationszahl (n)	335.82	171.20	183.49	147.74
Scanpath-Länge (px)	47857.54	21232.87	30308.00	19196.41
Sakkadendauer (ms)	24.82	1.98	41.35	34.03

Anmerkung. ^A Verwendung eines AOI-Sets, welches relevante Bereiche (Relationen und Konzepte) umfasst. ^B Bereiche außerhalb der relevanten Bereiche (White Spaces; für mehr Informationen siehe 4.5).

H3.1 Expert*innen wiesen auf hierarchisch strukturierten Concept Maps eine signifikant kürzere Fixationsdauer auf als Noviz*innen ($\chi^2[1]=7.92, p = .025$ [nach Bonferroni-Holm korrigiert], $d = 0.61$; siehe Tabelle 1). Dies konnte auch für die netzartig strukturierte Concept Map ermittelt werden ($\chi^2[1]=7.92, p = .025$ [nach Bonferroni-Holm korrigiert], $d = 0.61$; siehe Tabelle 1).

H3.2 Es wurde erwartet, dass sich die mittleren Fixationsdauern von Expert*innen und Noviz*innen bei der Rezeption von Informationen aus hierarchisch strukturierten Concept Maps weniger stark voneinander unterscheiden als bei der Rezeption von Informationen aus netzartig strukturierten Concept Maps. Ein Mediantest für Paardifferenzen ergab, dass dies nicht der Fall war, weshalb wir unsere Hypothese verwerfen ($\chi^2[1]=2.44, p = .118$ [nach Bonferroni-Holm korrigiert]).

Insgesamt ist anzumerken, dass die Werte der Noviz*innen für die meisten Parameter eine größere Streuung aufwiesen als die der Expert*innen (vgl. Tabelle 1).

Diskussion

Interpretation der Ergebnisse

In dieser Studie wurde untersucht, ob sich die während eines Concept Map-Trainings erworbene Expertise bei der Rezeption von Concept Maps in Blickbewegungscharakteristika manifestiert. Hierzu wurden die Blickbewegungen von trainierten Schüler*innen (Expert*innen) und untrainierten Schüler*innen (Noviz*innen) verglichen

(H1.1–H1.5). Weiterhin wurde untersucht, ob sich Expert*innen und Noviz*innen in der Betrachtung relevanter und irrelevanter Concept Map-Bereiche (H2.1+H2.2) und unterschiedlich strukturierter Concept Maps unterscheiden (H3.1+3.2).

Der Vergleich ausgewählter Blickbewegungen ergab hypothesenkonform, dass Expert*innen die Concept Maps länger betrachteten (H1.1) und sich durch höhere Fixationszahlen (H1.3) auszeichneten. Dieser Befund steht in Einklang mit metaanalytischen Befunden von Gegenfurtner et al. (2011) zum Betrachten von Visualisierungen. Die längere Betrachtungsdauer lässt darauf schließen, dass sich Expert*innen ausführlicher mit den bereitgestellten Concept Maps auseinandersetzten. Die höhere Fixationszahl lässt sich mit Hilfe der Theorie des *Long-Term Working Memory* erklären, wonach sich Expert*innen durch eine schnellere Informationsverarbeitung auszeichneten als Noviz*innen. Der erwartete Unterschied in der Fixationsdauer zwischen Expert*innen und Noviz*innen deutet sich zwar an, erwies sich jedoch als nicht signifikant (H1.2).

Anders als erwartet, wiesen Expert*innen längere Scanpaths auf als Noviz*innen (H1.4). Dies steht im Widerspruch der Theorie des *Long-Term Working Memory* und der hieraus abgeleiteten Hypothese. Gestützt wird der Befund jedoch durch andere empirische Studien (u. a. Jaarsma, Jarodzka, Nap, van Merriënboer & Boshuizen, 2014), bei denen ebenfalls längere Scanpaths mit steigendem Expertisegrad einhergingen. Geschlussfolgert wurde hier, dass Noviz*innen nicht „das ganze Bild“, sondern nur Teile davon verarbeiteten. Dies würde mit den Annahmen des *Holistic Model of Image Perception* übereinstimmen und könnte in der vorliegenden Untersuchung ein Anzeichen für einen Map Shock sein, welcher sich laut Blankenship und Dansereau (2000) in einer unvollständigen Betrachtung der Map äußern kann.

Im Widerspruch zu diesem Modell und der hierauf basierenden Hypothese steht das Ergebnis, dass Expert*innen kürzere Sakkadendauern aufwiesen als Noviz*innen (H1.5). Kok und Jarodzka (2017) merkten allerdings an, dass die Sakkadendauer von Expert*innen und Noviz*innen je nach Art der visuellen Aufgabe variiert und stellten die Eignung der Sakkadendauer zur Bestimmung des Expertisegrades in Frage.

In Anlehnung an die *Information-Reduction Hypothesis* wurde erwartet, dass sich Expert*innen und Noviz*innen in ihren mittleren Fixationsdauern auf relevanten (H2.1) und irrelevanten Bereichen der Concept Maps (H2.2) unterscheiden. In beiden Fällen

konnten jedoch keine entsprechenden Unterschiede beobachtet werden. Um Expertiseunterschiede sichtbar zu machen, reichte es wahrscheinlich nicht aus, lediglich zwischen Teilen der Concept Map und White Spaces zu unterscheiden. Hier hätten eventuell in Bezug auf die Inhalte der Concept Maps genauere Aufgabenstellungen formuliert und dazu passende AOIs definiert werden müssen.

Betrachtet man die Fixationsdauern von Expert*innen und Noviz*innen auf den hierarchischen und netzartigen Concept Maps, wird eine erwartungskonform kürzere Fixationsdauer der Expert*innen sichtbar (H3.1). Dieser Unterschied wird jedoch schwächer, wenn keine differenzierte Analyse nach Concept Map-Struktur erfolgt (siehe H1.2). Da sich dies weder theoretisch noch empirisch herleiten lässt, könnte es durch die geringe Stichprobengröße und hiermit verbundene geringe Power der verwendeten statistischen Verfahren erklärt werden.

Hinsichtlich der mittleren Fixationsdauern auf hierarchisch und netzartig strukturierten Concept Maps wurde ein *Aptitude-Treatment-Interaction Effect* erwartet (H3.2). Mittlere Fixationsdauern von Expert*innen und Noviz*innen sollten sich bei der Rezeption von Informationen aus hierarchisch strukturierten Concept Maps weniger stark voneinander unterscheiden als bei der Rezeption von Informationen aus netzartig strukturierten Concept Maps. Dies war jedoch nicht der Fall, was darauf hindeutet, dass Noviz*innen nicht stärker von der hierarchischen Struktur profitierten als Expert*innen.

Zu erwähnen ist, dass viele Messwerte der Noviz*innen eine größere Streuung aufwiesen als die der Expert*innen. Einerseits könnte dies als Artefakt interpretiert werden, das sich aus einer größeren Stichprobe von Noviz*innen ergibt. Andererseits könnten geringere Streuung der Messwerte auf Expert*innen-Seite auch auf geringere Kompetenzunterschiede in der Rezeption von Informationen aus Concept Maps hindeuten, was auf die Wirksamkeit des Trainings hindeuten könnte. Sollten sich die Befunde dieser Studie in nachfolgenden Studien erhärten, könnten auf die Expertise abgestimmte Hinweise zur praktischen Nutzung von Concept Maps abgeleitet werden.

Limitationen der vorliegenden Studie

Eine Limitation dieser Studie stellt die unterschiedlich große Stichprobenverteilung in den Gruppen dar ($n = 38$ Noviz*innen vs. $n = 14$ Expert*innen). Diese ergab sich hauptsächlich durch ein unterschiedlich großes Interesse an der Studien-Teilnahme. Wie

bei vielen Eye-Tracking Studien ist einschränkend zu erwähnen, dass die Power der angewandten statistischen Verfahren vergleichsweise gering ist (Gegenfurtner et al., 2011). Wir empfehlen daher eine Replikation der vorliegenden Studie an einer größeren und für die Populationen der Expert*innen und Noviz*innen repräsentativeren Stichprobe. Da in dieser Studie durch Vorabgespräche mit den Lehrkräften sichergestellt wurde, dass die Inhalte der Concept Maps noch nicht im Unterricht behandelt wurden, jedoch kein Vorwissenstest durchgeführt wurde, sollte in künftigen Studien das fachliche Vorwissen einbezogen werden.

Fazit und Implikationen für zukünftige Forschung

Im Rahmen der vorliegenden Studie wurden wichtige Hinweise dafür gefunden, dass Schüler*innen, die ein Concept Map-Training besuchten, eine größere Expertise in der Rezeption von Informationen aus Concept Maps aufwiesen als Schüler*innen, die im Rahmen eines Kontrolltrainings lediglich eine Kurzeinführung erhielten. Dieses Ergebnis deckt sich auch in den Leistungsdaten der vorausgehenden Trainingsstudie (Lenski et al., in Vorbereitung) – trainierte Schüler*innen erzielten eine höhere Lernleistung als untrainierte Schüler*innen. Gemeinsam mit diesen Befunden, untermauert die vorliegende Studie die Annahme, dass ein Concept Map-Training die Expertise in der Rezeption von Concept Maps steigert und diese Expertise auch notwendig ist, um Concept Maps gewinnbringend nutzen zu können. Sie leistet somit einen wichtigen Beitrag zur Debatte um die Wirkung und Relevanz von Trainingsmaßnahmen für die Nutzung von Concept Maps und trägt zu einem wichtigen Erkenntnisfortschritt an der Schnittstelle zwischen Expertiseforschung und anwendungsorientierter biologiedidaktischer Forschung bei.

Neben quantitativen Untersuchungen mit repräsentativeren Stichprobenzahlen, könnten sich zukünftige Studien der qualitativen Untersuchungen der Unterschiede in den Blickbewegungen von Expert*innen und Noviz*innen bei der Rezeption von Concept Maps widmen. Hier wäre insbesondere interessant, ob sich bei Noviz*innen weitere Hinweise für einen Map Shock finden lassen. Es könnten zudem nach für Experten typische Scanpaths gesucht werden mit dem Ziel diese in Form von sogenannten *Eye Movement Modelling Examples* (van Marlen, van Wermeskerken, Jarodzka & van Gog, 2018) Noviz*innen zur Verfügung zu stellen. Perspektivisch könnten Noviz*innen so bei der Navigation durch die Inhalte der Concept Maps unterstützt werden.

Das in dieser Studie eingesetzte Concept Map-Konstruktionstraining samt aller Materialien ist online verfügbar (Lenski & Großschedl, 2021).

Danksagung und Angaben zur finanziellen Förderung

Diese Studie ist Teil des Projekts „Biologie lernen durch Concept Mapping: Bedeutung eines Lernstrategietrainings für kognitive Belastung, kognitive Prozesse und Lernleistung“, welches von der Deutschen Forschungsgemeinschaft (DFG) gefördert wurde (Projektnummer GR 4763/2). Für die Unterstützung bei der Datenerhebung bedanken wir uns bei unseren Abschlussarbeitskandidat*innen Benjamin Bayer, Sven Nam Karsten und Regina Drevs.

Literaturverzeichnis

- Amadiou, F. & Salmerón, L. (2014). Concept maps for Comprehension and Navigation of Hypertexts. In D. Ifenthaler & H. R. (Hrsg.), *Digital Knowledge Maps in Education* (S. 41-59). New York: Springer.
- Amadiou, F., van Gog, T., Paas, F., Tricot, A. & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction* 19, 376-386. <https://doi.org/https://doi.org/10.1016/j.learninstruc.2009.02.005>
- Blankenship, J. & Dansereau, D. F. (2000). The effect of animated node-link displays on information recall. *Journal of Experimental Education*, 68(4), 293-308. <https://doi.org/https://doi.org/10.1080/00220970009600640>
- Bortz, J. & Lienert, G. A. (2008). *Kurzgefasste Statistik für die klinische Forschung: Leitfaden für die verteilungsfreie Analyse kleiner Stichproben*. Berlin: Springer-Verlag.
- Bortz, J. & Schuster, C. (2011). *Statistik für Human-und Sozialwissenschaftler*. Berlin, Heidelberg: Springer-Verlag.
- Cause, M., Lancelot, F., Maillant, J., Behrend, J., Cousy, M. & Schneider, N. (2019). Encoding decisions and expertise in the operator's eyes: Using eye-tracking as input for system adaptation. *International Journal of Human-Computer Studies*, 125, 55-65.
- Cohen, A. (1984). Effect of girdling date on fruit size of Marsh Seedless grapefruit. *Journal of Horticultural Science*, 59(4), 567-573.
- Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington Publishers.
- Dogusoy-Taylan, B. & Cagiltay, K. (2014). Cognitive analysis of experts' and novices' concept mapping processes: An eye tracking study. *Computers in Human Behavior*, 36, 82-93.
- Ericsson, K. A. & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211.
- Fischer, F. & Mandl, H. (2000). Strategiemodellierung mit Expertenmaps. In H. Mandl & F. Fischer (Hrsg.), *Wissen sichtbar machen. Wissensmanagement mit Mapping-Techniken* (S. 37-54). Göttingen: Hogrefe-Verlag.
- Funke, J. & Spering, M. (2006). Methoden der Denk- und Problemlöseforschung. In J. Funke (Hrsg.), *Denken und Problemlösen (=Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie II: Kognition, Band 8* (S. 675-744). Göttingen: Hogrefe.
- Gegenfurtner, A., Lehtinen, E. & Säljö, R. (2011). Expertise Differences in the Comprehension of Visualizations: A Meta-analysis of Eye-Tracking Research in Professional Domains. *Educational Psychology Review*, 23(4), 523-552.
- Großschedl, J. & Tröbst, S. (2018). Biologie lernen durch Concept Mapping: Bedeutung eines Lernstrategietrainings für kognitive Belastung, kognitive Prozesse und

- Lernleistung – Kurzdarstellung des DFG-Projekts. *Zeitschrift für Didaktik der Biologie (ZDB) - Biologie Lehren und Lernen*, 22(1), 20-30.
- Haider, H. & Frensch, P. A. (1999). Eye movement during skill acquisition: more evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(1), 172-190.
- Hemmerich, W. (2021). *Rechner zur Adjustierung des α -Niveaus*. StatistikGuru. Verfügbar unter: <https://statistikguru.de/rechner/adjustierung-des-alphaniveaus.html> [Zugriff: 11. Januar].
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*: Oxford University Press.
- Jaarsma, T., Jarodzka, H., Nap, M., van Merriënboer, J. J. G. & Boshuizen, H. P. A. (2014). Expertise under the microscope: Processing histopathological slides. *Medical Education*, 48(3), 292-300.
- Just, M. A. & Carpenter, P. A. (1984). Using eye fixations to study reading comprehension. In D. E. Kieras & M. A. Just (Hrsg.), *New Methods in Reading Comprehension Research* (S. 151-182). Hillsdale: Erlbaum.
- Kok, E. M. & Jarodzka, H. (2017). Before your very eyes: The value and limitations of eye tracking in medical education. *Medical Education*, 51(1), 114-122.
- Krupinski, E. A. (2000). The importance of perception research in medical imaging. *Radiation Medicine - Medical Imaging and Radiation Oncology*, 18(6), 329-334.
- Kundel, H. L., Nodine, C. F., Conant, E. F. & Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: gaze-tracking study. *Radiology*, 242(2), 396-402.
- Lenski, S., Elsner, S. & Großschedl, J. (in Vorbereitung). Working with concept maps in biology learning – development and evaluation of a construction and a study training.
- Lenski, S. & Großschedl, J. (2021). Concept Maps im Unterricht: eine Trainingseinheit für Schüler*innen der Sekundarstufe I. <https://doi.org/https://doi.org/10.17605/OSF.IO/48A5W>
- Lykins, A. D., Meana, M. & Kambe, G. (2006). Detection of differential viewing patterns to erotic and non-erotic stimuli using eye-tracking methodology. *Archives of Sexual Behavior*, 35(5), 569-575.
- Mintzes, J. J., Canas, A., Coffey, J., Gorman, J., Gurley, L., Hoffman, R. et al. (2011). Comment on "Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping" [Technical Comment]. *Science*, 334, 453-c - 453-d.
- Moser, T., Lohmeyer, Q., Meboldt, M., Distler, O. & Becker, M. O. (2020). Visual assessment of digital ulcers in systemic sclerosis analysed by eye tracking: implications for wound assessment. *Clinical and Experimental Rheumatology*, 38(3), 137-139.
- Novak, J. D. & Gowin, D. B. (1984). *Learning how to learn*. New York: Cambridge University Press.

- Patel, V. L. & Groen, G. J. (1986). Knowledge based solution strategies in medical reasoning. *Cognitive Science*, 10(1), 91-116.
- Sheridan, H. & Reingold, E. M. (2014). Expert vs. novice differences in the detection of relevant information during a chess game: evidence from eye movements. *Frontiers in Psychology*, 5, 941.
- Taupitz, J. (2001). Die Neufassung der Deklaration von Helsinki des Weltärztebundes vom Oktober 2000. *MedR-Medizinrecht*, 19(6), 277-286.
- van Marlen, T., van Wermeskerken, M., Jarodzka, H. & van Gog, T. (2018). Effectiveness of eye movement modeling examples in problem solving: The role of verbal ambiguity and prior knowledge. *Learning and Instruction*, 58, 274-283.
- Wendt, H., Bos, W., Selter, C., Köller, O., Schwippert, K. & Kasper, D. (2016). *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. New York: Waxmann.

9 Die Wirkung von Emotional Design-Abbildungen

Die **Publikationen III** und **IV** befassen sich mit dem Einfluss unterschiedlicher Gestaltungsformen auf kognitive und affektive Variablen. **Publikation III** widmete sich der Übertragung der *Emotional Design*-Prinzipien auf *Concept Maps*, während in **Publikation IV** psychophysiologische Messungen durchgeführt wurden, um tiefergehende Erkenntnisse zu dem Einfluss verschiedener Gestaltungsformen zu gewinnen.

9.1 Emotional Design in concept maps – No support but also no burden (Publikation III)

9.1.1 Zusammenfassung, Fragestellungen und methodischer Zugang

In der Forschung zum multimedialen Lernen wird davon ausgegangen, dass bestimmte Designmanipulationen (*Emotional Design*) den affektiv-motivationalen Zustand der Lernenden beeinflussen und dadurch das Lernen fördern können (Um et al., 2012). Im Rahmen dieser Studie wurde untersucht, ob die Integration von *Emotional Design*-Abbildungen die Lernförderlichkeit von *Concept Maps* verbessern kann. Hierzu wurden zwei Experimente durchgeführt. In Experiment 1 wurde eine Felduntersuchung im experimentellen Design mit $n = 249$ Schüler:innen durchgeführt. Es wurde die Wirkung von *Concept Maps* mit *Emotional Design*-Abbildungen (emotionales Design; ED), mit nicht-emotional gestalteten Abbildungen (neutrales Design; ND) und ohne Abbildungen (Kontrolldesign; CD) in Bezug auf die Lernleistung verglichen. In Experiment 1 wurde die kognitive Wirkung untersucht, in Experiment 2 stand der affektive Zustand der Schüler:innen im Vordergrund ($n = 47$).

Die konkreten Fragestellungen, welche im Rahmen dieser Studie beantwortet werden sollen, lauten:

F₁: Verbessert die Integration von *Emotional Design*-Abbildungen in *Concept Maps* die Lernleistung?

F₂: Können durch die Integration von *Emotional Design*-Abbildungen in *Concept Maps* die lernhinderlichen Formen der kognitiven Belastung reduziert werden?

F₃: Werden durch die Integration von *Emotional Design*-Abbildungen in *Concept Maps* positive Emotionen induziert und negative reduziert?

9.1.2 Eigenanteile

Die Autorin der vorliegenden Dissertation war an der Konzeption der Studie maßgeblich beteiligt und hat sie durchgeführt. Bei der Erstellung des Erhebungsmaterials, der Durchführung der Studie und der Dateneingabe wurde sie von Abschlussarbeitskandidat:innen⁶⁵ unterstützt. Die Autorin führte die Datenanalyse durch, verfasste das Manuskript und reichte es ein. Das Manuskript wurde von Herrn Prof. Dr. Großschedl überarbeitet, welcher auch die Supervision übernahm.

⁶⁵ Benjamin Bayer, Florian Schmitz, Hanna Kosman, Sven Karsten, Leonie Mennicken, Rebecca Reiffen, und Tanja Treinen.

9.1.3 Publikation III: Originalmanuskript

Emotional Design in concept maps – No support but also no burden

Sina Lenski & Jörg Großschedl

Abstract

A concept map is a powerful method that promotes meaningful learning and is highly recommended for use in biology classes. According to multimedia research, the effectiveness of concept maps could be improved by incorporating pictorial elements. Apart from using realistic images, a new field of research claims that specific design manipulations, including human-like features with appealing colors (*Emotional Design*), influence learners' affective state and improve learning. A positive affective state is assumed to evoke emotions and provoke deeper cognitive processing, which increases the cognitive resources available for a task. We conducted two experiments with a total of $N = 249$ junior high school students, comparing the effect of concept maps with *Emotional Design* illustrations (*Emotional Design*), with non-*Emotional Design* illustrations (neutral design), and without illustrations (control design). Experiment 1 examined the influence of these designs on students' *perceived affective state*, *perceived cognitive load* (*extraneous*, *intrinsic*, and *germane load*), *perceived task difficulty*, and *learning performance* ($n = 202$), experiment 2 focused on the *perceived affective state* of the students ($n = 47$). We found that *Emotional Design* led to a significant decrease in *perceived task difficulty*, but we neither found an effect on *learning performance* nor the *positive affective state*. Learning with pictorial concept maps (in emotional or neutral design) reduced the *negative affect* compared to learning with control concept maps. Other than expected, the neutral design led to reduced *perceived extraneous* and *intrinsic cognitive load*. Consequently, in terms of learning, *Emotional Design* in concept maps did not hamper learning but did not foster it either.

Keywords: Concept maps, multimedia learning, *Emotional Design*, anthropomorphisms, emotions

Introduction

Modern and global society is exposed to various information that we must filter, evaluate, and interpret. Most of them can only be assessed correctly when relationships are understood. Understanding concepts and their relationships (*conceptual knowledge*) are essential in everyday life, especially in learning contexts. This poses a significant challenge for students, especially for the subject of biology (Schmid & Telaro, 1990), as it tries to elucidate how organisms in supersystems interact with or influence each other and their environment (Reiners et al., 2018). The complexity of the biological subject matter causes high element interactivity which means that different aspects highly interact and cannot be learned in isolation (Chen et al., 2015). To help students obtain the required conceptual knowledge and handle the high element interactivity (Williams, 1998), concept maps are recommended for biology classes (Kinchin, 2000).

Concept maps are graphical tools that make knowledge and relationships between different elements of a system visible (Novak, 1990) and can be used for various purposes (e.g., planning, teaching, learning, and diagnosing; Schroeder et al., 2017). For the study of concept maps, Blankenship and Dansereau (2000) found that learners could experience a so-called *map shock* characterized by ‘bewilderment of not knowing where to start or how to penetrate the topography of the map’ (p. 294). Researchers assume that these cognitive challenges could produce an adverse affective reaction that demotivates and inhibits learning (Dees et al., 1994). A possible solution to this problem could be to enrich the text-based nodes of the concept maps with illustrations (Alpert & Grueneberg, 2001). According to the *cognitive theory of multimedia learning* (Mayer, 2001), the combined presentation of illustration and text can reduce cognitive load (e.g., Levie & Lentz, 1982).

A new research field assumes that distinct design manipulations (*Emotional Design*) influence learners’ affective-motivational state (e.g., positive emotions in general, situational interest, intrinsic motivation, and enjoyment) and thereby foster learning (Um et al., 2012). *Emotional Design* is characterized by human-like features, round shapes, and appealing colors. Several studies have been carried out to investigate the effect of *Emotional Design* on learning, affective state, and mental effort (Park et al., 2015; Plass et al., 2014; Plass & Kaplan, 2016; Um et al., 2012; Uzun & Yildirim, 2018). Some findings support the assumption that *Emotional Design* evokes emotions and enhances learning (Brom et al., 2018; Mayer & Estrella, 2014; Tien et al., 2018; Um et al., 2012; Wong

& Adesope, 2021), while other findings showed that *Emotional Design* did not impact emotions and learning (Münchow & Bannert, 2019; Park et al., 2015, Stárková et al., 2019). Overall, the meta-analysis by Brom et al. (2018) showed that *Emotional Design* improves *retention*, *comprehension*, and *transfer*. Furthermore, a reduced *perceived task difficulty*, a weak effect for *positive affect*, and no significant impact on *mental effort* were found. A recent meta-analysis supported these findings (Wong and Adesope, 2021). In *Emotional Design* research, mainly the positive scale of the positive and negative affect scale (PANAS) was used to investigate the affective impact of *Emotional Design*. Several other authors presume that the effect of *Emotional Design* can vary and may be dependent on factors such as context, learning domain, learning time, nature of the presentation, intensity of manipulation, or grade level (e.g., Münchow & Bannert, 2019; Stárková et al., 2019; Wong & Adesope, 2021).

The *Emotional Design* applied to concept maps has so far been little investigated. Tien et al. (2018) examined the effects of multidimensional concept maps (digital concept maps where videos, pictures, or sounds are linked to the concepts) on learning and the affective state in college students. They found that colorful and animated multidimensional concept maps lead to ‘higher learning well-being’ and better learning than achromatic multidimensional concept maps (Tien et al., 2018, p. 1). However, it should be noted that Tien et al. (2018) interpreted the *Emotional Design* differently than Um et al. (2012). Rather than using anthropomorphisms, they applied daily-life advertisement music, animations, and images to induce emotions. Thus, it is unclear whether *Emotional Design* features in concept maps (incorporation of colorful and anthropomorphized images into the concepts) increase the learning efficiency of concept maps or whether it increases cognitive load and impairs learning.

We examined whether concept maps in *Emotional Design* reduce inhibitory *perceived cognitive load*, influence the *perceived affective state*, and thereby improve *learning* of junior high school students in an authentic learning environment. The present study was designed for junior high school students as there is a discrepancy between the evidence-based potential of concept maps to improve students’ learning and the actual use in schools (Kinchin, 2001).

Theoretical Framework

Concept Maps

A concept map is a graphical tool that represents knowledge in a domain (Novak, 1990, see **Figure 1** for examples). External information (e.g., learning text and teaching) or own knowledge (retrieval practice; Blunt & Karpicke, 2014) guide the construction of a concept map. A concept map consists of textual elements (concepts) connected by labeled links that define their relationships. Arrowheads indicate the reading direction at one or both ends of the links. Combining at least two concepts connected via a link is called *proposition* and forms a meaningful unit (Novak & Gowin, 1984). If the concepts ‘chlorophyll’ and ‘leaf pigment’ are connected via the labeled arrow ‘is’, the proposition ‘chlorophyll is (a) leaf pigment’ can be formed. The number of propositions of a concept map is unlimited, but fifteen to twenty-five concepts are usually necessary to create a (good) concept map (Novak & Cañas, 2008). Novak and Cañas (2008) recommend structuring concept maps hierarchically. The concepts are sorted according to their relevance and essential terms at the top and more specific terms below in the concept maps (**Figure 1**).

To explain why the construction and the study of concept maps are effective for learning, Schroeder et al. (2017) classified three reasons: (1) concept maps promote meaningful learning, (2) concept maps reduce extraneous cognitive load, and (3) concept maps do both.

While learning with concept maps, existing knowledge structures are modified and linked with new information (*elaboration*), a fundamental component of *meaningful learning* (Renkl, 2010). According to Großschedl & Tröbst (2018), concept maps help organize new information and promote metacognitive learning. Metacognitive learning is encouraged since concept maps help to reflect individual learning progress and identify knowledge gaps (Mintzes et al., 1997). For studying concept maps, it is assumed that learners adopt the (experts’/teachers’) knowledge structure and integrate the information of the concept map into their knowledge (Gehl, 2013). The learning effectiveness of concept maps of both processes (constructing and studying) has been empirically confirmed by meta-analyses (Horton et al., 1993; Nesbit & Adesope, 2006; Schroeder et al., 2017).

Emotions, Emotional Design, and Learning

There are two opposing hypotheses about the possible effects of emotions on learning: the *emotions-as-facilitator-of-learning hypothesis* and the *emotions-as-suppressor-of-learning hypothesis* (Um et al., 2012). The *emotions-as-facilitator-of-learning hypothesis* suggests that experiencing positive emotions can enhance learning performance. According to Um et al. (2012), ‘experiencing positive emotions during the learning process can enhance learning outcomes, either through direct impact and learning or through mediating variables, such as interest and motivation’. Furthermore, (design) elements that elicit positive activating affective states are assumed to provoke deeper cognitive processing and thereby increase the cognitive resources dedicated to a task (see Stárková et al., 2019 for an explanatory graphic). By increasing available cognitive resources, aspects like affect and motivation can mediate the learning process at a higher level.

On the contrary, the *emotions-as-suppressor-of-learning hypothesis* postulates an impairing effect on learning. In this case, emotions are presumed to distract students from a task by, for example, processing information about one’s emotional state or other task-irrelevant thinking. This additional process competes for the limited capacity of the working memory (see Plass & Kalyuga, 2019 for a recent overview of the effect of emotions on cognitive processes).

Since studies on *Emotional Design* focus on the learners’ general emotional state rather than specific types of emotions, in this study, the generic term ‘affect’ is used to refer to emotion and mood (for a similar approach, see Park et al., 2015; Plass et al., 2014; Uzun & Yildirim, 2018). According to the *circumplex model of affect* by Russell (2003), emotions can be categorized in a two-dimensional system with *valence* as a continuum from positive to negative (affect) as one dimension and *activation* as a continuum from activated to deactivated as the other. Activating emotions are *enjoyment* or *curiosity* (positive) and *anger* (negative). Deactivating emotions include *relaxation* (positive) and *sadness* (negative; Pekrun et al., 2006) and are mainly related to an impairment of learning (e.g., Aspinwall, 1998). Empirically, positive affect is typically found to increase interest and motivation to learn, while negative affect is mainly shown to impair learning (see meta-analysis by Barroso et al., 2021). On the contrary, beneficial effects on learning for some negative emotions and inhibitory effects for positive affect has been found

(Münchow & Bannert, 2019). In a recent study by Mensink (2021), positive affect negatively predicted learning performance, while negative affect positively affected learning. In line, learners in a negative affective state before learning outperformed those in a positive affective state (Knörzer et al., 2016). Positive activating emotions are suspected of distracting from the learning material and acting detrimentally. Pekrun (2021) concluded that the type of effect (beneficial or inhibitory) of different emotions depends on, for example, the situation, the kind of task, and personal traits. Also, Knörzer et al. (2016) suspected that moderating trait variables are responsible for the inconsistent results. Beyond, it was found that the experience of multiple emotions (here, joy, anger, confusion, and frustration) could distract and negatively impact learning (Dever et al., 2021). The authors suspect that the processing of emotions causes additional cognitive load, reducing cognitive resources available for the learning task.

For multimedia designs, the effect of a design on the viewer ([visual] *appeal*), has been found to influence the affective state (Capota et al., 2007). *Emotional Design* manipulations are developed based on this assumption, however their effect on appeal has been little investigated so far (Mayer & Estrella, 2014). Contrary to expectations, learning material in *Emotional Design* was not perceived as more appealing compared to non-*Emotional Design* material (Mayer & Estrella, 2014). Besides examining the affective parameter *enjoyment*, Mayer and Estrella (2014) also assessed the *desire for additional lessons of a similar nature* which served as a proxy for positive affect. No difference between *Emotional Design* and non-*Emotional Design* material was found here either.

Cognitive Load and Multimedia Learning

In *Emotional Design* research, different theories are used to explain how *Emotional Design* influences learning and cognitive processes. Central to learning is the *cognitive load theory* (Sweller, 2005). It is assumed that the amount of information that can be processed simultaneously is restricted while different aspects can increase a person's cognitive load (*limited capacity assumption*; Mayer, 2001). Sweller (2005) distinguishes between three types of cognitive load, i.e., intrinsic, extraneous, and germane load. The intrinsic load is caused by the task or learning material and depends on the complexity of these and the learners' expertise. Consequently, it cannot be manipulated directly by instructional design. Germane load refers to the resources that learners invest in their learning process. In contrast, the extraneous load is not conducive for learning and arises from

a poor learning material design (e.g., unsuitable design). A *seductive detail*, for example, is assumed to impose extraneous cognitive load and can impair learning by competing with the learning object for cognitive resources (Harp & Mayer, 1997; Lenzner et al., 2013).

According to the *dual-channel assumption* by Paivio (1986), information is processed in verbal and visual channels. Based on this assumption, the *cognitive theory of multimedia learning* (Mayer, 2001) was developed, which describes multimedia as any material that consists of a combination of text and illustration elements. It is assumed to improve learning by reducing extraneous cognitive load. The text can be written or spoken, and illustrations are understood to include static pictures (e.g., illustrations, graphics, and photos) and moving pictures (e.g., videos and animations). The cognitive theory of multimedia learning is mainly focused on cognitive processes, but it was shown that emotional and motivational aspects and cognitive processes are inextricably linked (Plass & Kalyuga, 2019). The *cognitive-affective theory of learning with media* was established to extend the previous theory with non-cognitive elements (emotional and motivational aspects) which can influence the learning process (Moreno, 2006; Moreno & Mayer, 2005; Moreno & Mayer, 2007). It presumes that affect and motivation can mediate the learning process by manipulating the cognitive resources available for a specific task.

Two competing views exist regarding *Emotional Design* and its influence on extraneous cognitive load. As stated in section “**Emotions, Emotional Design, and Learning**”, emotions induced by *Emotional Design* could cause extra processing. This process is not considered relevant for the learning goal, it can be classified as extraneous cognitive load (Plass & Kalyuga, 2019). Another potential source of extraneous cognitive load is that to achieve an *Emotional Design*, certain features must be added to the learning material. According to the *coherence principle* by Mayer and Fiorella (2014), design elements that are described as interesting and entertaining but irrelevant in terms of content could cause extraneous cognitive load. However, this applies primarily to seductive details (*seductive detail effect*, Park et al., 2011), which does not include the *Emotional Design* according to Um et al. (2012). The *Emotional Design* does not consist of irrelevant, decorative elements. Instead, elements of the learning material are equipped with emotional features (Um et al., 2012). Thereby unnecessary extraneous cognitive load should be avoided. In addition, the *Emotional Design* is not expected to increase the extraneous cognitive load,

as it is assumed that it increases the cognitive resources that are made available for a task (as stated in section “Emotions, Emotional Design, and Learning”). Another aspect suggests that *Emotional Design* does not increase the extraneous load but even reduces it. This is because *Emotional Design* is supposed to guide attention (Le et al., 2021, Park et al., 2015; Peng et al., 2021). Search processes in a learning task are one source of extraneous cognitive load (Plass & Kalyuga, 2019), which can be reduced by guiding learners’ attention. Supporting this, students perceived the learning material adopting *Emotional Design* features to be less complicated (lower *perceived task difficulty*) than differently designed material (see meta-analysis by Brom et al., [2018]). The *perceived task difficulty* represents a persons’ estimation of effort needed to manage a certain task (Schneider et al., 2021). For *Emotional Design* features, a low perceived task difficulty is expected as they are supposed to ‘being perceived by the students as *easy to learn* because of their *entertaining appearance*’ (Uzun & Yildirim, 2018, p. 125).

In contrast, incorporating *Emotional Design* elements should not affect intrinsic cognitive load. Since the different design features are not presumed to affect aspects directly related to the complexity of the learning content. However, Le et al. (2021) reported that the *Emotional Design* decreased intrinsic load. Emotions, in general, could either increase or reduce the intrinsic cognitive load, which is context-dependent, according to Plass and Kalyuga (2019).

Regarding the germane cognitive load, it can be presumed that the *Emotional Design*, which potentially increases cognitive resources and lowers extraneous cognitive load, could increase the resources available for meaningful, effective learning (increase germane cognitive load).

Hypotheses

This study examined the influence of *Emotional Design* on learning performance, cognitive load, and affective state in a two-experiment approach. As part of **experiment 1**, the effect of *Emotional Design* on junior high school students’ learning performance, different dimensions of *perceived* cognitive load (e.g., in the form of *task difficulty*), and *perceived* affective state (e.g., *appeal*, *enjoyment*, and *desire for additional tasks of a similar nature*) were investigated. In **experiment 2**, the change of the *perceived affective state* was investigated in more detail. Concept maps with (1) bright and saturated colors as well as anthropomorphisms (*Emotional Design*, ED), (2) grayscale coloring and non-

anthropomorphic illustrations (neutral design, ND), and (3) without pictorial elements (control design, CD) were used.

Based on the theoretical foundation and previous research, the following hypotheses arise: We expect students who study with concept maps in *Emotional Design* to outperform students of the other conditions in terms of *learning performance* (H1). We further assume that students in the *Emotional Design* condition experience lower *extraneous cognitive load* (H2.1) and higher *germane cognitive load* (H2.2) compared to the other conditions. No difference is expected for the *intrinsic cognitive load* (H2.3). We presume students in the *Emotional Design* condition to experience a lower *task difficulty* (H2.4) and rate the *appeal, enjoyment, and desire for additional tasks of a similar nature* as higher (H3.1) than the other conditions. In terms of affect, we assume concept maps in *Emotional Design* to induce a rise in *positive affect* and a decrease in *negative affect*. (H3.2).

General Method

Two experiments were carried out in students' regular classrooms. For all experiments, the same learning material was used, and all students underwent concept map training to ensure comparable knowledge in the use of concept maps. Two investigators carried out the studies, one presenting the instruction and the other providing individual support if needed. Teachers were welcome to attend the lessons. All students were informed about the aims and course of the investigation, the option to stop participating at any time, guaranteed protection of data privacy, and the no-risk character of study participation. Written consents from parents and school principals were obtained before the study. During the experiments, we comply with the requirements of the school law, which, for example, determines that the learning material corresponds to the teaching specifications and does not promote a discriminatory understanding of the students (North Rhine-Westphalian Ministry of Education Science and Research, 2020) and followed the ethical principles and guidelines for the protection of human subjects of research (World Medical Association, 2001).

Learning Material

The learning material covered the topic ecosystem and consisted of three double pages (DIN A3) presented in a study book. The ecosystem is a central topic in biology

lessons (Lohmar & Eckhardt, 2014) and is well suited for the use of concept maps due to many possible interrelationships. Each double-page of the study book contained one of three concept maps referring to the aspects ‘Living Organisms in a Lake’, ‘Zones of a Lake’, and ‘Limnetic Zones in a Lake’. These concept maps followed the general principles of Novak and Cañas (2006), were hierarchically structured, and consisted of 16-25 concepts and 28-36 propositions. *CmapTools* software was used to construct the concept maps (available at <https://cmap.ihmc.us/>). Three study books were developed, each of them applied specific design principles (i.e., emotional, neutral, or control design) to the concept maps (see Figure 1).

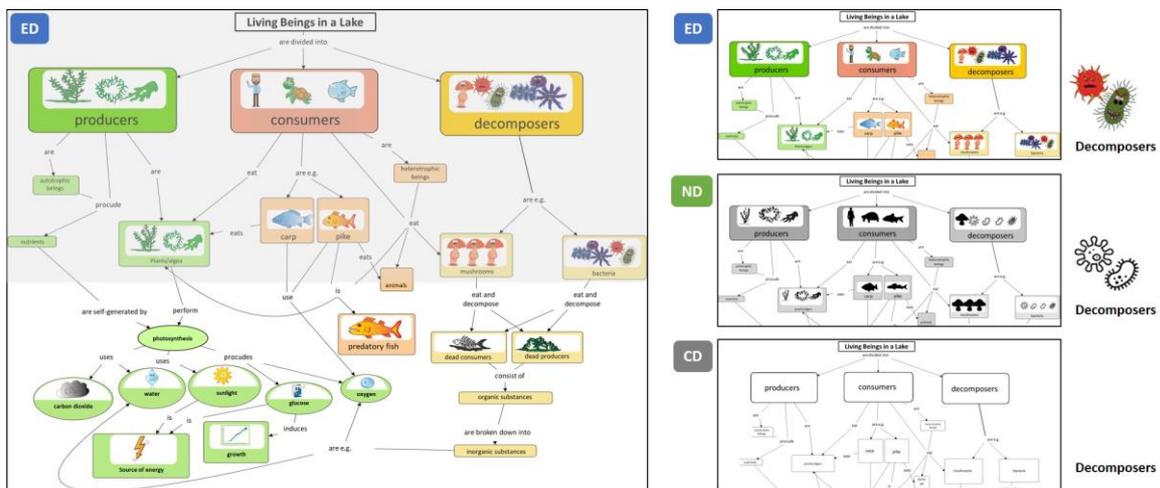


Figure 1. Concept map for ‘Living Organisms in a Lake’ (translated from German) in *Emotional Design* (ED) (left). Grey area of the concept map is displayed enlarged (middle) for the *Emotional Design*, the neutral design (ND), and the control design (CD). The concept ‘Decomposers’ is shown in detail (right).

The concept maps in the *Emotional Design* condition were constructed using bright and saturated colors and illustrations with anthropomorphisms based on the cognitive-affective theory of learning with media and along the lines of previous studies (e.g., Mayer & Estrella, 2014). Additionally, semantically related concepts were visually grouped by color (see *signaling principle*; Mautone & Mayer, 2001; for a similar approach see Nesbit & Adesope, 2011). For example, all concepts related to the main concept ‘Producers’ (e.g., ‘Photosynthesis’, ‘Plants/Algae’, and ‘Sunlight’) were colored green. Illustrations without emotionally relevant design features were incorporated for the neutral design condition, and monochromatic grayscale was used to group related concepts. Plain concept maps without any illustrations or design manipulations served as control design.

The conditions did not differ in content or structure (e.g., the number of objects shown, their orientation, location, or size). To avoid attention being divided (see *split-attention effect*; Mayer, 2001) for the neutral and *Emotional Design* concept maps, the illustrations were always presented in close spatial proximity to the textual element. Complete concept maps on ‘Living Organisms in a Lake’ are available as supporting information (see **Supplementary Learning Material**).

Concept Map Training

As training is recommended for the correct handling of concept maps (e.g., Allen & Tanner, 2003; Holley & Dansereau, 1984), we implemented a concept map training (135 min; Lenski et al., in preparation) before learning with concept maps. First, declarative knowledge about concept maps was conveyed (e.g., ‘What are concept maps?’, ‘What can they be used for?’, ‘What elements are concept maps made of?’). Learners were then provided with procedural knowledge. The training material was not specific to any school subject and covered various topics (e.g., ‘National Economy’ and ‘Different States of Water’).

Data Analysis

Normal distribution was identified visually and tested statistically using histograms and Kolmogorov-Smirnov tests. Levene’s tests or Box’s M tests were conducted to check for homogeneity of variances. Unless stated otherwise, all assumptions were met in the respective analyses. Parametric and nonparametric techniques were implemented, and if not otherwise specified an alpha significance level of .05 was set for all analyses and hypothesis testing.

Experiment 1

Method

Design and Procedure

An experimental intervention study was implemented with a pre- and post-test design. The study consisted of six lessons (à 45 min) delivered over three weeks (two lessons a week). In the first lesson, *prior knowledge*, *reading fluency*, and *demographic data* (age, gender, and grade in biology) were determined by questionnaires (45 min). In the following three lessons, the students received training in concept maps (see **Concept**

Map Training; General Method). In the learning phase (45 min), study books in the *Emotional Design*, neutral design, or control design were randomly distributed among the students (see **Learning Material; General Method**). The students were told they would be tested on the material's content afterward and were instructed to study each concept map carefully. After the learning phase, *cognitive load*, *task difficulty*, *appeal*, *enjoyment*, and *desire for additional tasks of a similar nature* were assessed by self-rate questionnaires included in the study books. Puzzles were handed out to those who completed their learning phase early to avoid disturbing others. In the following lesson, learning performance was measured using a post-test questionnaire (45 min).

Sample

An *a priori* power analysis was performed for sample size requirements. Based on the findings of former studies ($d_{recall} = 0.32$, $d_{transfer} = 0.33$, $d_{perceived\ difficulty} = -0.21$; Brom et al., 2018), we expected a medium effect (cf., Cohen, 2013) of our treatment on our primary dependent measure (*learning performance*). Power analysis was conducted using *G*Power 3.1.9.2* with the assumed medium effect ($f = .28$) and an alpha level of .05 at a power of .95. The result showed that a total sample of 223 students is required. In this context, Cunningham and McCrum-Gardner (2007) recommend estimating the drop-out rate in advance and adjusting the sample size. As to our knowledge, there are no benchmarks available on drop-out numbers from multi-day intervention studies in schools, we assumed a drop-out rate of about 15% based on previous studies of similar manner.

A total of 257 8th-grade students were recruited from four different German high schools and $N = 202$ students completed the entire experiment (47% female; age: $M = 13.18$, $SD = 0.49$; maximum classroom size: 33 students). All classes received 50 euros as compensation. Students who did not attend the training, the learning phase,

or the post-test were excluded in further analyses (for detailed information, see the flowchart of participants in **Figure 2**).

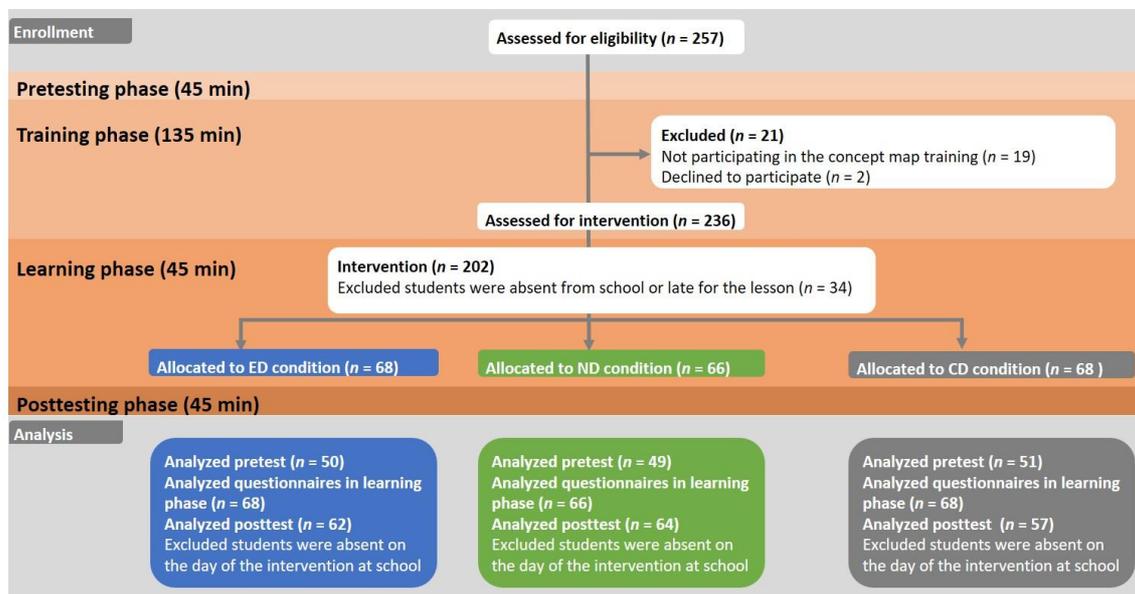


Figure 2. Flow of participants through **experiment 1** according to the criteria recommended in the CONSORT guidelines (Moher et al., 2012).

Instruments

Reading Fluency

As reading fluency was identified as a predictor for successful reading (National Institute of Child Health Human Development, 2000) and is known to influence learning performance (Bigozzi et al., 2017), we captured students' reading fluency by the *Salzburg Reading Screening SLS 2-9* (Wimmer & Mayringer, 2014), a classroom-administered standardized reading speed task (test-retest reliability for normative sample [grade 8] $r = .87$).

Prior Knowledge

Since prior knowledge can account for a large portion of a students' learning performance (Shapiro, 2004), we controlled for prior knowledge on 'Ecosystems' (including 'The Lake Ecosystem'). Prior knowledge was measured by a pre-test questionnaire consisting of 30 items that include general items about ecosystems and specific questions about the lake ecosystem. A general item deals, for example, with the question of why photosynthesis is important for humans (see **Supplementary Tests**, p. 2, item 3). An item focusing on the lake ecosystem covers, for example, the role of decomposers in the ecosystem lake (see **Supplementary Tests**, p. 6, item 14). Three of the items in the pretest

were taken from the third International Mathematics and Science Study TIMSS (Baumert et al., 1998; Harmon et al., 1997), as the items fit thematically well and have been validated for the 8th grade (see **Supplementary Tests**, items taken from the TIMSS study are marked accordingly). Nine items were excluded due to low items' discrimination $< .20$ (Durrheim & Tredoux, 2004). The remaining 21 items are single choice (13 items), multiple choice (4 items), and matching tasks (4 items). The complete questionnaire is available as supporting information (see **Supplementary Tests**). One point for each correct answer and zero points for incorrect answers were awarded; hence, the students' total scores obtained could range from a minimum score of zero to a maximum score of 21 points (Cronbach's $\alpha = .72$).

Perceived Cognitive Load

Regarding cognitive load, mostly the self-rating items on *perceived effort* (Paas, 1992) and *perceived (task) difficulty* were used in previous studies. According to Brom et al. (2018), these are proxies (subcomponents) of cognitive load and may not be suitable for this purpose. Therefore, contemporary cognitive load questionnaires are highly recommended for future *Emotional Design* research (Brom et al., 2018). Based on this, we decided to use the first version of the *naïve rating questionnaire* by Klepsch et al. (2017) to measure the three types of cognitive load separately (seven items). The items were not modified from the original except for the replacement of 'the task' with 'the concept map'. Items were rated on a 7-point Likert scale from 'not at all true' to 'very much true', with only the ends of the scale being labeled. Two items assessed the intrinsic cognitive load ('For this concept map, many things needed to be kept in mind simultaneously'; 'This concept map was very complex'; Cronbach's $\alpha = .51$). The germane cognitive load was assessed by two items as well ('For this concept map, I had to engage myself highly'; 'For this concept map, I had to think intensively about what things meant'; Cronbach's $\alpha = .67$). The extraneous cognitive load was investigated by three items ('When looking at concept maps, it was exhausting to find the critical information'; 'The design of the concept map was very inconvenient for learning'; 'When looking at concept maps, it was difficult to recognize and link the crucial information'; Cronbach's $\alpha = .77$).

Perceived task difficulty

To determine the *perceived task difficulty*, a single item was applied, which was used in previous studies (Mayer & Estrella, 2014; Um et al., 2012); 'Please rate how

difficult the concept maps were for you.’). This item was rated on a 5-point Likert scale from ‘*very easy*’ to ‘*very difficult*’.

Self-assessed Handling of Concept Maps

To determine whether all students had successfully passed the concept map training, they were asked to assess their handling of concept maps during the learning phase. The questionnaire was developed based on a self-regulation questionnaire in the context of a computer-based concept map training program for 10th-grade students (den Elzen-Rump & Leutner, 2007). It consists of four items on a 3-point Likert scale (e.g., ‘Did you look carefully at the arrow directions in the concept map?’ on a scale from ‘*seldom*’ to ‘*very often*’; Cronbach’s $\alpha = .70$). As the original questionnaire was designed for a construction task, the items were modified for an observation-based task (e.g., ‘arrows drawn’ was replaced with ‘looked carefully at the arrow directions’). Moreover, the emoticon-based rating scale was replaced by a number-based Likert scale. According to Pollock et al. (2018), emoticons should be used with caution in rating scales due to possible interpretation problems.

Perceived Appeal, Enjoyment, and Desire for additional Tasks of a similar Nature

To investigate affective parameters, we followed the approach of Mayer and Estrella (2014). They used a five-item questionnaire on *task difficulty*, *mental effort appeal*, *enjoyment*, and *desire for additional tasks of a similar nature* on a 5-point Likert scale. As in the present study a questionnaire on cognitive load is applied, only the four items covering appeal, enjoyment, and desire for additional tasks, and perceived task difficulty (see section “Perceived Task Difficulty”) were adopted. The question on mental effort was excluded to avoid redundancies. All items were translated into German and the term ‘the lesson’ was replaced by ‘the concept maps’ (*appeal*: ‘Please rate how appealing the concept maps were for you’ on a scale from ‘*very unappealing*’ to ‘*very appealing*’; *desire for additional tasks of a similar nature*: ‘I would like to learn from more concept maps like these’ on a scale from ‘*strongly disagree*’ to ‘*strongly agree*’; *enjoyment*: ‘I enjoyed learning from concept maps’ on a scale from ‘*strongly disagree*’ to ‘*strongly agree*’).

Knowledge of the Lake Ecosystem (Learning Performance)

Learning performance was measured by a 21-item post-test focusing on the ‘Lake Ecosystem’. Items were taken from the pre-test with two items slightly changed. Three items were excluded due to poor item parameters (items’ discrimination $< .20$; Durrheim & Tredoux, 2004). The final version of the questionnaire includes 15 single choice items and three matching items. The full questionnaire is available as supporting information (see **Supplementary Tests**). The maximum score on the pre-test was 18 points (Cronbach’s $\alpha = .78$).

Results

Preliminary Analyses

We checked whether the three conditions (*Emotional Design*, neutral design, and control design) differed in *age*, *reading fluency*, *prior knowledge* about ecosystems, *grades in biology*, and *gender* distribution. Since we assume that the null hypothesis is true for preliminary analyses, an alpha level of $\alpha = .10$ is set to increase test power and reduce the probability of committing a type II error (Döring & Bortz, 2016). One-way analysis of variance (ANOVA) did not result in any statistically significant differences between conditions in *age*, $F(2, 161) = 0.23$, $p = .798$, *reading fluency*, $F(2, 151) = 1.38$, $p = .255$ or *prior knowledge* about ecosystems, $F(2, 151) = 1.38$, $p = .255$. A chi-squared test revealed no statistically significant differences in *gender*, $\chi^2(4) = 6.23$, $p = .182$. Kruskal-Wallis tests showed that students did not differ referring to *self-assessed handling of concept maps*, $\chi^2(2) = 0.55$, $p = .761$, and *grades in biology*, $\chi^2(2) = 2.02$, $p = .364$. For descriptive data, see **Supplementary Table 1**. No unusual data distribution or outliers were observed.

Hypothesis Testing

Knowledge of the Lake Ecosystem (Learning Performance; H1)

We assumed that students in the *Emotional Design* condition outperform students in the neutral and control design conditions in terms of learning performance. A one-factorial analysis of covariance (ANCOVA) showed that, after adjustment for *prior knowledge* and *reading fluency* as covariates, the design conditions had no influence on *knowledge of the lake ecosystem* tested by means of a post-test, $F(2, 141) = 0.23$, $p = .796$ (lack of support for H1). **Table 1** displays mean values and standard deviations.

Table 1. Means and standard deviations of dependent variables

Dependent variables	ED		ND		CD	
	M	SD	M	SD	M	SD
Knowledge of the lake ecosystem (learning performance)	11.58	3.17	11.50	3.24	10.86	2.85
Appeal ^A	3.33	1.10	3.11	1.15	3.11	1.17
Enjoyment ^A	3.23	1.21	2.97	1.12	3.08	1.27
Desire for additional lessons of a similar nature ^A	3.12	1.09	3.05	1.17	2.97	1.29
Task difficulty ^A	2.61	1.05	2.71	0.90	3.08	1.05
Intrinsic cognitive load	4.92	1.28	4.36	1.31	4.96	1.42
Extraneous cognitive load	3.41	1.51	2.98	1.34	3.70	1.60
Germane cognitive load	4.32	1.42	3.86	1.41	4.48	1.56

ED = *Emotional Design*; ND = *Neutral design*; CD = *Control design*.

Note. ^AVariables are based on a single-item measures.

Cognitive Load and Task Difficulty (H2.1-2.4)

We expected the *Emotional Design* to affect the *perceived cognitive load*. The results of a bivariate correlation between all dependent variables showed moderate correlations between the cognitive load subscales (*intrinsic*, *extraneous*, and *germane load*) and *task difficulty* (bivariate correlation is available as supporting information [see **Supplementary Table 2**]). The results together with the classification of task difficulty as a proxy for cognitive load (Brom et al., 2018), a one-way multivariate analysis of variance (MANOVA) was conducted. *Design* was set as independent variable and the *cognitive load* subscales as well as the *perceived task difficulty* as dependent variables. MANOVA assumptions were met, except for the multivariate assumption of normality. Since MANOVAs are robust with respect to those deviations from normal distribution (Olson, 1974), we use the original MANOVA (for a similar argumentation see Stárková et al., 2019). Although *prior knowledge* correlates significantly with cognitive load, it was not found to be a significant covariate (Wilks' Lambda = 0.96, $F[2, 159] = 1.91$, $p = .13$). Thus, the variance analysis was performed without any covariate. A significant effect for design (Wilks' Lambda = 0.90, $F[2, 201] = 2.60$, $p = .009$, $\eta^2_p = .051$) was accompanied

by significant effects using univariate variance analyses (ANOVAs) on *extraneous cognitive load*, $F(2, 181) = 3.89$, $p = .022$, $\eta^2_p = .038$; *intrinsic cognitive load*, $F(2, 201) = 4.162$, $p = .017$, $\eta^2_p = .040$; *germane cognitive load*, $F(2, 201) = 3.16$, $p = .045$, $\eta^2_p = .031$, and *task difficulty*, $F(2, 201) = 4.32$, $p = .015$, $\eta^2_p = .042$.

Bonferroni post-hoc analyses were conducted. In the neutral design, intrinsic cognitive load was lower compared to the *Emotional Design*, $p = .050$ ($M_{\text{Diff}} = 0.56$, 95%-CI[1.12, 0.00]), and compared to the control design condition, $p = .031$ ($M_{\text{Diff}} = 0.60$, 95%-CI[1.16, 0.04]; lack of support for **H2.2**). See **Table 1** for mean values and standard deviations. The *extraneous cognitive load* was significantly lower in the neutral condition compared to the control design condition, $p = .018$ ($M_{\text{Diff}} = 0.72$, 95%-CI[1.34, 0.09]; lack of support for **H2.1**). The same pattern was observed for the *germane load*, $p = .049$ ($M_{\text{Diff}} = 0.61$, 95%-CI[0.00, 1.22]; lack of support for **H2.3**). The *perceived task difficulty* was highest for the control design condition and lowest for the *Emotional Design* condition, $p = .015$ ($M_{\text{Diff}} = 0.49$, 95%-CI[0.07, 0.91]; support for **H2.4**).

Appeal, Enjoyment, and Desire for Additional Tasks of a Similar Nature (H3.1)

Independent Samples Kruskal-Wallis tests showed that the students' perceived *appeal* (of the concept maps), $H(2, 199) = 1.09$, $p = .58$, *enjoyment*, $H(2, 201) = 1.99$, $p = .37$, and *desire for additional tasks of a similar nature*, $H(2, 200) = 0.71$, $p = .70$ did not differ significantly between conditions (lack of support for **H3.1**; see **Table 1**).

Experiment 2

Method

Design and Procedure

To investigate whether concept maps in *Emotional Design* influence *positive affect*, an experimental intervention study with a pre- and post-test design was conducted. First, demographic data (age, gender, and biology grade level) were determined, and students completed a questionnaire on positive and negative affect for the first time. Each student was then randomly assigned to one of the three study books implementing emotional, neutral, or control design (see **General Method**). In a learning phase, students were asked to study the concept maps carefully.

Sample

An *a priori* statistical power analysis using *G*Power 3.1.9.2* was performed for sample size requirements. For studies by Um et al. (2012) and Plass et al. (2014), large effect sizes of $d_{\text{PositiveAffect}} = 0.79$ and $d_{\text{PositiveAffect}} = 0.88$ were reported but in the meta-analysis by Brom et al. (2018) a marginal effect for positive affect ($d_{\text{PositiveAffect}} = 0.11$) was found. Since in the meta-analysis by Brom et al. (2018), studies were included in which only the color of a learning environment was varied, we consider the effect sizes gained by Um et al. (2012) and Plass et al. (2014) to be a better guide value for our study. A compromise was chosen, and a medium average effect size ($f = .30$) was anticipated. Accordingly, a total sample of 48 students was required for repeated measures ANOVA at a significance level of $\alpha = .05$ at a power of .95. A total of 54 students (grade 9) were recruited from a German high school. Since $n = 7$ were excluded because they missed the concept map training, $N = 47$ students were assigned to the treatment conditions (47% female; age: $M = 14.34$, $SD = 0.56$).

Instrument

The Positive and Negative Affect Scale (*PANAS*; Watson et al., 1988) was used in the German version (Breyer & Bluemke, 2016) to measure affect. For *positive affect*, ten different emotions related to positive affect (interested, excited, strong, enthusiastic, proud, alert, inspired, determined, attentive, and active) were presented to the students (Cronbach's $\alpha = .79$). For *negative affect*, ten different adjectives describing negative affect (distressed, upset, guilty, scared, hostile, irritable, ashamed, nervous, jittery, and afraid) were presented (Cronbach's $\alpha = .87$). Each subscale was rated on a 5-point Likert scale from 'not at all' to 'very much'. Ratings are averaged by mean as recommended by Breyer and Bluemke (2016).

Results

Data were analyzed using *IBM SPSS Statistics* (version 26) and *npard R* software package (The R Foundation for Statistical Computing, Vienna, Austria).

Preliminary Analysis

Prior to hypothesis testing, the three conditions (*Emotional Design*, neutral design, and control design) were checked for differences in *age*, *grades in biology*, and *gender distribution*. As described earlier, for preliminary analyses, an alpha level of $\alpha = .10$ is set

as recommended by Döring and Bortz (2016). A chi-squared test revealed no statistically significant results for *gender distribution*, $\chi^2(2) = 0.83, p = .659$. Furthermore, it was shown by means of Kruskal-Wallis tests that students in all conditions did not differ in terms of *grades in biology*, $H(2) = 2.31, p = .891$, or *age* $H(2) = 0.02, p = .990$. No unusual data distribution or outliers were observed.

Hypothesis Testing

Affective State (H3.2)

We expected students to experience more *positive affect* and less *negative affect* when learning with the *Emotional Design* compared to the neutral or the control design. As the data were not normally distributed, a nonparametric counterpart of the repeated measures analysis of variance was conducted (rANOVA; nparLD R package; see Noguchi et al., [2012] for more information) with *design* as an independent variable and *positive affect* as a repeated measures variable. Descriptive statistics are shown in **Table 2**. We found no significant main effect for *time of measurement* (pre vs. post), $F(1, 47) = 0.01, p = .921$. Also, no significant difference for the *design* could be determined, $F(2, 47) = 2.45, p = .087$. Furthermore, *time of measurement* and *design* did not interact with each other, $F(2, 47) = 1.11, p = .325$. To examine differences in the experienced *negative affect*, a rANOVA was calculated with *design* as an independent variable and *negative affect* as a repeated measures variable. No significant effect for the *time of measurement*, $F(1, 47) = 1.00, p = .315$ and non for the interaction (*time of measurement* : *design*) was observed, $F(1, 47) = 0.55, p = .577$. However, a significant main effect for the *design* could be found, $F(1, 47) = 4.82, p = .009$. Post-hoc analyses (Bonferoni) showed that learning with concept maps in the control design led to higher negative affect compared to concept maps in *Emotional Design*, $F(1, 47) = 4.36, p = .037$ and neutral design, $F(1, 47) = 7.59, p = .006$ (partly support for **H3.2**).

Table 2. Means and standard deviations of the positive and the negative affect

	ED		ND		CD	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Positive affect pre	2.99	0.64	2.71	0.79	3.13	0.80
Positive affect post	3.02	0.59	2.59	1.03	3.13	1.17
Negative affect pre	1.58	0.59	1.39	0.37	2.07	0.98
Negative affect post	1.51	0.60	1.69	0.82	2.10	1.12

ED = *Emotional Design*; ND = Neutral design; CD = Control design.

General Discussion

The purpose of this research was to examine the effect of *Emotional Design* in concept maps on junior high school students' *learning performance, perceived cognitive load, and perceived affective state*. The concept maps covered the biological topic of the lake ecosystem.

Learning Performance

Based on the cognitive-affective theory of learning with media and previous findings (e.g., Um et al., 2012), we assumed concept maps with *Emotional Design* to influence learners' affective-motivational state and thereby foster learning. Contrary to our assumption, *Emotional Design* did not foster learning performance. Our results contradict previous findings (Brom et al., 2018; Mayer & Estrella, 2014; Tien et al., 2018; Um et al., 2012; Wong & Adesope, 2021) but as current results in *Emotional Design* research are heterogeneous, our results complement other findings showing that *Emotional Design* did not influence recall, transfer, or comprehension (Park et al., 2015; Stárková et al., 2019; Uzun & Yıldırım, 2018). There are several possible reasons to explain conflicting results in the field. Varying intensity of *Emotional Design* could be one possible cause (Park et al., 2015) as features of *Emotional Design* could include variations in shape, color, and sound, as well as the use of anthropomorphic graphics. So far, studies that incorporate only one design feature (e.g., color; Heidig et al., 2015) and those applying various, potentially more intensive manipulations such as sounds, animations, and anthropomorphisms (Bülbül & Abdullah, 2021; Uzun & Yıldırım, 2018) are summarized under the umbrella term *Emotional Design*. Furthermore, the different ways in which the learning material is presented (e.g., paper-based or computer-based) must be considered. Previous studies were mainly computer-based (Münchow & Bannert, 2019; Münchow et al., 2017; Shangguan et al., 2020; Stárková et al., 2019; Um et al., 2012). Another aspect which needs to be considered is whether the learning material is animated or static (*nature of presentation*; Wong & Adesope, 2021). A meta-analysis by Berney and Bétrancourt (2016) revealed that animated graphics were more beneficial for learning compared to static graphics. Since our study was conducted in an authentic learning environment, the paper-based study books could not include sounds or animations. The resulting *Emotional Design* might, therefore, be 'weaker' compared to computer-based (animated) designs (see the *Emotional Designs* applied in Bülbül and Abdullah [2021] or Uzun and Yıldırım

[2018]). Another aspect could be the learning topic. In our study we used the topic of the lake ecosystem because it is well suited for the use of concept maps. However, the effect of *Emotional Design* manipulations could vary depending on the topic, which still needs to be investigated further.

Furthermore, although in the *Emotional Design* research a direct effect of emotions on learning is assumed (Um et al., 2012), there is evidence to question this direct connection (e.g., Linnenbrink, 2006). In fact, a large body of research supports the assumption that emotions influence learner's self-regulation, interest, intrinsic motivation, and their use of learning strategies and thereby 'exerting positive effects on overall performance' (Pekrun, 2017, p. 151). To further illustrate the complexity of the relationship between learning and emotions, it has to be mentioned that it is assumed that not only emotions can influence for example motivation and achievement but they in turn influence emotions (*reciprocal influence*; Pekrun, 2017).

As stated earlier, the cognitive theory of multimedia learning (Mayer, 2001), presumes that learning material providing both, pictures and text, should be more efficient for learning than material consisting only of pictures or only of texts. Therefore, concept maps with picture-text combinations (*Emotional Design* and neutral design) should have been more effective for learning than text-based concept maps (control design) which we could not observe in our study neither.

Cognitive Load

We expected *Emotional Design* manipulations to influence different types of perceived cognitive load. Other than expected, *Emotional Design* did not lead to significant differences in the perceived extraneous cognitive load compared to the non-*Emotional Designs*. This finding is in line with the results of previous studies (Park et al., 2015; Plass et al., 2014). Other than the *Emotional Design*, we found the neutral design leading to a significantly reduced extraneous cognitive load compared to the control design. *Emotional Design* features could impose more extraneous cognitive load than neutral design features since *Emotional Designs* come with additional seductive details, whereas neutral designs are – in the most literal sense – reduced to essentials. The findings support assumptions posed by the cognitive theory of multimedia learning and the emotions as extraneous cognitive load hypothesis but contradict the cognitive-affective theory of learning with media (Um et al., 2012).

Intrinsic cognitive load was presumed to be unaffected by design manipulations. Other than expected, the intrinsic load was reportedly lower in the neutral design condition compared to the *Emotional Design* and the control design condition. One potential source of intrinsic cognitive load can be lack of prior knowledge (Sweller, 2005). However, as the experimental conditions did not differ in terms of prior knowledge or the self-assessed handling of concept maps (see preliminary results, **experiment 1**), differently perceived complexity cannot be due to different degrees of familiarity with concept maps or different levels of prior knowledge. A possible explanation could be that learners were unable to distinguish between sources of extraneous cognitive load and sources of intrinsic cognitive load (Klepsch & Seufert, 2020). As students perceived lower levels of extraneous load, the intrinsic load of the learning material is felt less complex because more cognitive resources are available to cope with it (Klepsch & Seufert, 2020). For the settings where intrinsic cognitive load and extraneous cognitive load may merge, Klepsch et al. (2017) recommend using complex instruments to uncover underlying processes. It is also worth noting that the scale for intrinsic load had limited reliability in our experiment ($\alpha = .51$), consequently the results need to be interpreted with caution.

For germane cognitive load, we expected the *Emotional Design* to increase the perceived germane load compared to the other conditions. Contrary to these expectations, our results do not indicate such an effect. This is in line with the findings of several studies (Mayer & Estrella, 2014; Navratil et al., 2018; Park et al., 2015; Um et al., 2012; Uzun & Yıldırım, 2018). An explanation why *Emotional Design* does not lead to higher germane load invested in a task is offered by Uzun and Yıldırım (2018). They draw the conclusion, based on Pintrich et al. (2000) that *Emotional Design* elements could give the impression that the learning material is less difficult, resulting in less effort devoted to the task. Supporting this, the students' ratings of the perceived task difficulty in our study were significantly lower in the *Emotional Design* condition, compared to the control condition (Le et al., 2018).

Finally, methodological differences make it difficult to interpret the cognitive load results within the group of *Emotional Design* research. The self-rating questionnaire created by Paas (1992) has predominantly been used in *Emotional Design* research to measure mental effort. While, in some studies, mental effort is regarded as germane load (Um et al., 2012), in others it is equated with intrinsic load (DeLeeuw & Mayer, 2008). The

same applies to task difficulty. Um et al. (2012) use this to measure extraneous load, whereas DeLeeuw and Mayer (2008) assign it to germane load. Thus, direct comparison of the results in *Emotional Design* research is difficult.

Affective State

In contrast to the cognitive-affective theory of learning with media and respective previous findings (e.g., Um et al., 2012), we did not find *Emotional Design* to induce a *positive affective state* (see **experiment 2**). Furthermore, no difference was found for other affective variables like self-report ratings on how well the students enjoyed the learning phase, desired additional lessons of a similar nature, or found the learning material appealing (see **experiment 1**). These results are in line with findings by many other studies which also found no influence of the *Emotional Design* on positive affect (e.g., Heidig et al., 2015; Münchow & Bannert, 2019; Park et al., 2015).

However, both the neutral and the *Emotional Design* led to lower *negative affect* compared to the control design. Studying concept maps can be overwhelming for learners (*map shock*; Dansereau et al., 1994). The excessive demand could lead to an affective reaction characterized by frustration and results according to Blankenship and Dansereau (2000) in a loss of motivation. The findings of the present study suggest that adding (emotional or neutral) images to concept maps could help learners overcome negative feelings while studying concept maps.

According to the cognitive-affective theory of learning with media, learning is influenced by emotions. Thus, it is reasonable to assume that an *Emotional Design* that does not induce emotions may also be unable to influence learning performance. Stark et al. (2018) raised the question whether changes in the emotional state of learners need to be proven when classifying an applied design as emotional. However, they note that if a change in an emotional state is required for a design to be designated as *Emotional Design*, many studies that have been carried out so far could no longer be assigned to this research field. Heidig et al. (2015) and Münchow and Bannert (2019), for example, found that *Emotional Design* influenced learning performance even though their design did not induce positive affect. Stárková et al. (2019) suspected that the affective influence of *Emotional Design* is generally small and difficult to detect (supported by the meta-analysis by Brom et al., 2018). It must be noted that even replication studies using the same learning material and instruments yielded contradicting results for affect (Navratil

et al., 2018; Park et al., 2015; Plass et al., 2014). This suggests that the perceived affective state and learning performance could be influenced by other variables which still need to be identified.

Münchow and Bannert (2019) for example observed a time-dependent effect, whereby, after ten minutes of learning time, higher (albeit not statistically significant) positive emotions were found among students who learned with *Emotional Design* material compared to students who learned with a neutral design. After another ten minutes of learning, this was no longer observed. The authors concluded that *Emotional Design*, at least for the first ten minutes of learning, can prevent the reduction of positive emotions. In support of this observation, many of the studies that found a positive effect on emotional state, were based on a short learning time (≤ 15 min). Consequently, extended learning time (e.g., 20 minutes, as in our study) could cause these effects to fade. Wong and Adesope (2021) support these findings in their meta-analysis and suspect that ‘novelty of learning with *Emotional Designs* was most beneficial between short and moderate lengths of instructional time’ (p. 24).

Limitations

When measuring affective parameters using the self-reporting method, study participants may not be conscious of their current emotional state (Damasio, 2000) and, even if they are, they may encounter difficulties assigning it to specific categories (Salovey & Mayer, 1990). Since our study was carried out on junior high school students rather than college students (as in most studies), the students in this study might had problems answering the items on the PANAS scale. One possible reason for the failure to find any effects on for example, *appeal*, *enjoyment* or *perceived cognitive load* is that the concept maps provided in the form of study books may have been more elaborate than everyday worksheet-based learning material and, as the students only learned with one of the three designs, they lacked a direct comparison. Supporting this, Navratil et al. (2018) showed that an intended negative *Emotional Design* was only perceived as more negative when students had the opportunity to compare it directly with different designs.

Conclusion and Outlook

Our research pioneered to investigate the question of whether *Emotional Design* in paper-based concept maps improve learning for junior high school students. Together,

these results indicate no superiority of *Emotional Design* over non-*Emotional Design* concept maps. However, the use of *Emotional Design* seems to reduce the perceived task difficulty. The current state of research suggests that in the best-case scenario, emotions are triggered by an *Emotional Design* and learning is improved while, in the worst-case scenario, it has no effect on emotions and learning. As suggested by several authors, we agree that the effect of *Emotional Design* features could depend on factors like learning context, intensity of the design manipulation, the age of the learners or in which way the learning material is presented (e.g., Münchow & Bannert, 2019; Stárková et al., 2019; Wong & Adesope, 2021). The present study was conducted for a biological topic (as most studies in this field). Therefore, caution is needed when interpreting the results, and more research is required to transfer them to other cognitive domains or different learning conditions. To allow an assessment of the generalizability of our findings, we recommend broad and systematic examinations on *Emotional Design* in concept maps for other (biological) topics.

Furthermore, to make specific statements about the use of *Emotional Design* in concept maps, more basic research on *Emotional Design* manipulations is needed. From our point of view, this research should involve the following aspects.

Reliable measuring instruments for determining emotional influence should be established. We recommend more objective approaches (Brom et al., 2018), for example, heart rate variability or electrodermal activity measurements as implemented by, e.g., Lenski & Großschedl (submitted). Furthermore, the two dimensions of emotions (arousal and valence) should be examined differentially (Lenski & Großschedl, submitted) as it is suspected that the learning performance is especially affected by arousal (Irrazabal & Burin, 2021; Schneider et al., 2019). Additionally, further research is needed to identify the conditions in which *Emotional Design* manipulations successfully evoke emotions (see also Heidig et al., 2015). If the basis for objective measurement and emotional induction is established, effects of those *Emotional Design* features on learning performance and emotions can be analyzed and new insights gained. These insights could have a major impact on the design of learning material and learning success in general. Since learning success could be influenced by other factors, such as time spent with the learning material, personal characteristics like cultural background (Stárková et al., 2019) or different perceptions of aesthetics (Heidig et al., 2015), uncovering these aspects should be

focused on. As already stated, a large proportion of studies assume that emotions affect learning indirectly, which should also be reflected in *Emotional Design* research. Finally, the practicability of successful *Emotional Design* manipulations should be evaluated. If time-consuming and artistic manipulations are necessary to evoke emotions and influence educational variables, the use in everyday school life may be restricted (Lenski & Großschedl, submitted).

Data Availability Statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics Statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author Contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Funding

This study is part of the project “Learning Biology through Concept Mapping: Importance of a learning strategy training for cognitive load, cognitive processes, and learning performance,” which was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft; DFG), Grant (GR 4763/2).

Conflict of Interest

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.

Supplementary Material

The Supplementary Material for this article can be found online⁶⁶.

⁶⁶ <https://www.frontiersin.org/articles/10.3389/feduc.2022.807627/full#supplementary-material>

References

- Allen, D., & Tanner, K. (2003). Approaches to cell biology teaching: Mapping the journey – Concept maps as signposts of developing knowledge structures. *Cell Biology Education*, 2(3), 133-136. <https://doi.org/10.1187/cbe.03-07-0033>
- Alpert, S., & Grueneberg, K. (2001, June 25-30). Multimedia in concept maps: A design rationale and web-based application. *Proceedings of ED-MEDIA 2001, World Conference on Educational Multimedia, Hypermedia, and Telecommunication*, 31-36.
- Aspinwall, L. G. (1998). Rethinking the role of positive affect in self-regulation. *Motivation and Emotion*, 22(1), 1-32. <https://doi.org/10.1023/A:1023080224401>
- Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, 147(2), 134.
- Baumert, J., Lehmann, R., Lehrke, M., Clausen, M., Hosenfeld, I., Neubrand, J., ... & Günther, W. (1998). *Testaufgaben Naturwissenschaften TIMSS 7./8. Klasse (Population 2) [Test questions natural sciences TIMSS 7./8. Class (population 2)]*. Max-Planck-Institut für Bildungsforschung.
- Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education*, 101, 150-167. <https://doi.org/10.1016/j.compedu.2016.06.005>
- Bigozzi, L., Tarchi, C., Vagnoli, L., Valente, E., & Pinto, G. (2017). Reading fluency as a predictor of school outcomes across grades 4–9. *Frontiers in Psychology*, 8, 200. <https://doi.org/https://doi.org/10.3389/fpsyg.2017.00200>
- Blankenship, J., & Dansereau, D. F. (2000). The effect of animated node-link displays on information recall. *Journal of Experimental Education*, 68(4), 293-308. <https://doi.org/10.1080/00220970009600640>
- Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology*, 106(3), 849.
- Breyer, B., & Bluemke, M. (2016). Deutsche Version der Positive and Negative Affect Schedule PANAS [German version of the positive and negative affect schedule PANAS] (GESIS Panel) *Zusammenstellung sozialwissenschaftlicher Items und Skalen*. <https://doi.org/10.6102/zis242>

- Brom, C., Stárková, T., & D'Mello, S. K. (2018). How effective is *Emotional Design*? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review*, 25, 100-119. <https://doi.org/10.1016/j.edurev.2018.09.004>
- Bülbül, A. H., & Abdullah, K. (2021). *Emotional Design* of educational animations: Effects on emotion, learning, motivation and interest. *Participatory Educational Research*, 8(3), 344-355.
- Capota, K., van Hout, M., & van der Geest, T. (2007, August). Measuring the emotional impact of websites: A study on combining a dimensional and discrete emotion approach in measuring visual appeal of university websites. In *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces* (pp. 135-147).
- Chen, O., Kalyuga, S., & Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *Journal of Educational Psychology*, 107(3), 689.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic Press.
- Cunningham, J. B., & McCrum-Gardner, E. (2007). Power, effect and sample size using GPower: Practical issues for researchers and members of research ethics committees. *Evidence-Based Midwifery*, 5(4), 132-137.
- Damasio, A. R. (2000). A second chance for emotion. *Cognitive Neuroscience of Emotion*, 12-23.
- Dansereau, D. F., Dees, S. M., & Simpson, D. D. (1994). Cognitive modularity: Implications for counseling and the representation of personal issues. *Journal of Counseling Psychology*, 41(4), 513.
- Dees, S. M., Dansereau, D. F., & Simpson, D. D. (1994). A visual representation system for drug abuse counselors. *Journal of Substance Abuse Treatment*, 11(6), 517-523. [https://doi.org/10.1016/0740-5472\(94\)90003-5](https://doi.org/10.1016/0740-5472(94)90003-5)
- DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, 100(1), 223. <https://doi.org/10.1037/0022-0663.100.1.223>
- den Elzen-Rump, V., & Leutner, D. (2007). Naturwissenschaftliche Sachtexte verstehen – Ein computerbasiertes Trainingsprogramm für Schüler der 10. Jahrgangsstufe zum selbstregulierten Lernen mit einer Mapping-Strategie [Understanding

- scientific factual texts – A computer-based training program for 10th grade students for self-regulated learning with a mapping strategy]. In M. Landmann & B. Schmitz (Eds.), *Selbstregulation erfolgreich fördern* (pp. 251-268). Kohlhammer.
- Dever, D. A., Wiedbusch, M. D., Cloude, E. B., Lester, J., & Azevedo, R. (2021). Emotions and the comprehension of single versus multiple texts during game-
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation [Research methods and evaluation]*. Springer.
- Durrheim, K., & Tredoux, C. (2004). *Numbers, hypotheses & conclusions: A course in statistics for the social sciences*. Juta and Company Ltd.
- Gehl, D. (2013). *Vom Betrachten zum Verstehen [About viewing to understanding]*. Springer.
- Großschedl, J., & Tröbst, S. (2018). Biologie lernen durch Concept Mapping: Bedeutung eines Lernstrategietrainings für kognitive Belastung, kognitive Prozesse und Lernleistung – Kurzdarstellung des DFG-Projekts [Learning biology by concept mapping: the importance of learning strategy training for cognitive load, cognitive processes and learning performance - brief description of the DFG project]. *Zeitschrift für Didaktik der Biologie (ZDB) - Biologie Lehren und Lernen*, 22(1), 20-30. <https://doi.org/10.4119/zdb-1630>
- Harmon, M., Smith, T. A., Martin, M. O., Kelly, D. L., Beaton, A. E., Mullis, I. V., Gonzales, E. J., & Orpwood, G. (1997). *Performance assessment: IEA's third international mathematics and science study (TIMSS)*. International Association for the Evaluation of Educational Achievement.
- Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, 89(1), 92. <https://doi.org/10.1037/0022-0663.89.1.92>
- Heidig, S., Müller, J., & Reichelt, M. (2015). *Emotional Design* in multimedia learning: Differentiation on relevant design features and their effects on emotions and learning. *Computers in Human Behavior*, 44, 81-95. <https://doi.org/10.1016/j.chb.2014.11.009>

- Holley, C. D., & Dansereau, D. F. (1984). Networking: The technique and the empirical evidence. In D. F. Dansereau & C. D. Holley (Eds.), *Spatial learning strategies* (pp. 81-108). Elsevier.
- Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, 77(1), 95-111.
<https://doi.org/10.1002/sce.3730770107>
- Irrazabal, N., & Burin, D. (2021). Effects of emotional valence and arousal on comprehension and assembly of instructions. *Trends in Psychology*, 29(1), 104-122.
- Kinchin, I. M. (2000). Concept mapping in biology. *Journal of Biological Education*, 34(2), 61-68. <https://doi.org/10.1080/00219266.2000.9655687>
- Kinchin, I. M. (2001). If concept mapping is so helpful to learning biology, why aren't we all doing it? *International Journal of Science Education*, 23(12), 1257-1269.
- Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01997>
- Klepsch, M., & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science* 48(1), 45-77. <https://doi.org/10.1007/s11251-020-09502-9>
- Knörzer, L., Brünken, R., & Park, B. (2016). Facilitators or suppressors: Effects of experimentally induced emotions on multimedia learning. *Learning and Instruction*, 44, 97-107.
<https://doi.org/https://doi.org/10.1016/j.learninstruc.2016.04.002>
- Le, N., Nguyen, K., Nguyen, A., & Le, B. (2021). Global-local attention for emotion recognition. *ArXiv preprint ArXiv:2111.04129*.
- Le, Y., Liu, J., Deng, C., & Dai, D. Y. (2018). Heart rate variability reflects the effects of *Emotional Design* principle on mental effort in multimedia learning. *Computers in Human Behavior*, 89, 40-47. <https://doi.org/10.1016/j.chb.2018.07.037>
- Lenski, S., Elsner, S., & Großschedl, J. (in preparation). Comparing construction and study of concept maps - An intervention study on cognitive, metacognitive and emotional effects of training and learning.

- Lenski, S., & Großschedl, J. (submitted). *Emotional Design* pictures – Pleasant but too weak to evoke arousal and attract attention? *BMC Psychology*.
- Lenski, S. & Großschedl, J. (2021). Concept Maps im Unterricht: eine Trainingseinheit für Schüler*innen der Sekundarstufe [Concept maps in the classroom: a training session for secondary school students].
<https://doi.org/https://doi.org/10.17605/OSF.IO/48A5W>
- Lenzner, A., Schnotz, W., & Müller, A. (2013). The role of decorative pictures in learning. *Instructional Science*, *41*(5), 811-831.
- Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communication and Technology*, *30*(4), 195-232.
<https://doi.org/10.1007/BF02765184>
- Linnenbrink, E. A. (2006). Emotion research in education: Theoretical and methodological perspectives on the integration of affect, motivation, and cognition. *Educational Psychology Review*, *18*(4), 307-314.
- Lohmar, B., & Eckhardt, T. (2014). The education system in the Federal Republic of Germany 2012/2013: A description of the responsibilities, structures and developments in education policy for the exchange of information in Europe.
- Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology*, *93*(2), 377.
<https://doi.org/https://psycnet.apa.org/doi/10.1037/0022-0663.93.2.377>
- Mayer, R. E. (2001). A cognitive theory of multimedia learning In R. D. Mayer (Ed.), *Multimedia Learning* (pp. 41 - 62). Cambridge University Press.
- Mayer, R. E., & Estrella, G. (2014). Benefits of *Emotional Design* in multimedia instruction. *Learning and Instruction*, *33*, 12-18.
<https://doi.org/https://doi.org/10.1016/j.learninstruc.2014.02.004>
- Mayer, R. E., & Fiorella, L. (2014). 12 principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. *The Cambridge Handbook of Multimedia Learning*, 279.
- Mensink, M. C. (2021). Emotional responses to seductive scientific texts during online and offline reading tasks. *Discourse Processes*, 1-18.

- Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (1997). Meaningful learning in science: The human constructivist perspective. *Handbook of Academic Learning* 405-447. <https://doi.org/10.1016/B978-012554255-5/50014-4>
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., ... & Altman, D. G. (2012). CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *International journal of surgery, 10*(1), 28-55.
- Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning, 22*(3), 149-158. <https://doi.org/10.1111/j.1365-2729.2006.00170>
- Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology, 97*(1), 117.
- Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(3), 309-326. <https://doi.org/10.1007/s10648-007-9047-2>
- Münchow, H., & Bannert, M. (2019). Feeling good, learning better? Effectivity of an *Emotional Design* procedure in multimedia learning. *Educational Psychology, 39*(4), 530-549. <https://doi.org/10.1080/01443410.2018.1524852>
- Münchow, H., Mengelkamp, C., & Bannert, M. (2017). The better you feel the better you learn: Do warm colours and rounded shapes enhance learning outcome in multimedia learning? *Education Research International, 2017*. <https://doi.org/10.1155/2017/2148139>
- National Institute of Child Health Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769)*. U.S. Government Printing Office.
- Navratil, S. D., Köhl, T., & Heidig, S. (2018). Why the cells look like that – The influence of learning with *Emotional Design* and elaborative interrogations. *Frontiers in Psychology, 9*, 1653. <https://doi.org/10.3389/fpsyg.2018.01653>
- Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413-448. <https://doi.org/10.3102%2F00346543076003413>

- Nesbit, J. C., & Adesope, O. O. (2011). Learning from animated concept maps with concurrent audio narration. *The Journal of Experimental Education, 79*(2), 209-230. <https://doi.org/10.1080/00220970903292918>
- Noguchi, K., Gel, Y. R., Brunner, E., & Konietschke, F. (2012). nparLD: an R software package for the nonparametric analysis of longitudinal data in factorial experiments. *Journal of Statistical Software, 50*(12).
- North Rhine-Westphalian Ministry of Education Science and Research. Schulgesetz für das Land Nordrhein-Westfalen [School law for the state of North Rhine-Westphalia], (2020). Retrieved November 19, 2021, from: <https://bass.schulwelt.de/6043.htm>
- Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27*(10), 937-949.
- Novak, J. D., and Cañas, A. J. (2008). *The Theory Underlying Concept Maps And How To Construct Them*. Technical Report IHMC CmapTools 2006-01 Rev 01-2008. Florida Institute for Human and Machine Cognition, Pensacola, Florida, USA. Available online at: <http://www.ssu.ac.ir/fileadmin/templates/fa/Moavenatha/Moavenate-Amozeshi/edicupload/olymp-3.pdf> (accessed March 11, 2021).
- Novak, J. D., and Cañas, A. J. (2006). *The Theory Underlying Concept Maps And How To Construct Them*. Technical Report IHMC CmapTools 2006-01. Available online at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.8995&rep=rep1&type=pdf> (accessed March 11, 2021).
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press.
- Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. *Journal of the American Statistical Association, 69*(348), 894-908.
- Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429. <https://doi.org/10.1037/0022-0663.84.4.429>
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.

- Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). *Emotional Design* and positive emotions in multimedia learning: An eyetracking study on the use of anthropomorphisms. *Computers & Education*, 86, 30-42.
<https://doi.org/10.1016/j.compedu.2015.02.016>
- Park, B., Moreno, R., Seufert, T., & Brünken, R. (2011). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior*, 27(1), 5-10. <https://doi.org/10.1016/j.chb.2010.05.006>
- Pekrun, R. (2021). Emotions in reading and learning from texts: Progress and open problems. *Discourse Processes*, 58, 1-10.
- Pekrun, R. (2017). Achievement emotions. *Emotions in Late Modernity*, 142.
- Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology*, 98(3), 583. <https://doi.org/10.1037/0022-0663.98.3.583>
- Peng, X., Xu, Q., Chen, Y., Zhou, C., Ge, Y., & Li, N. (2021). An eye tracking study: positive emotional interface design facilitates learning outcomes in multimedia learning?. *International Journal of Educational Technology in Higher Education*, 18(1), 1-18.
- Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 43-97). Buros Institute of Mental Measurements.
- Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). *Emotional Design* in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction*, 29, 128-140.
<https://doi.org/10.1016/j.learninstruc.2013.02.006>
- Plass, J. L., & Kalyuga, S. (2019). Four ways of considering emotion in cognitive load theory. *Educational Psychology Review*, 1-21. <https://doi.org/10.1007/s10648-019-09473-5>
- Plass, J. L., & Kaplan, U. (2016). *Emotional Design* in digital media for learning. In *Emotions, technology, design, and learning* (pp. 131-161). Elsevier.
- Pollock, G., Ozan, J., Goswami, H., Rees, G., & Stasulane, A. (2018). *Measuring youth well-being*. Springer.
- Renkl, A. (2010). *Lehren und Lernen [Teaching and learning]*. Springer.

- Reiners, C. S., Großschedl, J., Meyer, M., Schadschneider, A., Schäbitz, F., & Struve, H. (2018). Using the terms experiment, theory, model, and law in mathematics and natural sciences. *ChemKon*, 25(8), 324-333.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145.
- Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, 9(3), 185-211. <https://doi.org/10.2190%2FDUGG-P24E-52WK-6CDG>
- Schmid, R. F., & Telaro, G. (1990). Concept mapping as an instructional strategy for high school biology. *The Journal of Educational Research*, 84(2), 78-85.
- Schneider, S., Nebel, S., Meyer, S., & Rey, G. D. (2021). The interdependency of perceived task difficulty and the choice effect when learning with multimedia materials. *Journal of Educational Psychology*. Advance online publication. <https://doi.org/10.1037/edu0000686>
- Schneider, S., Häbler, A., Habermeyer, T., Beege, M., & Rey, G. D. (2019). The more human, the higher the performance? Examining the effects of anthropomorphism on learning with media. *Journal of Educational Psychology*, 111(1), 57.
- Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review*, 30(2), 431-455. <https://doi.org/10.1007/s10648-017-9403-9>
- Shangguan, C., Gong, S., Guo, Y., Wang, X., & Lu, J. (2020). The effects of *Emotional Design* on middle school students' multimedia learning: The role of learners' prior knowledge. *Educational Psychology*, 1-18. <https://doi.org/10.1080/01443410.2020.1714548>
- Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal*, 41(1), 159-189.
- Stark, L., Park, B., & Brünken, R. (2018). Emotionen beim Lernen mit Multimedia [Emotions while learning with multimedia]. In S. Ladel, J. Knopf, & A. Weinberger (Eds.), *Digitalisierung und Bildung* (pp. 141-158). Springer.
- Stárková, T., Lukavský, J., Javora, O., & Brom, C. (2019). Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning? *Journal of Computer Assisted Learning*, 35(4), 555-568. <https://doi.org/10.1111/jcal.12359>

- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer (Ed.), *Cambridge Handbook of Multimedia Learning*. Cambridge University Press.
- Tien, L.-C., Chiou, C.-C., & Lee, Y.-S. (2018). *Emotional Design* in multimedia learning: Effects of multidimensional concept maps and animation on affect and learning. *EURASIA Journal of Mathematics, Science and Technology Education*, 14(12), em1612. <https://doi.org/10.29333/ejmste/94229>
- Um, E., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). *Emotional Design* in multimedia learning. *Journal of Educational Psychology*, 104(2), 485. <https://doi.org/10.1037/a0026609>
- Uzun, A. M., & Yıldırım, Z. (2018). Exploring the effect of using different levels of *Emotional Design* features in multimedia science learning. *Computers & Education*, 119, 112-128. <https://doi.org/10.1016/j.compedu.2018.01.002>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063.
- Williams, C. G. (1998). Using concept maps to assess conceptual knowledge of function. *Journal for Research in Mathematics Education*, 29(4), 414-421. <https://doi.org/10.5951/jresmetheduc.29.4.0414>
- Wimmer, H., & Mayringer, H. (2014). *SLS 2-9: Salzburger Lese-Screening für die Schulstufen 2-9 [SLS 2-9: Salzburg reading screening for grades 2-9]*. Hogrefe.
- Wong, R. M., & Adesope, O. O. (2021). Meta-analysis of *Emotional Designs* in multimedia learning: A replication and extension study. *Educational Psychology Review*. <https://doi.org/10.1007/s10648-020-09545-x>
- World Medical Association. (2001). World medical association declaration of helsinki. Ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, 79(4), 373.

9.2 Emotional Design pictures – Pleasant but too weak to evoke arousal and attract attention? (Publikation IV)

9.2.1 Zusammenfassung, Fragestellungen und methodischer Zugang

In der neueren Forschung zum Lernen mit Multimedia wird angenommen, dass bestimmte Design-Prinzipien, wie z. B. anthropomorphe Merkmale mit ansprechenden Farben (*Emotional Design*), lernförderlich sein können, da sie den affektiven Zustand (d. h. Valenz, Aktivierung) und die Aufmerksamkeitslenkung der Lernenden positiv beeinflussen können. Bislang wurde die emotionale Wirkung hauptsächlich durch die Verwendung von Fragebögen ermittelt. Da sich in Metaanalysen ein heterogenes Befundmuster in Bezug auf die affektive Wirkung zeigte, wurden zur Schaffung eines aussagekräftigeren Befundmusters objektivere Messmethoden empfohlen. Um perspektivisch die Manipulation von Emotionen durch entsprechende Gestaltung gezielt einsetzen zu können, ist zudem die differenzierte Untersuchung von Valenz und Aktivierung notwendig (Irrazabal & Burin, 2021). **Publikation IV** umfasst zwei Experimente, in denen die Wirkung von unterschiedlich gestalteten Abbildungen (emotionales Design [ED], neutrales Design [ND] und Kontrolldesign [CD]⁶⁷) auf die Valenz, Aktivierung und die Aufmerksamkeit von Schüler:innen untersucht wurde. Die Untersuchung bezüglich der Aktivierung ($N = 18$) und der Aufmerksamkeit ($N = 27$) erfolgte durch psychophysiologische Messungen. Die Valenz ($N = 15$) wurde im Rahmen einer separaten Online-Umfrage unter der Verwendung derselben Abbildungen ermittelt.

Die konkreten Fragestellungen, welche im Rahmen dieser Studie beantwortet werden sollen, lauten:

F₁: Unterscheidet sich die Wirkung von emotional gestalteten Abbildungen und nicht-emotional gestalteten Abbildungen in Bezug auf die Valenz und die Aktivierung?

F₂: Generieren emotional gestaltete Abbildungen mehr Aufmerksamkeit, als nicht-emotional gestaltete Abbildungen?

⁶⁷ Detaillierte Informationen zu den unterschiedlichen Designs und eine Beispiel-Abbildung finden sich in der Original-Publikation (**Kapitel 9.2.3**).

9.2.2 Eigenanteile

Die Autorin der vorliegenden Dissertation konzipierte die Studie und führte sie durch. Bei der Erstellung des Erhebungsmaterials, der Durchführung der Studie und der Dateneingabe wurde sie von Abschlussarbeitskandidat:innen⁶⁸ unterstützt. Die Autorin führte die Datenanalyse durch, verfasste das Manuskript und reichte es ein. Das Manuskript wurde von Herrn Prof. Dr. Großschedl überarbeitet, welcher auch die Supervision übernahm.

⁶⁸ Benjamin Bayer, Florian Schmitz, Regina Drevs und Tanja Treinen.

9.2.3 Publikation IV: Originalmanuskript

Emotional Design pictures – Pleasant but too weak to evoke Arousal and Attract Attention?

Sina Lenski & Jörg Großschedl

Abstract

Background: A new field of research builds on the theoretical assumption that distinct design manipulations, such as human-like features with appealing colors (*Emotional Design*), foster multimedia learning by influencing the learners' affective state (i.e., valence, arousal) and attention. Empirical studies, however, provide inconsistent findings of the affective potential of *Emotional Design*, underlining the need for psychophysiological measurements to assess the affective state and attention more objectively.

Methods: Two experiments were conducted to discover the effect of differently designed stimuli (picture-text combinations with anthropomorphic pictures in bright and saturated colors [*Emotional Design*], picture-text combinations with non-anthropomorphic pictures in grayscale [neutral design], and plain text [control design]) on junior high school students' valence ($N = 15$), arousal ($N = 18$), and attention ($N = 27$). Valence was determined by students' judgments on a rating scale; arousal and attention were assessed by psychophysiological parameters (electrodermal activity of students and their dwell time on stimuli, respectively).

Results: Results show that students perceived the valence of stimuli that represented positive (e.g., sunlight) or neutral content (e.g., consumer) more pleasantly in *Emotional Design* than in neutral and control designs. No differences were observed for stimuli representing negative content (e.g., dead consumer). Moreover, results show that *Emotional Design* did not induce higher arousal than neutral and control designs and did not attract

more attention. Instead, the text within picture-text combinations (emotional and neutral designs) attracted more attention when pictures were presented in neutral than in *Emotional Design*.

Conclusion: By measuring valence and arousal using psychophysiological parameters, the present study helps to understand the heterogeneous findings of previous studies.

Keywords: Arousal, attention electrodermal activity, *Emotional Design*, eye-tracking, valence

Introduction

In the context of the Covid-19 pandemic, a large amount of educational material (videos, brochures, posters, flyers, and books) has been developed to educate people about the virus and to communicate the code of conduct (e.g., Molina, 2020; Save the Children, 2020; United Nations Children's Fund [UNICEF], 2020). This material often implements design principles that humanize the virus by depicting it with, for example, eyes or mouths (e.g., Centers for Disease Control and Prevention [CDC], 2020). Design principles such as the implementation of human-like features or round shapes and appealing colors lead to a design known as *Emotional Design*. *Emotional Design* is believed to positively impact learners' affective state, affecting students' interest, cognitive load, and learning performance (Um, Plass, Hayward, & Homer, 2012). Meta-analyses, however, report only weak effects of *Emotional Design* on students' affective states (Brom, Stárková, & D'Mello, 2018; Wong & Adesope, 2021). In contrast to the results of the two meta-analyses, the effects reported in individual studies appear to be quite inconsistent. While some studies found the *Emotional Design* to change the affective state (e.g., Bülbül & Abdullah, 2021; Shangguan, Gong, Guo, Wang, & Lu, 2020; Um et al., 2012; Uzun & Yıldırım, 2018), other studies found no effects (e.g., Königschulte, 2015; Münchow &

Bannert, 2019; Münchow, Mengelkamp, & Bannert, 2017; Peng et al., 2021; Stárková et al., 2019). Various factors can cause inconsistent findings. (a) Valence and arousal (activation) were rarely observed as separate dimensions of the affective state, making the multiple results difficult to compare. (b) Self-rate questionnaires were the predominant way to gather insight into the affective state, which is error-prone as participants may not be sensitive to their current affective state or changes (Shangguan et al., 2020). Consequently, several authors call for more objective measures in studying the affective state (e.g., Brom et al., 2018; Bülbül & Abdullah, 2021; Irrazabal & Burin, 2021; Peng et al., 2021). Moreover, previous research (e.g., Königshulte, 2015; Moreno, 2006; Moreno & Mayer, 2007; Park et al., 2015) allows the conclusion that the affective state influences attention, and attention, in turn, influences learning. Our study aimed to provide a deeper insight into whether *Emotional Design* affects valence, arousal, and attention in junior high school students. The present study results are intended to suggest the practical design of multimedia learning environments in schools.

Affective State and Emotional Design

According to Russell's (2003) *circumplex model*, a spatial metaphor of two orthogonal dimensions can describe individuals' affective states. The individual's current position in this two-dimensional space determines the emotion (e.g., anger, happiness, sadness, relaxation) the individual currently perceives. The first dimension (*valence*) characterizes the value of a stimulus and spans a continuum from the endpoints *pleasant* and *unpleasant*. The second dimension (*arousal*) describes the degree of activation of the central nervous system and spans a continuum between *activated* and *deactivated*. Activating emotions are enjoyment or curiosity (pleasant) and anger or fear (unpleasant). Deactivating emotions include relaxation (pleasant) and sadness (unpleasant, Pekrun, Elliot,

& Maier, 2006). Arousal of the nervous system can be measured by the degree of activation experienced by an individual (e.g., by measuring electrodermal activity [EDA] or heart rate variability).

Only a few psychophysiological measurements have been used in *Emotional Design* research. For example, Uzun and Yıldırım (2018) observed participants' heart rate variability and concluded that *Emotional Design* favors activating emotions, increasing arousal. Le, Liu, Deng, and Dai (2018) used a similar methodological approach but regarded heart rate variability as an indicator of *mental effort*. They conclude that an *Emotional Design* perpetuates mental effort compared to a neutral design. Another methodological approach consists of EDA measurements. EDA measurements have been used in neuromarketing and affective neuroscience for many years but have only recently gained attention in educational research (Braithwaite, Watson, Jones, & Rowe, 2013; Caruelle, Gustafsson, Shams, & Lervik-Olsen, 2019; Pijeira-Díaz, Drachsler, Järvelä, & Kirschner, 2016; Posada-Quintero & Chon, 2020). Increased EDA has been measured for stimuli that generate positive emotions, such as preferred products (Ohira & Hirao, 2015), happy music (Krumhansl, 1997), positive facial expressions (Levenson, Ekman, & Friesen, 1990), but for unpleasant stimuli as well (Lang & Bradley, 2007). Findings from these studies suggest that EDA is an indicator of arousal (i.e., the degree of activation) rather than an indicator of valence (Correia, Morais, Quintão, Quaresma, & Vigário, 2021).

Attention and Emotional Design

Visual *attention* is defined as “the mechanism the nervous system uses to highlight specific locations, objects or features within the visual field” (Bisley, 2011, p. 49). According to the *eye-mind hypothesis* (Just & Carpenter, 1980), it is assumed that the location of eye fixation reflects the place of attention, while fixation duration (so-called *dwell time*) is regarded as an indicator of the amount of attention (Holmqvist et al., 2011).

A long-lasting fixation indicates deep information processing or difficulty in processing information (Rayner, 1998), whereas a short-lasting fixation does not. Thus, eye-tracking should give more direct and objective access to cognitive processes than self-reports (Alemdag & Cagiltay, 2018). Eye-tracking has been frequently used in multimedia research and provided insight into the processing of information (e.g., Bayram & Bayraktar, 2012; Mwambe, Tan, & Kamioka, 2021; van Gog & Scheiter, 2010; von Reumont & Budke, 2020). A study by Rayner (2009), for example, has shown that the average fixation duration on pictures is longer than fixation duration on text elements. However, Alemdag and Cagiltay (2018) meta-analysis revealed an opposite effect, indicating that learners pay more attention to the text than pictures.

For emotional stimuli, it is assumed that the cognitive system is “likely to be complemented by a perceptual mechanism that is biased to readily detect and process emotional stimuli among other, competing stimuli” (Nummenmaa, Hyönä, & Calvo, 2006, p. 257). Eye-tracking research has shown that stimuli in an *Emotional Design* attract more attention compared with neutral stimuli (Nummenmaa et al., 2006; Peng et al., 2021). This effect is known as *emotion-related attentional bias* and has been reported in numerous studies (e.g., Harrison, Sullivan, Tchanturia, & Treasure, 2010; Lea, Qualter, Davis, Pérez-González, & Bangee, 2018). Attention is presumably not caused by valence but by the arousal intensity (e.g., Vogt, de Houwer, Koster, van Damme, & Crombez, 2008). However, findings in *Emotional Design* research appear to be inconsistent: While Park et al. (2015) found *Emotional Design* features to attract attention, Stárková et al. (2019) could not find such an attention-grabbing effect for stimuli adopting *Emotional Design* principles compared with non-*Emotional Designs*.

Problem Statement

Recent research investigates the effect of *Emotional Design* on valence and arousal as the defining dimension of affective state (Chung & Cheon, 2020; Irrazabal & Burin, 2021), but a systematic comparison of emotional and non-*Emotional Designs* regarding valence, arousal, and attention is still missing. The present study compares the effect of three differently designed stimuli (picture-text combinations with anthropomorphic pictures in bright and saturated colors [*Emotional Design*], picture-text combinations with non-anthropomorphic pictures in grayscale [neutral design], and plain text [control design]) on valence (**experiment I**), arousal, and attention (**experiment II**) in samples of junior high school students. Whereas valence was determined by students' judgments on a rating scale, arousal and attention were assessed by psychophysiological parameters (electrodermal activity of students and their dwell time on stimuli, respectively).

- Hypothesis 1 (**H1**): Stimuli in *Emotional Design* will be perceived as more pleasant (measured by self-rating) than those in neutral or control designs.
- Hypothesis 2 (**H2**): Stimuli in *Emotional Design* lead to a higher state of arousal (measured by EDA) than those in neutral or control designs.
- Hypothesis 3 (**H3**): Stimuli in *Emotional Design* will attract more attention (measured by dwell time) than those in a neutral design.
- Exploratory Goal (**EG**): Textual elements in picture-text combinations attract more attention than pictorial elements. An explorative approach aimed to investigate whether the pictorial element's *Emotional Design* (vs. neutral design) affects this preference.

Experiment I

Method

Sample

We used G*Power software (Version 3.1.9.2) to determine an adequate sample size (Faul, Erdfelder, Lang, & Buchner, 2007). We considered a medium effect⁶⁹ ($f = .30$) to be desirable for our treatment, defined an α error level of .05, and a power of $1 - \beta = .95$ (according to Whitley & Ball, 2002). The *a priori* power analysis showed that a total sample of 13 students is adequate for repeated-measures ANOVA. A total of 23 German junior high school students from grades 8 and 9 were recruited. Eight students had to be excluded as they did not complete the questionnaire. The remaining sample consists of $N = 15$ students (53 % female; $M_{age} = 13.67$ years, $SD = 0.62$).

Stimulus Material, Design, and Procedure

We selected twelve concepts (e.g., *water*, *fish*, and *consumer*) of a biological topic (Lake Ecosystem) and applied three design principles to represent them. The application of the design principles resulted in 3 x 12 stimuli (nota bene: we refer to a concept that is presented in a particular design as a stimulus), of which 12 represented an *Emotional Design*, 12 a neutral design, and 12 the control design (stimulus material is available as supplementary material [S1]). Emotional stimuli are picture-text combinations with anthropomorphic pictures in bright and saturated colors, neutral stimuli are picture-text combinations with non-anthropomorphic pictures in grayscale, and control stimuli are plain text without a pictorial representation. An online survey was carried out on the *SoSci Survey* platform and shared with the participants on www.soscisurvey.com. The survey started with instructional information, followed by demographic data (age, gender, and

⁶⁹ The effect sizes reported in the literature differ greatly. Some studies report small effects (e.g., Brom et al., 2018, Heidig et al. 2015) and others large effects (e.g., Plass et al., 2014; Um et al., 2012; Uzun & Yıldırım, 2018).

grade level). Then the stimuli were presented in a fixed randomized order with one stimulus per page. Below each stimulus, a slider scale was displayed to gain information about perceived valence. The program did not allow to continue without answering, and the survey took an average time of $M = 3.39$ min to complete ($SD = 0.59$ min, range 2.42 – 4.42 min).

Instrument

To examine how far different stimuli and design principles (emotional, neutral, and control design) induce valence, a ten-point slider scale from *unpleasant* (rating = 1) to *pleasant* (rating = 10) was placed below each stimulus. A similar rating scale was used by Yao, Wu, Zhang, and Wang (2017).

Results

According to our first hypothesis (**H1**), stimuli in *Emotional Design* are more pleasant than stimuli in neutral design or control designs. A two-way repeated-measures ANOVA (rANOVA) with Greenhouse–Geisser adjustment was conducted to test **H1** and to explore the effect of different concepts (e.g., sunlight, consumer, dead consumer; see **Figure 2**) on valence ratings (see **Table 1** for descriptive data). The rANOVA revealed a significant main effect of *design* on participants' valence ratings, $F(1.17, 16.32) = 69.03$, $p < .001$, $\eta_p^2 = .831$ (strong effect with $f = 2.22$). Bonferroni-corrected pairwise comparisons showed that emotional stimuli ($M = 6.78$, $SD = 0.87$) appear more pleasant (i.e., higher valence ratings) than neutral stimuli ($M = 4.48$, $SD = 0.47$) and control stimuli ($M = 4.73$, $SD = 0.66$), thereby supporting **H1**.

Table 1. Mean valence ratings and standard deviations for design ($n = 15$).

Variable	ED		ND		CD	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Valence	6.78	0.87	4.48	0.47	4.73	0.66

ED = *Emotional Design*; ND = *Neutral design*; CD = *Control design*.

The effect of *concepts* is large ($f = 1.56$) and significant as well, $F(3.08, 43.09) = 33.95$, $p < .001$, $\eta_p^2 = .708$, which indicates that the concept of a stimulus represents also influences students' valence rating. Valence ratings for some concepts (e.g., dead consumer) are significantly lower than for other concepts (e.g., sunlight; see **Figure 1**). In addition, the rANOVA revealed a significant interaction effect between design and concept, $F(5.30, 74.15) = 12.39$, $p < .001$, $\eta_p^2 = .470$ (strong effect with $f = 0.942$). A closer look at the data showed that the design affects valence ratings for some concepts (e.g., sunlight) but not for others (e.g., dead consumer; see **Figure 1**, grey highlights). It is worth mentioning the high scatter of the data for the concepts of bacteria and predatory fish in the *Emotional Design*. Descriptive data and post-hoc analyses for the concepts are provided in the supplementary material (S2).

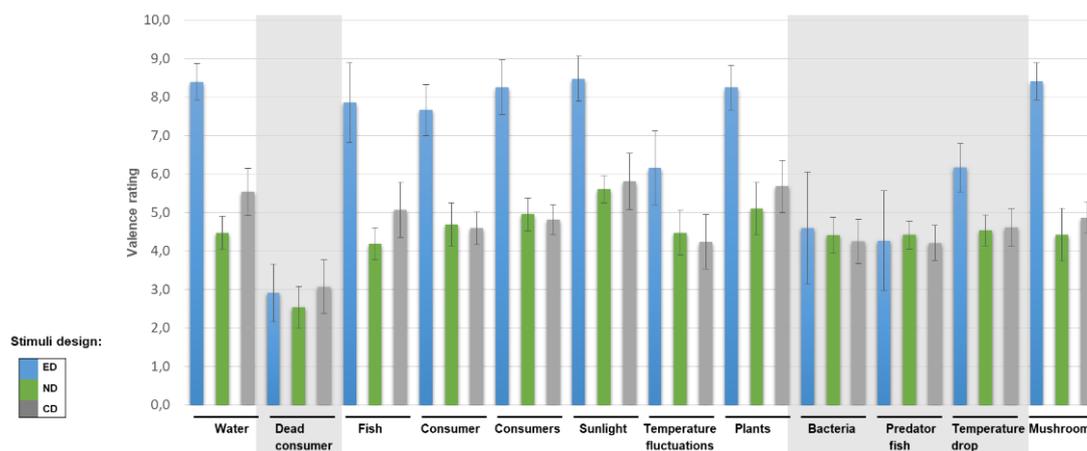


Figure 1. Mean valence rating and standard deviations for different concepts in *Emotional Design* (ED), neutral design (ND), and control design (CD). The higher the valence ratings, the more pleasant the concept is perceived. Concepts that differ significantly in valence ratings from the other concepts are highlighted in grey. Highlighting is based on post-hoc analyses which can be viewed in the supplementary material (S2).

Interim Discussion

The valence measurement aimed to investigate whether stimuli in *Emotional Design* are perceived as more pleasant than non-*Emotional Design* stimuli. As expected, the *Emotional Design* increased the perceived pleasantness (higher valence ratings) compared to the neutral and control designs (cf. H1). Interestingly, this is true only for concepts representing positive or neutral content (e.g., sunlight, plants, and fish). Valence ratings did not differ between emotional, neutral, and control designs in the case of concepts representing negative content (e.g., dead consumer, bacteria, and predator fish).

This suggests that the *Emotional Design* increases perceived pleasantness (i.e., higher valence ratings) through positive facial expressions (e.g., warm colors and round face-like shapes). In contrast, negative facial expressions (e.g., saturated red hues and angry eyes) did not increase the perceived pleasantness of a stimulus to the same extent. Moreover, in *Emotional Design*, concepts representing positive content were judged more pleasantly than concepts representing negative content. A similar result has been reported by Navratil, Köhl, and Heidig (2018), who observed that students rated the pleasantness of intended-positive *Emotional Design* higher than the pleasantness of intended-negative *Emotional Design*. This effect, however, was only observed when the students were exposed to both conditions (within-subjects design). If the students were assigned to one condition only (between-subjects design), the intended-positive and intended-negative *Emotional Designs* were perceived equally positively. Furthermore, Navratil et al. (2018) presume that learners prefer design manipulations that fit the design (e.g., a concept representing negative content is depicted in an intended-negative *Emotional Design*).

Experiment II

Method

Sample

A *a priori* power analysis was performed to determine sample size requirements for testing hypotheses 2 (**H2**) and 3 (**H3**) by rANOVA. We specified an α error level of .05 and a power of $1 - \beta = .95$ for both analyses (cf. Whitley & Ball, 2002). As Uzun and Yıldırım (2018) reported an effect size of $f = .40$ for a similar problem, we entered this effect size into our computations concerning **H2**. Concerning **H3**, we were guided by a study by Park et al. (2015), who observed effect size of $d = .55$. The *a priori* power analyses showed adequate sample sizes of 18 and 31 participants for **H2** and **H3** testing, respectively.

A total of 30 German junior high school students from grades 8 and 9 (58 % female; age: $M = 14.31$, $SD = 0.54$) were recruited for **experiment II**. They received small gifts as an expense allowance (e.g., pens or candy bars). Three students refused the EDA measurements, another three students produced incomplete data, and six students were non-responders (no usable SCR, see below). This results in an adequate sample size of $n = 18$ for **H2** testing. Three students had to be excluded for eye-tracking measurements because of measurement errors or disturbances during the experiment, which resulted in a sample size of $n = 27$ students for **H3** testing.

Stimulus Material, Design, and Procedure

In **experiment II**, students were randomly assigned to one of three sets of 12 stimuli (nota bene: we refer to a concept [e.g. *water*; see **Figure 2**] that is presented in a particular design [emotional, neutral, or control design] as a stimulus). In each set of stimuli, the order of the concepts (water, dead consumer, and fish, etc.) was identical. Still, the design changed in a fixed order so that each concept was presented for eight seconds in each set in a different design (e.g., the first concept *water* in set 1 in neutral, in set 2 in emotional, and in set 3 in control design; see **Figure 2**). A neutral stimulus was

followed by an emotional stimulus, an emotional stimulus followed by a control stimulus, and a control stimulus followed by a neutral stimulus. Before the experiment started, the students were given a brief overview of the procedure, and the EDA sensor was attached to the wrist of the student's non-dominant arm. The eye-tracking system was then calibrated for each student using 5-point-calibration (viewing distance was approximately 15.7 to 19.7 inches). Immediately after the calibration, the instructor started the experiment. The duration of the whole session was about 15 minutes per student. We chose this comparatively complex research design to avoid order, minimize fatigue effects, and reduce the influence of the concept represented in evaluating the impact of the design. The stimulus material is available as supplementary material (S1).

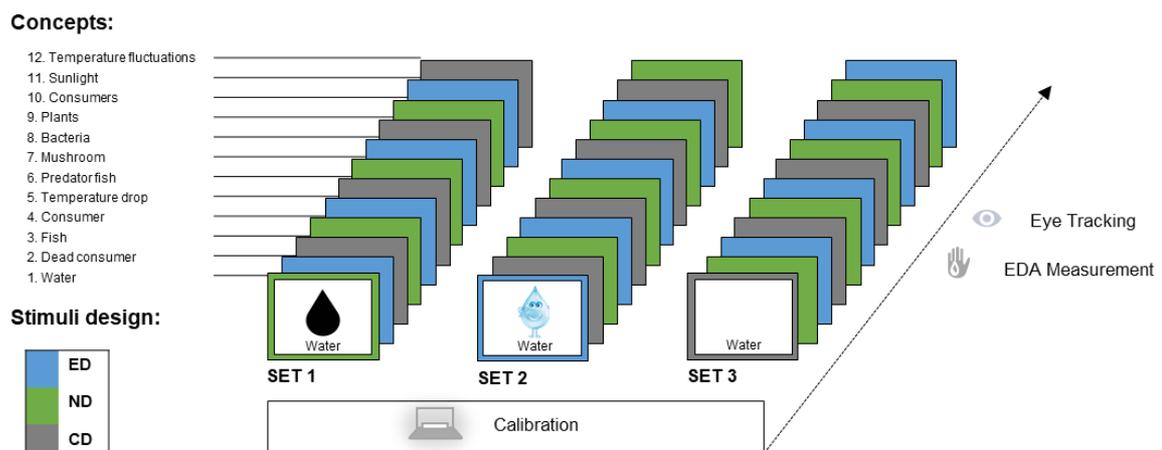


Figure 2. Research design and procedure. Students were assigned to one of three sets of stimuli. The twelve concepts (left) were displayed, with their design varying from one stimulus to the next (*Emotional Design* [ED], neutral design [ND], and control design [CD]).

Psychophysiological Measurements and Instruments

Several techniques (e.g., EDA measurements) are currently available which use psychophysiological parameters to measure arousal (Dzedzickis, Kaklauskas, & Bucinskas, 2020). EDA reveals implicit emotional responses that may occur without conscious awareness and is believed to provide an objective assessment of arousal

(Braithwaite et al., 2013). EDA is considered the only psychophysiological parameter that is sensitive to changes in emotional state and is not influenced by the activity of the parasympathetic system (Braithwaite et al., 2013). When an emotionally arousing stimulus is experienced, the skin moisture level changes due to an increase in the activity of eccrine sweat glands. The resulting amplification in electrical properties of the skin, the EDA, can reveal changes in the sympathetic nervous system.

EDA consists of two essential signals: the slowly changing tonic signal, skin conductance level (SCL), and the rapid changing phasic signal, skin conductance response (SCR). The SCL is a background signal, as only slight changes occur over minutes (Dzedzickis et al., 2020). SCRs arise in response to a single stimulus (event-related, ER-SCRs) but can also occur spontaneously (nonspecific skin conductance responses [NS-SCRs]; Boucsein, 2012). The onset of the ER-SCR typically occurs between 1 and 4 s after stimulus presentation (Dawson, Schell, & Filion, 2017), which means that “any SCR that begins between one and three (or four) seconds following stimulus onset is considered to be elicited by the stimulus” (Boucsein, 2012, p. 207). The parameters *ER-SCR amplitude* and *ER-SCR count* are used as indicators for arousal (Braithwaite et al., 2013). An *EdaMove 4* sensor (Movisens GmbH, Karlsruhe, Germany) was used with two self-adhesive Ag/AgCl electrodes. EDA was recorded with a sampling rate of 32 Hz by a constant voltage system (0.5 V).

Eye-tracking measurements were carried out to examine the influence of different design conditions on attention, with *dwelt time* as an indicator (see Stárková et al. [2019] for a similar approach). Here, dwell time is understood as the sum of durations of all fixations and saccades that hit an *Area of Interest* (AOI). An AOI is a delimited two-dimensional space on a stimulus for which metrics should be extracted (see **Figure 4**). To account for

different sized AOIs, all dwell time values reported are normalized (dwell time in milliseconds [ms] divided by AOI coverage). This way, it is a reliable measurement, as it adjusts the time spent to process an AOI to its relative surface in the screen (Ceravolo et al., 2019). The eye-tracking system *SMI RED250 mobile* (SensoMotoric Instruments GmbH, Teltow, Germany) was used on a 17-inch screen to measure the students' dwell time on different AOIs. As it is common in eye-tracking research, only the right eye position was analyzed (Lykins, Meana, & Kambe, 2006). According to the manufacturer, data were recorded at 60 Hz with an accuracy of 0.4°, and head movements were not restrained. *IViewX* was used for data collection, and stimuli were presented by *Experiment Center 3.0* (SensoMotoric Instruments GmbH, Teltow, Germany). Fixations were detected using the default dispersion-based algorithm (100 pixels and a minimum duration of 80 ms).

Signal Processing

EDA data (*SCR amplitude* and *SCR count*) were imported to IBM SPSS Statistics (version 24.0) using *SensorManager*, viewed and preprocessed using *UnisensViewer*, and analyzed using *DataAnalyzer* (Movisens GmbH, Karlsruhe, Germany). To test **H2**, each student's SCR amplitude and SCR count related to a specific design (emotional, neutral, or control design) were averaged. The first, second, and twelfth stimuli presented to the students were excluded from the analyses because the beginning and the end of the EDA measurements needed to be marked by tapping the device, which could cause SCRs not related to the treatment (see **Figure 3**). Hence, *SCR amplitude* and *SCR count data* for nine stimuli and each student (3 x emotional, 3 x neutral, and 3 x control design) remained for analyses (e.g., for neutral design in set 1: consumer, mushroom, consumers).

To extract SCR data, raw data were exported to *UnisensViewer*, automatically detecting SCR. Default settings by *Movisens* firmware to see SCRs (minimal rise time of

.05 $\mu\text{S/s}$, minimal amplitude of SCR of .1 μS , and a maximal rise time of SCR of .9 s) were used based on Dawson et al. (2017) and Boucsein (2012). Output intervals were set, and datasets were checked for artifacts that led to the exclusion of the respective students. ER-SCRs were classified and considered in the analyses when they occurred within a response window of 1 to 3 s following stimulus onset (Braithwaite et al., 2013; Dawson et al., 2017). This narrow response window was applied to reduce the possible contamination of measurements by NS-SCR signals. The parameters (*ER-SCR amplitude* and *ER-SCR count*) in the response window were chosen for hypothesis testing (see Figure 3) as Dawson et al. (2017) recommended.

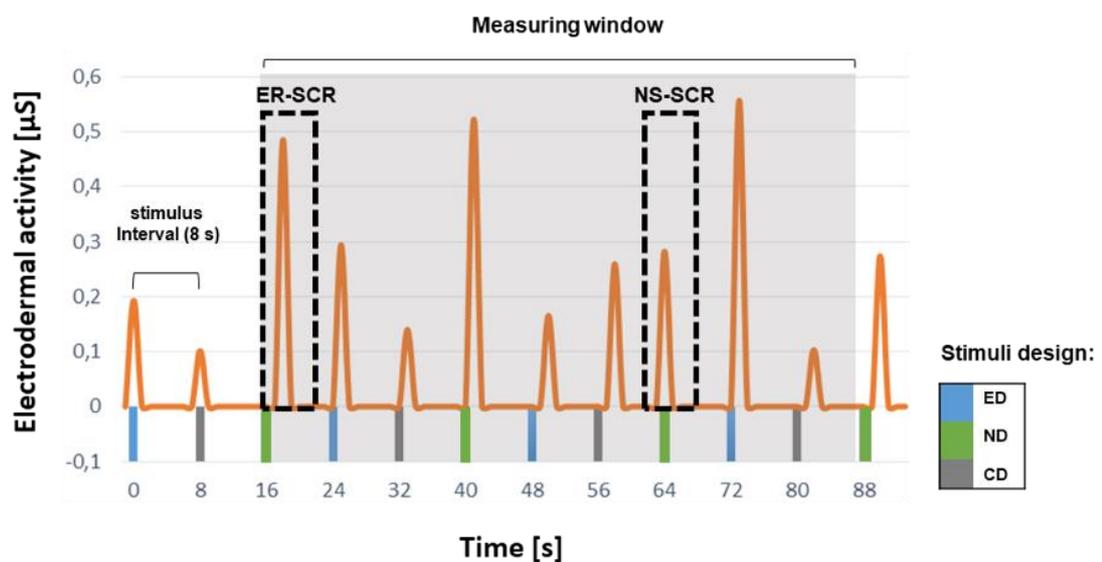


Figure 3. Exemplary SCR signals for set 2. The measurement window comprises nine stimuli (3 x emotional [ED], 3 x neutral [ND], and 3 x control design [CD]). Each stimulus is being viewed for 8 s. A signal which occurred within the chosen response window (1 to 3 s) is rated as an event-related SCR signal (ER-SCR). In contrast, a signal that occurred outside the response window is considered a non-specific SCR signal (NS-SCR).

Eye-tracking data were viewed and analyzed using *BeGaze* (*SensoMotoric Instruments GmbH*, Teltow, Germany) for all twelve stimuli. AOIs for the stimuli consisting of pictorial and textual elements (*Emotional Design* and neutral design) were pre-defined. Pictures were defined as $\text{AOI}_{\text{pictorial}}$, while all textual elements were defined as $\text{AOI}_{\text{textual}}$

(see **Figure 4**). Overlapping of AOIs was avoided, and margins were left between different AOIs (Holmqvist et al., 2011).

All students' eye-tracking and EDA data (from the three sets) were combined into one data set. Mean values of the different parameters (SCR count, SCR amplitude, and dwell time) were calculated for each design (*Emotional Design*, neutral design, and control design). This approach is considered valid, as the set allocation was randomized, and all test conditions were kept constant.

To test **H3**, attention (measured by *dwell time*) on pictorial and textual elements was averaged for each student resulting in four dependent variables (dwell time ED-AOI_{pictorial}, dwell time ED-AOI_{textual}, dwell time ND-AOI_{pictorial}, dwell time ND-AOI_{textual}). Dwell time ED-AOI_{pictorial} and dwell time ND-AOI_{pictorial} provide information about a student's attention to pictorial elements in *Emotional Design* (ED) and neutral design (ND), respectively. Dwell time ED-AOI_{textual} and dwell time ND-AOI_{textual} describe a student's attention spent on textual elements presented in combination with a pictorial element in emotional or neutral design (see **Figure 4**).

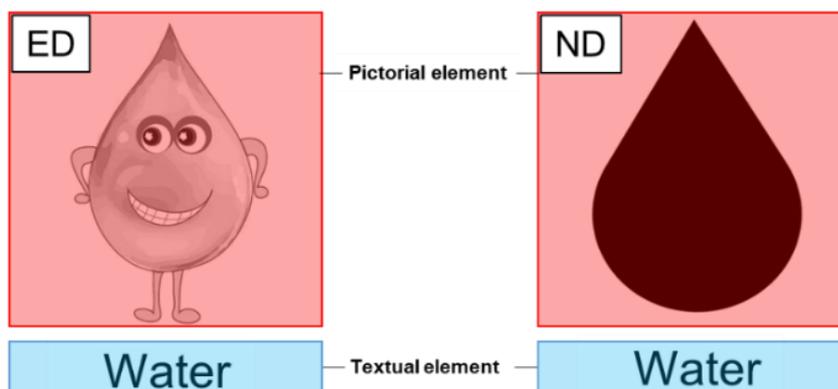


Figure 4. Sample presentation of areas of interest (AOIs) for the stimulus *water* (translated from German) in the *Emotional Design* (ED) and the neutral design (ND). Red: AOI_{pictorial}; blue: AOI_{textual}.

Results

Arousal

EDA measurements investigate whether different design principles affect arousal. According to **H2**, stimuli in the *Emotional Design* are expected to evoke emotions, leading to a higher state of arousal compared with concepts in the neutral and the control designs. Friedman's test revealed no significant difference for *SCR amplitude* ($\chi^2[2] = .03, p = 0.220$) and *SCR count* ($\chi^2[2] = 3.06, p = 0.217$) among the design conditions (lack of support for **H2**). Descriptive data are shown in **Table 2**.

Table 2. Mean ER-SCR amplitude, count, and standard deviations for design ($n = 18$).

Variables	ED		ND		CD	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ER-SCR amplitude (μs)	0.16	0.14	0.11	0.14	0.23	0.21
ER-SCR count (n)	0.35	0.27	0.26	0.27	0.44	0.38

ED = *Emotional Design*; ND = Neutral design; CD = Control design; ER-SCR = Event-related skin conductance response.

Attention

Regarding **H3**, it is presumed that pictorial elements in the *Emotional Design* attract more attention than pictorial elements in a neutral design. Wilcoxon signed-rank tests revealed that students' dwell time on pictorial elements in the *Emotional Design* ($M = 21862.92$ ms/coverage, $SD = 966.41$) was not greater than that on pictorial elements in the neutral design ($M = 22054.07$ ms/coverage, $SD = 997.95$), $Z = -1.92, p = .848$ (lack of support for **H3**).

Referring to the exploratory goal (**EG**) of our study, results show textual elements attract more attention ($M = 31770.42$ ms/coverage, $SD = 3307.93$) compared to pictorial elements ($M = 21958.50$ ms/coverage, $SD = 859.95$), $Z = -2.31, p = .042$ (strong effect with $d = 0.992$). Additionally, textual elements which were co-represented with neutral

design pictures ($M = 36091.51$ ms/coverage, $SD = 4025.91$) attracted significantly more *attention* than those that were presented together with *Emotional Design* pictures ($M = 27449.33$ ms/coverage, $SD = 2909.25$), $Z = -3.22$, $p = .002$ (strong effect with $d = 1.579$). Bonferroni-Holm correction was applied. Descriptive data are shown in **Table 3**.

Table 3. Mean dwell times and standard deviations for design ($n = 27$).

Design	ED		ND	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dwell time [ms/coverage] AOIpictorial	21862.92	966.41	22054.07	997.95
Dwell time [ms/coverage] AOItextual	27449.33	2909.25	36091.51	4025.91

ED = *Emotional Design*; ND = Neutral design; AOIpictorial = Area of interest represents attention on pictorial elements; AOItextual = Area of interest represents attention on textual elements.

Interim Discussion

This experiment aimed to provide additional objective insights into whether *Emotional Design* influences *arousal* and *attention*.

Arousal

The aim of the EDA measurements was to infer affective state from psychophysiological responses of the skin. Stimuli in the *Emotional Design* were expected to cause higher arousal measured by skin conductivity. However, the obtained results of this study did not support the hypothesis. The *Emotional Design* did not lead to a higher arousal than the neutral or the control designs (lack of support for H2). Similar results have been reported from several other studies (Münchow & Bannert, 2019; Park et al., 2015) who were also unable to find effects of *Emotional Design* on affective status. Stárková et al. (2019) assumed the affective influence of *Emotional Design* to be small in general and difficult to detect (supported by the meta-analysis of Wong and Adesope [2020]) and that the affective effect might be influenced by other variables that have not been identified

yet. It could also play a role whether the *Emotional Design* is presented statically or animated and what degree of humanization is used.

Previous studies in which *Emotional Design* increased the arousal (measured via heart rate variability) applied either animated stimuli with narration (Le et al., 2018) or static stimuli with sound effects (Bülbül & Abdullah, 2021; Uzun & Yıldırım, 2018). They could be seen as more complex *Emotional Designs* compared to the basic *Emotional Design* manipulation used in this present study and in other studies (e.g., Münchow et al., 2017; Um et al., 2012). At present, it is not entirely clear which design criteria must be fulfilled by an *Emotional Design* to evoke emotions. Characteristics that may vary in pictorial *Emotional Design* include colors, sounds, shapes, and humanization of objects. Thus, studies summarized under the term *Emotional Design* incorporate investigations in which only one characteristic was manipulated (color, Heidig, Müller, & Reichelt, 2015) or subtle emotional interface design was applied (Peng et al., 2021) but also those where many characteristics were varied (Bülbül & Abdullah, 2021; Uzun & Yıldırım, 2018). The study by Bülbül and Abdullah (2021) underlines this as they examined the effects of *Emotional Design* ranging from basic manipulations covering shape, simple anthropomorphisms and warm colors as suggested by, for example, Um et al. (2012) to complex animations with “emotionally connectible characters ... [and]...emotionally complementary sound effects” (pp. 348-349). While they found more complex *Emotional Design* animations to induce emotions, the basic *Emotional Design* features which were applied in most studies in the field “were not proven to be effective to increase educational outcomes and even to induce emotions” (Bülbül & Abdullah, 2021, p. 353). Furthermore, although all material that have been examined so far in *Emotional Design* research can be classified as multimedia learning material, a more detailed classification is needed. Weidenmann (2002) proposed a classification into the three categories *multimediality*,

multicodality, and *multimodality*. Accordingly, multimedial is characterized by the integrated use of technical devices, such as a computer and DVD player, while multicodal refers to the use of various symbolic systems (e.g., combinations of picture and text). Meanwhile, a single picture or text is classified as monocodal. Audio-visual learning software consisting of pictures and sounds would be categorized as multimodal because various sensory modalities are addressed (sight and hearing; for more information see Weidenmann [2002]). According to this categorization, it is not only the complexity/intensity of the *Emotional Design* that must be considered but also the way in which it is presented. If this classification is applied to previous studies on *Emotional Design*, mainly monomedial (computer-based) and primarily multicodal settings (Münchow & Bannert, 2019; Münchow et al., 2017; Shangguan et al., 2020; Stárková et al., 2019) or multicodal-multimodal settings (Le et al., 2018; Plass et al., 2014; Um et al., 2012; Uzun & Yıldırım, 2018) were used. Future studies and meta-analyses could take these classifications into account.

Why arousal was not increased for the *Emotional Design* could also have a different reason. As mentioned above, words (as in the control design) alone may elicit strong emotional reactions as the word leaves room for a stronger emotional association. The word “sunlight” for example could trigger a positive association with the last summer vacation and thus cause a strong emotional reaction, possibly even stronger than the anthropomorphized sun depicted in the *Emotional Design*. This could explain why the descriptive arousal measurements of the control design (plain text) are higher than those of the *Emotional Design*.

Another possible explanation could be that the *Emotional Design* induced changes in the affective state, but the EDA measurement was not able or not sensitive enough to capture them (see Limitations). In this light, findings and related interpretations should be treated

context-related and further research is needed in this area.

Attention

The aim of the eye-tracking measurements was the examination whether stimuli in *Emotional Design* attract more attention compared to stimuli in neutral design. Contrarily to the expectation, results show no differences in dwell time between pictures in the emotional and the neutral design (lack of support of **H3**). These results are conflicting with the *emotion-related attentional bias* but are in line with findings by Stárková et al. (2019), who could not find a significant difference for the total dwell time comparing *Emotional Design* and neutral design pictures. However, they observed that *Emotional Design* pictures attracted significantly more attention during initial observation (first two seconds of viewing). Park et al. (2015) found longer overall fixation durations for *Emotional Design* pictures (expressive anthropomorphisms) compared with pictures in a neutral design. There are various possible reasons for these conflicting results. Differences in the implementation of *Emotional Design* could cause differences in arousal intensity, as already discussed for the EDA measurements. Furthermore, the results suggest that textual elements attracted more attention than pictorial elements in picture-text combinations, corresponding with findings by Stárková et al. (2019) and the meta-study by Alemdag and Cagiltay (2018). Interestingly, the explorative examination revealed that textual elements attracted more attention when the co-represented pictures were shown in a neutral design compared with pictures in an *Emotional Design* (**EG**). One possible explanation could be that the abstractness of the neutral design pictures may cause difficulty in recognizing the concept visualized. To compensate this lack of information, an intense study of the textual element is necessary. *Emotional Design* pictures include colors and more details which could facilitate the recognition of the depicted concept. Supporting this, Mason, Pluchino, Tornatora, and Ariasi (2013) claimed that “more detailed

iconic diagrams ... are less abstract and closer to the referent they depict ... [and] are potentially easier to interpret” (p. 358). Inline, Lin, Holmqvist, Miyoshi, and Ashida (2017) found that “readers focused more on the text when learning with simplified illustrations” (p. 578). Another explanation could be based on the limited capacity assumption. Colors and humanizations presented in the *Emotional Design* pictures may be more cognitively demanding regarding picture-text integration compared to the neutral design. Therefore, the extraneous cognitive load could be higher for *Emotional Design* stimuli reducing the capacity available for text processing (see e.g., Sweller et al. [2005] for more information on cognitive load). Similarly, Koedinger, Alibali, and Nathan (2008) found that abstract representations put fewer demands on working memory compared to detailed representations.

Summary

From the present study, it can be concluded that the *Emotional Design* influences the affective state via the valence dimension. However, this does not seem to apply to negative content and correspondingly designed stimuli (use of negative facial expressions). In agreement with the presumption that the *Emotional Design* should be coherent with the content (Navratil et al., 2018), future studies should address a potential content- and coherence- dependent effect of *Emotional Design* manipulations. Furthermore, although *Emotional Design* stimuli were perceived as more pleasant, they did not seem to attract more attention. No differences could be found for the arousal dimension compared to non-*Emotional Designs*. This is supported by the presumption of Vogt et al. (2008) that attention is not driven by valence but by arousal intensity.

Limitations

Strength and limitation at the same time is the fact that this study was not carried out as part of a learning setting. Disruptive influences generated by a learning process

could be reduced but comparing the results of the present study to those of previous studies is only possible to a limited extent. Since a learning process is more complex than the mere viewing of different stimuli, the *Emotional Design* could have a beneficial, for example, an attention-directing effect in learning settings, i.e., a *cueing effect* (van Gog & Scheiter, 2010) which was not possible to examine in the present study. EDA measurements are “a sensitive index of emotional processing” (Braithwaite et al., 2013, p. 3) but aspects of analysis could reduce validity. To assign arousal to the individual stimuli (identify ER-SCR), a too strict sorting out of supposedly unspecific signals (NS-SCRs) can harbor the risk of missing ER-SCRs. This misclassification of NS-SCR signals could lead to a reduction of sensitivity (Braithwaite et al., 2013). To avoid those errors in future studies, data analysis using software that is particularly suited to determining ER-SCRs (e.g., *BIOPAC MP36R* or *Acq Knowledge Software*) is recommended. To enable a better assignment of the ER-SCRs to the individual stimuli, intermissions between the individual stimuli (blank screens) should be applied. Referring to the a priori power analysis, this investigation had sufficient power and the effect found was even greater than previously assumed. However, for future studies on static *Emotional Design*, we recommend larger sample sizes. Thereby the high drop-out rate (non-responders and measuring faults) should be considered in sample acquisition. In this light, the findings of the present study and related interpretations should be treated context-related and further research is needed in this area.

Conclusion

The *Emotional Design* of educational material is particularly interesting for school learning as its effect on affective state and learning is supposed to be high especially in younger age groups (Schneider, Häßler, Habermeyer, Beege, & Rey, 2019; Wong & Adesope, 2021). Most studies in the field have so far been carried out on college students

(Brom et al., 2018). As Valsiner and Connolly (2003) assume that younger students are less able to perceive their affective state than college students (Shangguan et al., 2020), it is unclear whether previous results are applicable to younger students.

In this study, static pictures applying a basic *Emotional Design* as developed and evaluated by Um et al. (2012) were created. This design was observed to influence valence but not arousal. However, the degree of arousal is considered as particularly relevant in the context of learning (Chung & Cheon, 2020; Irrazabal & Burin, 2021; McGaugh & Roozendaal, 2002; Schneider et al., 2019). Schneider et al. (2019) suspect that there is “an optimum range of arousal, which is not too low (no activation) and not too high (avoidance behavior), but instead activates learners” (p. 73). An *Emotional Design* apparently needs a certain intensity to induce arousal, but should not be too strong, otherwise the stimulus could have a distracting effect (see seductive detail effect; Park, Moreno, Seufert, & Brünken, 2011). Supporting this, Bülbül and Abdullah (2021) achieved convincing effects on emotions and educational variables such as transfer, interest, and intrinsic motivation for a complex, intense *Emotional Design* animation. Static pictures based on the *Emotional Design* criteria of Um et al. (2012), however lead to conflicting results regarding their influence on emotions and educational variables and is suspected to be not intense enough to induce emotions (e.g., Bülbül & Abdullah, 2021; Stárková et al., 2019).

Consequently, it can be suspected that the degree of arousal has a decisive influence on how emotionally a design is perceived. Therefore, future studies in the field should be based on the principle of *finding the correct balance*. Furthermore, understanding the effects and interactions of valence and arousal is still quite limited. More insights into these aspects could lead to more homogeneous results, not only about the effect on the

affective state but also on educational variables.

Moreover, the aspect of practicability should also be evaluated in future studies. The designs created by Bülbül and Abdullah (2021) that successfully induced emotions and positively influenced educational variables are very artistic and seem extremely laborious. If such claims are made on an effective *Emotional Design*, the practical implementation (in everyday school life) seems to be difficult. Thus, the feasibility should be considered, and costs should be weighed against benefits in future research.

Emotional Design represents an approach with great potential within multimedia learning. The present study represents an important step towards more objective measurements and sets new impulses in the field of *Emotional Design*. It would be desirable if future *Emotional Design* research will use more (psycho)physiological measurements despite the practical challenges involved.

List of Abbreviations

AOI: Area of interest; **CD**: Control design; **ED**: *Emotional Design*, **EDA**: Electrodermal activity; **ER-SCR**: Event-related skin conductance response; **ND**: Neutral design; **NS-SCR**: Nonspecific skin conductance responses; **SCR**: Skin conductance response; **SCL**: Skin conductance level.

Declarations

Ethics approval and consent to participate

Ethical review and approval were not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian / next of kin.

Availability of data and materials

The data presented in this study are available on request from the corresponding author.

Competing interests

The authors declare no conflict of interest. The authors have no relevant financial or non-financial interests to disclose.

Funding

This research received no external funding.

Authors' contributions

Conceptualization, S.L. and J.G.; methodology, S.L. and J.G.; formal analysis, S.L.; investigation, S.L.; writing: original draft preparation, S.L.; writing: review and editing, J.G.; visualization, S.L.; supervision, J.G.; All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The authors would like to thank our bachelor and master candidates (Judith Keusch, Sophie Telaak, Benjamin Bayer, Florian Schmitz, Sven Karsten, Leonie Mennicken, Rebecca Reiffen, and Tanja Treinen) for their participation in material development, data collection and data acquisition.

References

- Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education, 125*, 413-428.
- Bayram, S., & Bayraktar, D. M. (2012). Using eye tracking to study on attention and recall in multimedia learning environments: The effects of design in learning. *World Journal on Educational Technology, 4*(2), 81-98.
- Bisley, J. W. (2011). The neural basis of visual attention. *The Journal of Physiology, 589*(1), 49-57.
- Boucsein, W. (2012). *Electrodermal activity*. New York: Springer Science & Business Media.
- Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. *Psychophysiology, 49*(1), 1017-1034.
- Brom, C., Stárková, T., & D'Mello, S. K. (2018). How effective is *Emotional Design*? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review, 25*, 100-119. doi:10.1016/j.edurev.2018.09.004
- Brom, C., Stárková, T., Lukavský, J., Javora, O., & Bromová, E. (2016). Eye tracking in *Emotional Design* research: What are its limitations? *Proceedings of the 9th Nordic Conference on Human-Computer Interaction, 1-6*.
- Bülbül, A. H., & Abdullah, K. (2021). *Emotional Design* of educational animations: Effects on emotion, learning, motivation and interest. *Participatory Educational Research, 8*(3), 344-355.
- Caruelle, D., Gustafsson, A., Shams, P., & Lervik-Olsen, L. (2019). The use of electrodermal activity (EDA) measurement to understand consumer emotions – A literature review and a call for action. *Journal of Business Research, 104*, 146-160.
- Centers for Disease Control and Prevention (CDC). (2020). Handwashing is your superpower! [Poster]. Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/downloads/superhero_poster.pdf

- Ceravolo, M. G., Cerroni, R., Farina, V., Fattobene, L., Leonelli, L., Mercuri, N. B., & Raggetti, G. (2019). Attention allocation to financial information: the role of color and impulsivity personality trait. *Frontiers in Neuroscience, 13*, 818.
- Chung, S., & Cheon, J. (2020). *Emotional Design* of multimedia learning using background images with motivational cues. *Journal of Computer Assisted Learning, 36*(6), 922-932.
- Cohen, J. (1983). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Correia, P., Morais, P., Quintão, C., Quaresma, C., & Vigário, R. (2021). Assessing the emotional reaction to negative pictures through electrodermal activity data. *International Conference on Applied Human Factors and Ergonomics*, 116-124.
- Dawson, M. E., Schell, A. M., & Filion, D. L. (2017). *The electrodermal system*. New York: Cambridge University Press.
- Dzedzickis, A., Kaklauskas, A., & Bucinskas, V. (2020). Human emotion recognition: Review of sensors and methods. *Sensors, 20*(3), 592.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191.
- Harrison, A., Sullivan, S., Tchanturia, K., & Treasure, J. (2010). Emotional functioning in eating disorders: Attentional bias, emotion recognition and emotion regulation. *Psychological Medicine, 40*(11), 1887-1897.
- Heidig, S., Müller, J., & Reichelt, M. (2015). *Emotional Design* in multimedia learning: Differentiation on relevant design features and their effects on emotions and learning. *Computers in Human Behavior, 44*, 81-95. doi:10.1016/j.chb.2014.11.009
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.
- Irrazabal, N., & Burin, D. (2021). Effects of emotional valence and arousal on comprehension and assembly of instructions. *Trends in Psychology, 29*(1), 104-122.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329.

- Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science*, 32(2), 366-397.
- Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, 51(4), 336.
- Lang, P., & Bradley, M. M. (2007). The international affective picture system (IAPS) in the study of emotion and attention. *Handbook of Emotion Elicitation and Assessment*, 29, 70-73.
- Le, Y., Liu, J., Deng, C., & Dai, D. Y. (2018). Heart rate variability reflects the effects of *Emotional Design* principle on mental effort in multimedia learning. *Computers in Human Behavior*, 89, 40-47. doi:10.1016/j.chb.2018.07.037
- Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, 27(4), 363-384.
- Lin, Y. Y., Holmqvist, K., Miyoshi, K., & Ashida, H. (2017). Effects of detailed illustrations on science learning: An eye-tracking study. *Instructional Science*, 45(5), 557-581.
- Lykins, A. D., Meana, M., & Kambe, G. (2006). Detection of differential viewing patterns to erotic and non-erotic stimuli using eye-tracking methodology. *Archives of Sexual Behavior*, 35(5), 569-575.
- 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.
- Molina, M. (Producer). (2020). COVIBOOK: Supporting and reassuring children around the world. Retrieved from https://resourcecentre.savethechildren.net/node/17159/pdf/covi-en_0.pdf?embed=1
- Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*, 22(3), 149-158. doi:10.1111/j.1365-2729.2006.00170.x

- Moreno, R., & Mayer, R. E. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, 19(3), 309-326. doi:10.1007/s10648-007-9047-2
- Münchow, H., & Bannert, M. (2019). Feeling good, learning better? Effectivity of an *Emotional Design* procedure in multimedia learning. *Educational Psychology*, 39(4), 530-549. doi:10.1080/01443410.2018.1524852
- Münchow, H., Mengelkamp, C., & Bannert, M. (2017). The better you feel, the better you learn: Do warm colours and rounded shapes enhance learning outcome in multimedia learning? *Education Research International*, 2017. doi:10.1155/2017/2148139
- Mwambe, O. O., Tan, P. X., & Kamioka, E. (2021). Endogenous eye blinking rate to support human–automation interaction for e-learning multimedia content specification. *Education Sciences*, 11(2), 49.
- Navratil, S. D., Köhl, T., & Heidig, S. (2018). Why the cells look like that – The influence of learning with *Emotional Design* and elaborative interrogations. *Frontiers in Psychology*, 9, 1653. doi:10.3389/fpsyg.2018.01653
- North Rhine-Westphalian Ministry of Education Science and Research. (2020). Schulgesetz für das Land Nordrhein-Westfalen [School law for the state of North Rhine-Westphalia]. Retrieved from <https://bass.schul-welt.de/6043.htm>
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion*, 6(2), 257.
- Ohira, H., & Hirao, N. (2015). Analysis of skin conductance response during evaluation of preferences for cosmetic products. *Frontiers in Psychology*, 6, 103.
- Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). *Emotional Design* and positive emotions in multimedia learning: An eyetracking study on the use of anthropomorphisms. *Computers & Education*, 86, 30-42. doi:10.1016/j.compedu.2015.02.016
- Park, B., Moreno, R., Seufert, T., & Brünken, R. (2011). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior*, 27(1), 5-10. doi:10.1016/j.chb.2010.05.006
- Park, B., Plass, J. L., & Brünken, R. (2014). Cognitive and affective processes in multimedia learning.

- Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology, 98*(3), 583. doi:10.1037/0022-0663.98.3.583
- Peng, X., Xu, Q., Chen, Y., Zhou, C., Ge, Y., & Li, N. (2021). An eye tracking study: Positive emotional interface design facilitates learning outcomes in multimedia learning? *International Journal of Educational Technology in Higher Education, 18*(1), 1-18.
- Pijeira-Díaz, H. J., Drachsler, H., Järvelä, S., & Kirschner, P. A. (2016). Investigating collaborative learning success with physiological coupling indices based on electrodermal activity. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, 64-73*.
- Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). *Emotional Design* in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction, 29*, 128-140. doi:10.1016/j.learninstruc.2013.02.006
- Plass, J. L., Homer, B. D., MacNamara, A., Ober, T., Rose, M. C., Pawar, S., . . . Olsen, A. (2019). *Emotional Design* for digital games for learning: The effect of expression, color, shape, and dimensionality on the affective quality of game characters. *Learning and Instruction, 101194*.
- Posada-Quintero, H. F., & Chon, K. H. (2020). Innovations in electrodermal activity data collection and signal processing: A systematic review. *Sensors, 20*(2), 479.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372.
- Rayner, K. (2009). Eye movements in reading: Models and data. *Journal of Eye Movement Research, 2*(5), 1.
- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145.
- Save the Children. (2020). What is coronavirus? Comic – Ages under 7. Retrieved from <https://resourcecentre.savethechildren.net/library/what-coronavirus-comic-ages-under-7?embed=1>
- Schneider, S., Häßler, A., Habermeyer, T., Beege, M., & Rey, G. D. (2019). The more human, the higher the performance? Examining the effects of anthropomorphism on learning with media. *Journal of Educational Psychology, 111*(1), 57.

- Shangguan, C., Gong, S., Guo, Y., Wang, X., & Lu, J. (2020). The effects of *Emotional Design* on middle school students' multimedia learning: The role of learners' prior knowledge. *Educational Psychology*, 1-18. doi:10.1080/01443410.2020.1714548
- Stárková, T., Lukavský, J., Javora, O., & Brom, C. (2019). Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning? *Journal of Computer Assisted Learning*, 35(4), 555-568. doi:10.1111/jcal.12359
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer (Ed.), *Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press.
- Um, E., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). *Emotional Design* in multimedia learning. *Journal of Educational Psychology*, 104(2), 485. doi:10.1037/a0026609
- United Nations Children's Fund (UNICEF) (Producer). (2020, 27.07.2021). Activity book: Talking about COVID-19 with kids. A child-friendly guide to help young children cope with coronavirus. Retrieved from <https://www.unicef.org/jamaica/documents/activity-book-talking-about-covid-19-kids>
- Uzun, A. M., & Yıldırım, Z. (2018). Exploring the effect of using different levels of *Emotional Design* features in multimedia science learning. *Computers & Education*, 119, 112-128. doi:10.1016/j.compedu.2018.01.002
- Valsiner, J., & Connolly, K. J. (2003). The nature of development: The continuing dialogue of processes and outcomes. *Handbook of Developmental Psychology*, 9-18.
- van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, 20, 95-99.
- Vogt, J., de Houwer, J., Koster, E. H., van Damme, S., & Crombez, G. (2008). Allocation of spatial attention to emotional stimuli depends upon arousal and not valence. *Emotion*, 8(6), 880.
- von Reumont, F., & Budke, A. (2020). Strategies for successful learning with geographical comics: An eye-tracking study with young learners. *Education Sciences*, 10(10), 293.

- Weidenmann, B. (2002). Multicodierung und Multimodalität im Lernprozess [Multicoding and multimodality in the learning process]. In L. J. Issing & P. Klimsa (Eds.), *Information und Lernen mit Multimedia und Internet* [Information and Learning with Multimedia and Internet] (pp. 45-62): Beltz PVU.
- Wong, R. M., & Adesope, O. O. (2021). Meta-analysis of *Emotional Designs* in multimedia learning: A replication and extension study. *Educational Psychology Review*. doi:10.1007/s10648-020-09545-x
- World Medical Association. (2001). World medical association declaration of helsinki. Ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, 79(4), 373.
- Yao, Z., Wu, J., Zhang, Y., & Wang, Z. (2017). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behavior Research Methods*, 49(4), 1374-1385.

Gesamtdiskussion und Ausblick

10 Gesamtdiskussion

Ein zentraler Beweggrund für das Verfassen dieser Dissertation war, dass *Concept Maps* zwar eine effiziente Methode zur Vermittlung von Zusammenhangswissen darstellen können, diese Lernmethode sich aber im schulischen Kontext bisher nicht etablieren konnte.

Ein Erklärungsansatz für diesen Umstand ist die unklare Forschungslage in Bezug auf den benötigten Expertisegrad für einen lernförderlichen Einsatz von *Concept Maps*. Zudem wurde ein möglicher Unterschied bezüglich der benötigten Expertise für die Rezeption und Konstruktion von *Concept Maps* noch nicht differenziert betrachtet (Schroeder et al., 2017).

Bei der Rezeption von *Concept Maps* wurde beobachtet, dass Lernende, insbesondere Noviz:innen, einen sogenannten *Map Shock* erleben können, da sie mit dieser Darstellung der Informationen nicht vertraut sind (Blankenship & Dansereau, 2000). Angenommen wird, dass dieser *Map Shock* auf eine hohe kognitive Belastung zurückzuführen ist. Basierend auf der neueren Forschung im Bereich des multimedialen Lernens wird davon ausgegangen, dass die Vermittlung von Informationen in bestimmten Gestaltungsformen die kognitive Belastung senkt, die kognitiven Ressourcen erhöhen und somit die Lernleistung steigern kann.

Ziel dieser Dissertation war es, tiefere Erkenntnisse darüber zu erlangen, welche Faktoren das Lernen mit *Concept Maps* erleichtern können. Dieses Ziel wurde im Rahmen zweier Projekte verfolgt, welche sich unterschiedlichen Aspekten widmeten, die das Lernen mit *Concept Maps* potenziell verbessern können. Zum einen wurden neue Trainingsmaßnahmen entwickelt und hinsichtlich ihrer Lernförderlichkeit evaluiert und zum anderen wurde die Lernwirksamkeit unterschiedlich gestalteter *Concept Maps* untersucht.

Während die einzelnen Ergebnisse in den jeweiligen Publikationen bereits umfassend diskutiert wurden, werden die Befunde im Folgenden in Zusammenhang gebracht.

10.1 Die Relevanz von Concept Map-Trainingsmaßnahmen (Projekt 1)

Um die Relevanz von *Concept Map*-Trainingsmaßnahmen zu untersuchen, wurde im Rahmen dieses Projekts ein *Concept Map*-Konstruktionstraining und ein *Concept Map*-Rezeptionstraining entwickelt. Diese Trainingsmaßnahmen orientierten sich an kognitionspsychologischen Ansätzen und basieren auf der aktuellen Befundlage der kognitiven Lernforschung⁷⁰ sowie auf den Erkenntnissen und Empfehlungen bestehender Trainingsmaßnahmen (Jonassen et al., 1993; Nückles et al., 2004; Quinn et al., 2003; Sumfleth et al., 2010). Sie enthielten zum einen theoretische Informationen zu *Concept Maps* und boten zum anderen Übungsgelegenheiten zu einem der beiden *Concept Map*-Formate (Konstruktion oder Rezeption). Als Vergleich diente eine theoretische Kurzeinführung⁷¹, welche weder Informationen noch Übungsgelegenheiten zum Erlernen von *Concept Maps* bot.

Es zeigte sich, dass das Konstruktionstraining für die Fähigkeit⁷² *Concept Maps* zu erstellen förderlicher war, als das Rezeptionstraining oder die theoriebasierte Kurzeinführung. Daraus kann geschlossen werden, dass Übungsgelegenheiten notwendig sind, um kohärente *Concept Maps* erstellen zu können. Somit unterstützen die Ergebnisse dieses Projekts vorherige Befunde darin, dass eine ausschließlich theoriebasierte Einführung nicht ausreicht, um die Konstruktion kohärenter *Concept Maps* zu gewährleisten (z. B. Becker et al., 2021; den Elzen-Rump & Leutner, 2007; Großschedl & Tröbst, 2018; Sumfleth et al., 2010). Die Bestätigung dieser Befunde konnte um die Erkenntnis erweitert werden, dass in diesem Projekt auch die Übung im Umgang mit vorgefertigten *Concept Maps* keinen Vorteil für das Konstruieren von *Concept Maps* bot.

Das Konstruktionstraining wirkte sich nicht nur in Bezug auf die Erstellung von *Concept Maps* förderlich aus, sondern auch auf die Lernleistung. Hier war zu beobachten, dass Übung in der Konstruktion nicht nur die Lernleistung durch die Erstellung von *Concept Maps*, sondern auch die Lernleistung durch die Rezeption vorgefertigter *Concept Maps*

⁷⁰ Die Struktur der *Concept Map*-Trainingsmaßnahmen orientiert sich an der *Cognitive Apprenticeship Theory* (CAT; Collins et al., 1988). Hiernach lernen Noviz:innen indem das selbständige Lernen kontinuierlich gesteigert wird, bis sie in der Lage sind, eine Aufgabe selbstständig zu lösen. Im Laufe der Zeit nehmen die Aufgaben an Komplexität und Vielfalt zu, gleichzeitig werden Unterstützung und Feedback reduziert.

⁷¹ Die theoretische Kurzeinführung vermittelte dasselbe deklarative Wissen, was auch im Rahmen der *Concept Map*-Trainingsmaßnahmen vermittelt wurde, sie enthielt jedoch keine Übungsgelegenheit zur Rezeption oder Konstruktion von *Concept Maps*.

⁷² Die Analyse der erstellten *Concept Maps* (*Concept Map*-Qualität) diente als Indikator hierfür.

verbesserte. Es gibt sogar Hinweise dafür, dass das Konstruktionstraining die Rezeption besser fördert als das Rezeptionstraining.

Durch die im vorliegenden Projekt durchgeführten *Eye-Tracking*-Untersuchungen konnte die Beobachtung, dass sich ein Konstruktionstraining vorteilhaft auf die Rezeption auswirkt, untermauert werden. Es konnten Anhaltspunkte dafür gefunden werden, dass Lernende, welche die Konstruktion von *Concept Maps* trainierten, eine größere Expertise in der Rezeption aufwiesen, als Lernende, die im Rahmen eines Kontrolltrainings lediglich eine Kurzeinführung erhielten.

In Bezug auf die Struktur von *Concept Maps* wird angenommen, dass hierarchisch strukturierte *Concept Maps* für Noviz:innen vorteilhafter sein können, da im Vergleich zu einer netzartigen Struktur die Betrachtungsabfolge stärker vorgegeben ist (Blankenship & Dansereau, 2000). Diese Annahme konnte im Rahmen dieses Projekts jedoch nicht bestätigt werden. Die *Eye-Tracking*-Messungen ergaben keine Hinweise auf eine erleichterte Informationsaufnahme für Noviz:innen aus hierarchisch strukturierten *Concept Maps*.

Eine lernförderliche Wirkung der Expertise in der Rezeption auf die Konstruktion von *Concept Maps* konnte in diesem Projekt nicht beobachtet werden. Da beide Trainingsmaßnahmen hinsichtlich Zusammensetzung, Inhalt und Zeit identisch waren, scheinen die durch die verschiedenen Formate geschulten Fertigkeiten entscheidend zu sein. Fertigkeiten, die im Rahmen des Konstruktionstrainings erlernt wurden, veränderten höchstwahrscheinlich die Informationsverarbeitungsstrategie der Lernenden und ermöglichten ihnen so auch durch die Rezeption erfolgreich zu lernen. Da dies andersherum nicht der Fall war, kann geschlussfolgert werden, dass bei *Concept Maps* die Übung in dem aktiveren Prozess den passiveren Prozess ebenfalls schult. Auch in anderen Studien konnte ein sogenannter *Transfer Effect* beobachtet werden, welcher besagt, dass die Vertrautheit mit einem bestimmten Lernformat die Lernleistung in ähnlichen Formaten positiv beeinflusst (z. B. Royer, 1979). Hier wird einer aktiven Lernaktivität ebenfalls ein höheres Transferpotential zugeschrieben, als einer passiven (ebd.).

Die eingangs beschriebenen Befunde lassen die Schlussfolgerung zu, dass die Fähigkeit *Concept Maps* zu erstellen die Lernwirksamkeit dieser Methode maßgeblich bestimmt⁷³.

⁷³ Diese Aussage bezieht sich auf das Lernen mit *Concept Maps* durch deren Erstellung.

Diese Annahme ist auch im Forschungsgebiet verbreitet. Die Qualität der erstellten *Concept Maps* wird häufig als Indikator für den Lernerfolg verwendet (z. B. Austin & Shore, 1995). Eine noch nicht publizierte Untersuchung von Lenski et al. (in Vorbereitung) deutet jedoch darauf hin, dass nicht ohne weiteres von der *Concept Map*-Qualität auf die Lernleistung geschlossen werden kann. In dieser Studie zeigte sich, dass aus dem Gedächtnis erstellte *Concept Maps* eine geringere *Concept Map*-Qualität aufwiesen als *Concept Maps*, die einen Text als Grundlage hatten. In Bezug auf die Lernleistung schnitten Lernende, welche die *Concept Maps* aus dem Gedächtnis erstellten, allerdings trotz geringerer *Concept Map*-Qualität besser ab⁷⁴.

Basierend auf den Ergebnissen des vorliegenden Projekts könnte der Schluss, dass die *Concept Map*-Qualität die Lernleistung bedingt für das Konstruktions- und das Rezeptionstraining gezogen werden. Für die Kontrollgruppe ergab sich ein widersprüchliches Bild. Lernende, die das Kontrolltraining absolviert haben und in Bezug auf die Fähigkeit *Concept Maps* zu erstellen am schlechtesten abschnitten, waren im Hinblick auf die Lernleistung nicht unterlegen. Widererwartend war die Lernleistung bei den Lernenden am geringsten, welche am Rezeptionstraining teilnahmen und anschließend *Concept Maps* konstruierten.

Diese Befundlage ist verwunderlich, da sich das Rezeptionstraining und das Kontrolltraining hinsichtlich der Übungsgelegenheit zur Erstellung von *Concept Maps* nicht unterschieden. Es durfte daher erwartet werden, dass Lernende, die das Rezeptionstraining absolviert haben, mindestens dieselbe Lernleistung aufweisen würden, wie Lernende des Kontrolltrainings. Um dies zu erklären, lassen sich basierend auf den vorliegenden Ergebnissen nur Vermutungen anstellen. Lernende der Kontrollgruppe könnten aus Mangel an adäquater *Concept Map*-Expertise auf bereits bekannte Lernstrategien, wie beispielsweise das Notizenmachen, das Unterstreichen oder das wiederholte Lesen zurückgegriffen haben. Ein Hinweis für diese Annahme liefert die niedrige Qualität der von dieser

⁷⁴ Einschränkend muss zu der vorgestellten Studie jedoch gesagt werden, dass hier die Unterschiede in der Qualität nicht auf unterschiedliche Fertigkeiten in der Anwendung von *Concept Maps* beruhen, da alle Lernende im Vorfeld an derselben Trainingsmaßnahme teilnahmen. In diesem Fall beruhte der Unterschied laut den Autor:innen eher auf dem unterschiedlichen Grad der Vorwissensaktivierung und Elaboration welche in beiden Formaten gefordert wurde.

Gruppe erstellten *Concept Maps*. Diese niedrige Qualität könnte zum einen auf die mangelnde Fähigkeit *Concept Maps* zu erstellen zurückgeführt werden, zum anderen könnte sie dem Umstieg auf eine bevorzugte Lernstrategie geschuldet sein.

Somit könnten Lernende des Kontrolltrainings die Konstruktion entweder frühzeitig beendet haben oder sich während der Erstellung hauptsächlich mit dem Text und weniger mit dem Prozess der Erstellung der *Concept Map* beschäftigt haben. Diese Vermutung wird dadurch gestützt, dass das Notizenmachen und das Schreiben von Zusammenfassungen im Unterricht verbreitete Strategien sind, welche von den Lernenden somit besser beherrscht und daher wahrscheinlich bevorzugt eingesetzt werden (Hilbert & Renkl, 2008; Renkl & Nückles, 2006). Nach Renkl und Nückles (2006) müssen Lernende eine Lernstrategie zunächst selbst als produktiv wahrnehmen, um in diese zu vertrauen und sie anzuwenden. Dieses Ziel wurde möglicherweise durch die theoretische Kurzeinführung bei Lernenden des Kontrolltrainings nicht erreicht.

Durch die zunächst ernüchternd wirkenden Befunde drängt sich die Frage auf, warum das vermutliche Ausweichen auf eine bereits bekannte Lernstrategie zu besseren Lernergebnissen führt, als die Verwendung von *Concept Maps*. Aus der Qualität der *Concept Maps* kann geschlossen werden, dass Lernende, die an einer *Concept Map*-Trainingsmaßnahme teilgenommen haben, sich auch der Konstruktion gewidmet haben und wohl nicht auf eine andere Lernstrategie ausgewichen sind. Warum die Trainingsmaßnahmen im Vergleich zu dem Kontrolltraining nicht in einer besseren Lernleistung resultierten, könnte durch die Tatsache erklärt werden, dass diese neue Lernmethode für die Lernenden trotz Übungseinheiten ungewohnt war. Für trainierte Lernende, insbesondere die des Rezeptionsstrainings, ist der Automatisierungsprozess wahrscheinlich noch nicht abgeschlossen. Dies könnte erklären, warum diese Methode lerntechnisch noch nicht überlegen war.

Das beschriebene Phänomen wird in der Fachliteratur als *Nutzungsdefizit* bezeichnet (Artelt, 2000). Ein Nutzungsdefizit liegt vor, wenn eine spezifische Lernstrategie verwendet, aber nicht korrekt oder ineffizient ausgeführt wird, sodass die Lernenden nicht von ihr profitieren können (ebd.). Dies könnte laut Pressley et al. (1989) daran liegen, dass insbesondere jüngere Lernende die Qualität des eigenen Lernstrategiegebrauchs nicht korrekt einschätzen können. In Bezug auf den Zusammenhang zwischen der Expertise und der Selbsteinschätzung konnten Kruger und Dunning (1999) zeigen, dass Noviz:innen häufig nicht in der Lage sind ihre eigene Leistung korrekt einzuschätzen und dazu neigen

können, ihre eigenen Fähigkeiten zu überschätzen. Sie konnten auch zeigen, dass die Fähigkeit, die eigene Leistung akkurat einzuschätzen mit Zunahme der Fähigkeiten akkurater ausfällt. Im Rahmen dieses Projekts konnte dargelegt werden, dass Lernende ihre eigene Leistung besser einschätzen konnten, wenn sie am Konstruktionstraining teilgenommen haben, als Lernende, die das Rezeptionstraining absolvierten.

Dies liefert weitere Evidenz für die Annahme, dass das Konstruktionstraining, im Gegensatz zum Rezeptionstraining, dem Nutzungsdefizit eher entgegenwirken konnte. Trainingsmaßnahmen, die von kurzer Dauer oder nur theoriebasiert sind, könnten daher in einem Nutzungsdefizit resultieren. Da dieses Defizit nicht zwingend von den Lernenden als solches wahrgenommen wird, besteht die Gefahr, dass eine Art Kompetenzillusion entwickelt wird (Fürstenau, 2011).

In einer Studie von Karpicke und Blunt (2011) erzielten Lernende schlechtere Leistungen, wenn sie mit *Concept Maps* lernten, als Lernende, die durch das Abrufen vorher aufgenommener Informationen lernten (*Retrieval Practice*)⁷⁵. Dennoch waren die Lernenden, welche mit *Concept Maps* lernten, der Auffassung, dass sie besser abschneiden würden (mögliche Kompetenzillusion). Die Studie wurde jedoch dafür kritisiert, dass Lernende hier nur eine kurze Einführung zu *Concept Maps* erhielten (z. B. Mintzes et al., 2011). Es wird argumentiert, dass das Befundmuster auf die unzureichende Vertrautheit (Nutzungsdefizit) mit *Concept Maps* und weniger auf die Überlegenheit des Abrufprozesses zurückzuführen sei.

Um für dieses Projekt das gezeichnete Bild eines noch nicht abgeschlossenen Automatisierungsprozesses, also eines Nutzungsdefizites zu komplettieren, wäre für die kognitive Belastung zu erwarten gewesen, dass Lernende, die das Rezeptions- oder Kontrolltraining absolvierten, bei der Konstruktion von *Concept Maps* eine höhere kognitive Belastung erleben würden, als Lernende, die am Konstruktionstraining teilnahmen. Es konnte jedoch kein Unterschied in der kognitiven Belastung zwischen den verschiedenen Trainingsmaßnahmen festgestellt werden. Dies scheint auf methodische Limitationen zurückzuführen zu sein. Auch Becker et al. (2021), welche unterschiedliche *Concept Map*-Trainingsmaßnahmen hinsichtlich ihrer Wirksamkeit bei Studierenden mit dem gleichen Instrument untersuchten, beschrieben ähnliche methodische Einschränkungen. Somit bleibt

⁷⁵ Unter *Retrieval Practice* werden Abrufübungen verstanden. Hierunter fällt zum Beispiel die Wiedergabe gelesener Lerntexte aus dem Gedächtnis (Karpicke & Blunt, 2011).

unklar, inwiefern Trainingsmaßnahmen die kognitive Belastung beim Lernen reduzieren können.

In der Literatur konnte noch nicht eindeutig geklärt werden, ob das Lernen mittels der Konstruktion oder der Rezeption von *Concept Maps* lernförderlicher ist (Schroeder et al., 2017). Im Rahmen des vorliegenden Projekts konnte kein Lernunterschied zwischen dem Lernen durch das Konstruieren oder das Rezipieren von *Concept Maps* festgestellt werden. Wenn man ein mögliches Nutzungsdefizit berücksichtigt, könnte dieser Unterschied sich erst zeigen, wenn für beide Formate die Automatisierungsprozesse abgeschlossen sind und sie ihr volles Wirkpotenzial entfaltet haben.

10.2 Zusammenfassung und Ausblick in Bezug auf Concept Map-Trainingsmaßnahmen

Die bisherige methodische Ausrichtung im Forschungsbereich zu *Concept Map*-Trainingsmaßnahmen war auf die Konstruktion fokussiert und vernachlässigte die Frage, ob der Prozess des Rezipierens überhaupt trainiert werden muss. Die Ergebnisse dieses Projekts zeigen, dass unabhängig von der Wahl des Formats (Rezeption oder Konstruktion) ein vorgelagertes Konstruktionstraining die Lernleistung fördern kann (*Transfer Effect*).

Da die aktuelle Befundlage in Bezug auf die Wirksamkeit und Notwendigkeit von *Concept Maps*-Trainingsmaßnahmen widersprüchlich ist, liefert dieses Projekt wertvolle Erkenntnisse für künftige Studien und den Einsatz von *Concept Maps* im (Biologie-)unterricht.

Es kann daher festgehalten werden, dass das im Rahmen dieses Projekts konzipierte *Concept Map*-Training (mit all seinen bereits erwähnten Eigenschaften in Bezug auf z. B. Methode und Dauer) Lernende zur Konstruktion von *Concept Maps* befähigt. Aufgrund der vielen eingangs erwähneter Stellschrauben, an denen im Rahmen der Konzeption von Trainingsmaßnahmen gedreht werden kann, ist zu vermuten, dass die hier für die Sekundarstufe I konzipierten Trainingsmaßnahmen noch weiter optimiert werden können und sollten.

Renkl und Nückles (2006) merken diesbezüglich an, dass selbst kleine Variationen in der Nutzung von *Concept Maps* „zu erstaunlichen Unterschieden“ in Bezug auf die kogniti-

ven Prozesse und die Lernleistung führen können. Dies hat zur Folge, dass sich ein komplexes Wirkungsgeflecht ergibt, was auf der Forschungsebene zu widersprüchlichen Befunden führen kann. Dieses komplexe Wirkungsgeflecht zu ergründen und so die widersprüchlichen Befunde aufzulösen, sollte daher ein Hauptaugenmerk zukünftiger Forschung im Bereich der *Concept Map*-Trainingsmaßnahmen sein. Hierzu könnten die in **Kapitel 2.4** erwähnten unterschiedlichen Stellschrauben, welche im Rahmen einer *Concept Map*-Trainingsmaßnahme manipuliert werden können, systematisch untersucht werden. Ziel sollte es sein die optimale Dauer, Kombination aus theoretischer Anleitung, praktischer Übung, Themenwahl und dem Grad der nötigen Unterstützung zu ermitteln. In diesem Zusammenhang würde eine Übersichtsarbeit helfen, in der bereits entwickelte Trainingsmaßnahmen gesammelt, gegenübergestellt und hinsichtlich der genannten Aspekte miteinander verglichen werden. Da abhängig vom Alter und Lernniveau der Lernenden unterschiedliche Fördermaßnahmen empfohlen werden (Mandl & Friedrich, 2006; Renkl & Nückles, 2006), sollten auch diese Parameter in die Betrachtung miteinbezogen werden. Die hieraus gewonnenen Erkenntnisse könnten anschließend die Basis für zukünftige experimentelle Untersuchungen bilden.

Klärungsbedarf herrscht auch bezüglich der wichtigen Frage, welchen Einfluss die Fähigkeit *Concept Maps* zu nutzen, bzw. der vorhandene Automatisierungsgrad, auf die kognitive Belastung hat. Um diese Frage klären zu können sollten Messinstrumente verwendet werden, die eine zuverlässige Beurteilung der kognitiven Belastung während der Lernaktivität zulassen. DeLeeuw und Mayer (2008) empfehlen hierfür die Verwendung sogenannter *Dual-task* Messungen⁷⁶.

Obwohl gezeigt werden konnte, dass ein Konstruktionstraining die Lernleistung bei dem Rezipieren von *Concept Maps* verbessern kann, ist unklar, ob dies dabei helfen kann einen sogenannten *Map Shock* zu verhindern (Blankenship & Dansereau, 2000).

Aufgrund der in diesem Projekt durchgeführten *Eye-Tracking*-Untersuchungen wird für zukünftige Studien die Identifikation von für Expert:innen typischen *Scanpaths*⁷⁷ empfohlen.

⁷⁶ Bei *Dual-task* Messungen müssen Lernende neben der eigentlichen Lernaufgabe eine weitere Aufgabe erledigen. Aus der Analyse der Performanz, bei der Bewältigung dieser Aufgaben kann auf die durch die Lernaufgabe verursachte kognitive Belastung geschlossen werden.

⁷⁷ Ein *Scanpath* bezeichnet in der *Eye-Tracking*-Forschung die Visualisierung der Betrachtungsabfolge (Holmqvist et al., 2011).

Ein mögliches Ziel könnte die Entwicklung sogenannter *Eye Movement Modelling Examples*⁷⁸ (van Marlen et al., 2018) sein, welche Noviz:innen bei der Navigation durch die Inhalte von *Concept Maps* unterstützen.

Unter Berücksichtigung der Ergebnisse dieses Projekts sollte ein besonderes Augenmerk auf den Automatisierungsprozess gelegt werden. Es sollte in Zukunft zwischen dem Zeitpunkt, an dem Lernende in der Lage sind eine Methode korrekt einzusetzen und ab wann sich die Lernförderlichkeit der Methode einstellt, unterschieden werden. Auch Artelt et al. (2006, S. 17) betonen, dass eine lernförderliche Verwendung von Lernstrategien nicht durch ein kurzes Strategietraining erfolgt, sondern vielmehr eine „langfristige Gewohnheitsbildung“ nötig ist. Hierzu könnte genauer untersucht werden, ob und wie lange Lernende auf andere Lernstrategien ausweichen oder diese als Ergänzung zum Lernen mit *Concept Maps* einsetzen.

Da es im Forschungsbereich zu *Concept Maps* an systematischen Untersuchungen hinsichtlich der Lernwirksamkeit von Rezeption und Konstruktion mangelt, wird diese für künftige Studien empfohlen. Erst dann wäre ein direkter Vergleich der beiden *Concept Map*-Formate und eine Aussage in Bezug auf ihre Lernförderlichkeit möglich. In diesem Projekt konnte kein derartiger Unterschied festgestellt werden. Es besteht jedoch die Möglichkeit, dass dieser erst in Erscheinung tritt, wenn beide *Concept Map*-Formate ihr volles Wirkpotenzial erreicht haben.

Bei Vergleichen der beiden *Concept Map*-Formate sollte bedacht werden, dass der Rezeptionsprozess naturgemäß oftmals kürzer ausfällt, als der Konstruktionsprozess. Dieser Umstand sollte im Hinblick auf die Zeit, in der sich die Lernenden mit den *Concept Maps* beschäftigten, berücksichtigt werden. Zudem kann diese Beschäftigungszeit, welche ein Rezeptions- oder ein Konstruktionsprozess in Anspruch nimmt, mit dem Schweregrads der Thematik und der Form der Aufgabenstellung variieren.

10.3 Praktische Implikationen in Bezug auf Concept Map-Trainingsmaßnahmen

Auch wenn sich Lehrende bewusst sind, dass für Lernende das Bilden eines breiten Repertoires an Lernstrategien im Laufe der schulischen Laufbahn angestrebt werden sollte,

⁷⁸ Unter sogenannten *Eye Movement Modelling Examples* werden Videos verstanden, die das Blickverhalten einer eines Experten / einer Expertin (beim Lösen einer Aufgabe) visualisieren (Holmqvist et al., 2011).

tritt die Vermittlung von Lernstrategien im schulischen Kontext häufig in den Hintergrund (Mandl & Friedrich, 2006). Das vorliegende Projekt ermöglicht die Ableitung evidenzbasierter Handlungsempfehlungen zur Integration von *Concept Maps* in den Schulunterricht.

Wenn im Unterricht *Concept Maps* rezipiert oder konstruiert werden sollen, wird ein vorgelegertes Konstruktionstraining, das aus einer deklarativen Informationsvermittlung, gefolgt von einer prozeduralen Übungsphase besteht, empfohlen. Die Erkenntnis, dass man für die Verwendung beider Formate keine unterschiedlichen Trainingsmaßnahmen benötigt, könnte die Hemmschwelle *Concept Maps* zu nutzen, reduzieren.

Da die Nutzbarmachung der Lernförderlichkeit von *Concept Maps* für die Schulpraxis ein zentraler Beweggrund für die Durchführung dieser Dissertation war, wurde bei der Konzeption des Konstruktionstrainings explizit die praktische Umsetzbarkeit berücksichtigt⁷⁹. Um die Trainingsmaßnahmen in verschiedenen Schulfächern einsetzen zu können, wurden unterschiedliche Themenbereiche abgedeckt. Ohne ein entsprechendes Training sollten Lehrende den Schüler:innen vorgefertigte (Experten-) *Concept Maps* vorlegen oder sie mit Lücken-*Concept Maps* arbeiten lassen (Fischer & Mandl, 2000). Bei der Verwendung von vorgefertigten *Concept Maps* sollten zunächst solche verwendet werden, die eine geringe Anzahl an Konzepten aufweisen, um die Komplexität gering zu halten. Da in diesem Projekt für keine der beiden Strukturvarianten (hierarchisch oder netzartig) eine erhöhte Eignung festgestellt werden konnte, wird an dieser Stelle keine Einschränkung hinsichtlich der Struktur gemacht.

Im Rahmen dieses Projekts konnte gezeigt werden, dass eine ungefähr zweistündige Trainingsmaßnahme Lernende darin befähigen kann, *Concept Maps* zu erstellen. Zu vermuten ist jedoch, dass diese erst nach längerer Verwendung ihr ganzes Potential für Lehrende und Lernende entfalten. Renkl und Nückles (2006) raten zudem bei der Verwendung von Visualisierungstechniken, wozu auch *Concept Maps* zählen, zu einem reflektierten Praxiseinsatz. Dieser sieht eine an Lernziele und Lernniveaus der Schüler:innen angepasste Nutzung von Lernstrategien durch die Lehrenden vor.

⁷⁹ Das im Rahmen dieses Projekts für die Sekundarstufe I konzipierte Konstruktionstraining wurde für den Einsatz in der Schule aufbereitet und steht Lehrenden online in Form einer Handreichung mitsamt Lernmaterialien zur Verfügung (<https://doi.org/10.17605/OSF.IO/48A5W> ; Lenski & Großschedl; 2021).

10.4 Die Wirkung von Emotional Design-Abbildungen (Projekt 2)

Nach der *Cognitive Theory of Multimedia Learning* (Mayer, 2001), ist eine kombinierte Darstellung aus textlichen und bildlichen Informationen lernförderlicher und es tritt eine größere Reduzierung der kognitiven Belastung ein, als bei Informationen, die nur in Text- oder Bildform vermittelt werden. Ein relativ neuer Forschungszweig im Rahmen der Multimediaforschung spricht zudem motivationalen und affektiven Komponenten einen Einfluss auf das Lernen zu (Um et al., 2012). Fußend auf diesen Annahmen wurde das *Emotional Design* entwickelt. Hierbei handelt es sich um eine bestimmte Art der Gestaltung von Lernmaterialien zur Induktion von positiven Emotionen.

Orientiert an der Forschung zu multimedialem Lernen wurden in diesem Projekt verschiedene Abbildungen in textbasierte *Concept Maps* integriert. Ziel hierbei war es die kognitive Belastung der Lernenden beim Rezipieren von *Concept Maps* zu reduzieren und die Lernförderlichkeit zu erhöhen. Dabei wurden Abbildungen im *Emotional Design* und nicht-emotional gestaltete Versionen hiervon im Hinblick auf ihre kognitive und affektive Wirkung untersucht. Die ermittelte Wirkung wurde anschließend mit der von rein textbasierten *Concept Maps* verglichen. Zudem erfolgte bei den eingesetzten Bildern eine differenzierte Untersuchung in Bezug auf die Emotionsdimensionen Valenz und Aktivierung.

Die Valenz wurde durch Fragebögen ermittelt, während der Grad der Aktivierung durch Hautleitfähigkeitsmessungen ermittelt wurde. Um nähere Informationen in Bezug auf kognitive Prozesse zu gewinnen, wurden zudem *Eye-Tracking*-Messungen durchgeführt.

Bezogen auf die Lernleistung und die kognitive Belastung konnten in diesem Projekt keine Vorteile für *Emotional Design-Concept Maps* gegenüber Nicht-*Emotional Design-Concept Maps* ermittelt werden. Allerdings konnte beobachtet werden, dass die Integration von *Emotional Design*-Elementen die wahrgenommene Aufgabenschwierigkeit reduzieren kann, was in diesem Projekt jedoch nicht auf einen Unterschied in der Aufmerksamkeitslenkung zurückgeführt werden konnte. In Bezug auf den emotionalen Zustand,

wurde durch den Einsatz des in dem Forschungsgebiet gängigen Fragebogens⁸⁰ keine Induktion von positiven Emotionen, jedoch eine Reduktion des negativen Affekts ermittelt. Diesen Effekt erzielten jedoch auch nicht-emotional gestaltete Abbildungen.

Erwartungsgemäß wäre jedes *Emotional Design*, welches nicht in der Lage ist, positive Emotionen auszulösen, nicht als solches klassifizierbar. Wenn keine positiven Emotionen induziert werden, ist ein Ausbleiben der lernförderlichen Wirkung nicht verwunderlich. Stark et al. (2018) geben allerdings zu bedenken, dass bei einem derart engen Verständnis von *Emotional Design* viele der bisher entwickelten *Emotional Designs* nicht als solche klassifiziert werden dürften. Hier muss berücksichtigt werden, dass Emotionen nur schwer zu erfassen sind (Damasio, 2000). Dies spiegelt sich auch in bisherigen Befunden wider. *Emotional Designs*, welche nachweislich keine positiven Emotionen induzierten, konnten die Lernleistung steigern (Heidig et al. 2015; Münchow & Bannert; 2019), während keine Lernwirksamkeitssteigerung für nachweislich emotionsinduzierende Designs gefunden wurde (Uzun & Yildirim, 2018).

Ein möglicher Erklärungsansatz ist die in diesem Projekt vorgenommene differenzierte Betrachtung der Emotionsdimensionen. Es konnte ein unterschiedlicher Einfluss auf die beiden Dimensionen (Valenz und Aktivierung) beobachtet werden. Während die *Concept Maps* im *Emotional Design* die Valenz beeinflussten, konnte keine Veränderung bezüglich des Grads der Aktivierung festgestellt werden. Diesbezüglich wird angenommen, dass ein *Emotional Design* eine gewisse Intensität benötigt, um Lernende zu aktivieren. Die Art der Umsetzung kann somit die Wirkung von *Emotional Design* stark beeinflussen. Es gibt viele unterschiedliche Möglichkeiten Abbildungen im *Emotional Design* zu gestalten. Diese können in drei Kategorien eingeteilt werden und umfassen die *Nutzung der Gestaltungsprinzipien* (1), das *Lernmedium* (2) und die *Art der Darstellung* (3).

In Bezug auf die visuelle Gestaltung können lernrelevante Elemente in Form und Farbe variiert sowie anthropomorphisiert werden. Diese Variationen können allein (z. B. Farbe; Heidig et al., 2015) oder in Kombination eingesetzt werden (z. B. Um et al., 2012; Uzun & Yildirim, 2018).

⁸⁰ Im Rahmen der *Emotional Design*-Forschung wurde bislang hauptsächlich die *Positive and Negative Affect Scale* (PANAS; Watson et al., 1988) und Abwandlungen dessen verwendet (z. B. Um et al., 2012; Wong & Adesope, 2021).

Im Hinblick auf das Lernmedium kann beispielsweise zwischen papierbasiertem und digitalem Lernmaterial unterschieden werden⁸¹. Die Wahl des Lernmediums hat einen direkten Einfluss auf die Art der Darstellung, da digitale Formate etwa die Integration von akustisch dargebotenen Informationen (z. B. Töne oder eine Audiospur) sowie die Animation von Bildern ermöglichen. Papierbasiertes Lernmaterial ist auf die Darstellung statischer Abbildungen limitiert. Es wird vermutet, dass statische Bilder im *Emotional Design* nicht intensiv genug sind, um Emotionen zu induzieren. (z. B. Bülbül & Abdullah, 2021; Stárková et al., 2019). Eine Studie von Bülbül und Abdullah (2021) bestätigte dies. Hier konnte gezeigt werden, dass ein statisches *Emotional Design* die Emotionen und die Lernwirksamkeit kaum beeinflusste, während komplexe *Emotional Design*-Animationen sowohl positive Emotionen induzierten, als auch lernförderlich wirkten.

Da der Grad der Aktivierung im Kontext des Lernens als besonders relevant eingestuft wird (Irrazabal & Burin, 2021; Schneider et al., 2019), ist zu vermuten, dass dieser die Lernförderlichkeit des *Emotional Designs* maßgeblich bestimmt. Schneider et al. (2019) vermuten, dass es einen optimalen Aktivierungsbereich gibt. Falls von diesem Optimum nach oben oder nach unten abgewichen wird, besteht die Gefahr, dass dies ablenkend auf die Lernenden wirkt (Park et al., 2011).

Basierend auf den zusammengetragenen Erkenntnissen und Annahmen kann vermutet werden, dass die in diesem Projekt eingesetzten Abbildungen, wahrscheinlich nicht aktivierend genug waren, um die Lernwirksamkeit positiv zu beeinflussen. Ihre Wirkung auf die Valenz könnte jedoch dazu geführt haben, dass sie, verglichen mit den anderen *Concept Map*-Designs, als leichter empfunden wurden.

10.5 Zusammenfassung und Ausblick in Bezug auf Emotional Design-Abbildungen

Die Befürchtung, dass emotional gestaltete Bilder in *Concept Maps* eine zusätzliche (extrinsische) kognitive Belastung hervorrufen und somit die Lernleistung verschlechtern können, konnte in diesem Projekt entkräftet werden. Die Befunde legen nahe, dass die Integration von *Emotional Design*-Abbildungen im besten Fall positive Emotionen auslöst und das Lernen verbessert, während sie im schlimmsten Fall keine Auswirkungen

⁸¹ Bislang wurden ausschließlich computerbasierte Lernmaterialien im Rahmen der *Emotional Design*-Forschung untersucht (Münchow & Bannert, 2019; Münchow et al., 2017; Shangguan et al., 2020; Stárková et al., 2019; Um et al., 2012).

auf die Emotionen und das Lernen hat. Dieses Projekt konnte zudem zeigen, dass auf dem noch recht jungen Forschungsgebiet des *Emotional Designs* mehr Grundlagenforschung erforderlich ist, um spezifischere Aussagen über die lernförderliche Nutzung des *Emotional Designs* in *Concept Maps* zu ermöglichen.

Die Interpretation der bisherigen Befundlage zusammen mit den Befunden dieses Projekts legt die Vermutung nahe, dass die aktivierende Wirkung entscheidend für die Lernförderlichkeit ist. Ziel zukünftiger Studien sollte somit sein den erforderlichen Grad an Aktivierung zu finden, der für ein erfolgreiches Lernen mit *Emotional Design* notwendig ist.

Um dies gewährleisten zu können, müssen jedoch zunächst zuverlässige Messinstrumente zur Bestimmung des emotionalen Einflusses etabliert werden. Brom et al. (2018) empfehlen hierzu den stärkeren Einbezug psychophysiologischer Messinstrumente, wie die Herzfrequenzvariabilitäts- oder die Hautleitfähigkeitsmessung.

Ist die Grundlage für eine objektive Messung geschaffen, können unterschiedliche Merkmale des *Emotional Designs* im Hinblick auf ihre emotionale und lernförderliche Wirkung untersucht werden. Zukünftige Untersuchungen könnten sich beispielsweise mit dem Aspekt der Intensität befassen und nähere Informationen dazu liefern, welche Merkmale die Intensität des *Emotional Designs* beeinflussen. Darüber hinaus sind die Wechselwirkungen von Valenz und Aktivierung größtenteils noch unbekannt. Mehr Klarheit bezüglich dieses Zusammenspiels und weitere Erkenntnisse über die Wirkung einzelner Merkmale des *Emotional Designs* könnten zu einer homogenen Befundlage in Bezug auf die emotionale und lernbezogene Wirkung führen. Auch andere, noch nicht systematisch untersuchte Faktoren, wie der Lernkontext, das Alter der Lernenden oder die Art und Weise, wie das Lernmaterial den Lernenden präsentiert wird, können Einfluss auf die Wirkung des *Emotional Designs* haben (z.B. Brom et al., 2018; Münchow & Bannert, 2019; Stárková et al., 2019; Wong & Adesope, 2020). Zudem wurden persönliche Merkmale, wie der kulturelle Hintergrund der Lernenden (Stárková et al., 2019) oder die unterschiedlichen Wahrnehmungen von Ästhetik (Heidig et al., 2015), bislang noch nicht in die Untersuchungen miteinbezogen.

Gerade für den Einsatz in der Schule, spielt der Aufwand der Gestaltung eines Lernmaterials eine zentrale Rolle (Ballstaedt, 2007). In der Schulpraxis kann ein potentieller

Lernerfolg noch so groß sein, wenn der Zeitaufwand im Vergleich zu anderen (bekannteren) Methoden in einem Missverhältnis steht, ist eine Implementierung unwahrscheinlich. Deshalb sollten künftige Studien auch den Aspekt der Praktikabilität evaluieren. Das von Bülbül und Abdullah (2021) gestaltete Lernmaterial im *Emotional Design* induzierte zwar positive Emotionen und steigerte die Lernleistung. Aufgrund der anspruchsvollen künstlerischen Gestaltung und der Animationen war die Konzipierung vermutlich sehr zeitaufwändig.

Werden zu hohe Ansprüche an ein effektives *Emotional Design* gestellt, scheint die praktische Umsetzung im Schulalltag schwierig. Daher sollte die Prüfung der Umsetzbarkeit und eine Kosten-Nutzen-Abwägung Teil künftiger Forschung sein.

Im Kontext der Covid-19-Pandemie wurden Lernmaterialien (z. B. Videos, Poster und Bücher) entwickelt, um Schüler:innen über das Coronavirus aufzuklären und um Verhaltensregeln zu kommunizieren (Stiftung Jugend und Bildung in Zusammenarbeit mit dem Bundesamt für Bevölkerungsschutz und Katastrophenhilfe [BBK], n. d.). In diesen Materialien wurde das Coronavirus häufig anthropomorphisiert dargestellt, indem menschliche Gesichtsmarkmalen (z. B. Augen oder Münder) oder Extremitäten hinzugefügt wurden. Diese Darstellungen folgten häufig, jedoch wahrscheinlich unbeabsichtigt, den Gestaltungsprinzipien des *Emotional Designs*. Unklar ist jedoch, inwiefern diese Verwendung Fehlvorstellungen bei Lernenden ausgelöst hat. Ob ein *Emotional Design* Fehlvorstellungen hervorrufen kann, wurde in der Forschung noch nicht adressiert.

Insbesondere bei nicht direkt erfahrbaren Inhalten, wie beispielweise mikrobiologischen Lerninhalten (Wissen über Bakterien, Viren) scheinen Schüler:innen häufig fehlerhafte Vorstellungen zu entwickeln (Hilge, 1999; Hörsch & Kattmann, 2005). Die Verwendung des *Emotional Design* könnte besonders bei jüngeren Lernenden zu anthropomorphen Vorstellungen führen (Piaget, 1978). Hierunter versteht man die Übertragung von menschlichen Eigenschaften auf Tiere, Pflanzen und unbelebte Objekte (Johannssen, 2005).

Da anthropomorphe Denkweisen im Widerspruch zu (natur)wissenschaftlichen Erkenntnissen stehen, wird versucht diesen Vorstellungen entgegenzuwirken (ebd.). Daher nimmt die Untersuchung von Alltagsvorstellungen (Fehlvorstellungen; *Misconceptions*) eine wichtige Rolle in der fachdidaktischen Forschung ein (Menz & Seifried, 2022). Die

Frage, ob Schüler:innen durch die Verwendung von Lernmaterial im *Emotional Design* fehlerhafte Vorstellungen entwickeln, sollte daher in künftigen Studien adressiert werden.

Was die Übertragung des *Emotional Designs* auf *Concept Maps* betrifft, scheint es in Anbetracht der Forschungslage sinnvoll, zunächst die grundlegenden Fragestellungen, welche die Wirksamkeit des *Emotional Designs* betreffen, zu klären. Erst wenn hierzu mehr Erkenntnisse vorliegen erscheinen weitere Untersuchungen, welche sich der Wirksamkeit von *Emotional Design*-Elementen in *Concept Maps* widmen, sinnvoll.

10.6 Praktische Implikationen in Bezug auf Emotional Design-Abbildungen

Während für die Forschung die Generierung von evidenzbasiertem Wissens im Vordergrund steht, sind für Praktiker pädagogische Lösungen, die es ermöglichen konkrete Lernziele zu erreichen, von entscheidender Bedeutung. Solches Wissen erfordert Interventions- und Implementationsstudien, an denen es in Bezug auf die Verwendung des *Emotional Designs* (in *Concept Maps*) bislang noch mangelt. Die Frage, ob *Emotional Design-Concept Maps* im Unterricht verwendet werden sollten, kann somit nicht eindeutig beantwortet werden. Falls dennoch der Wunsch besteht, Abbildungen in *Concept Maps* aufzunehmen, wird an dieser Stelle dazu geraten, zunächst auf die Basis-Gestaltungsprinzipien der *Cognitive Theory of Multimedia Learning* (Mayer, 2001) zurückzugreifen. Dies wird empfohlen, da die Erstellung von *Emotional Design-Concept Maps* sehr aufwändig sein kann und ihre Überlegenheit gegenüber nicht-emotional gestalteten *Concept Maps* nicht erwiesen ist.

11 Fazit

Diese Dissertation beteiligt sich am aktuellen wissenschaftlichen Diskurs über die Relevanz von Trainingsmaßnahmen und den Nutzen von Multimedia-Gestaltungen für das Lernen.

Zusammen mit vorherigen Befunden leisten die im Rahmen dieser Dissertation gewonnenen Erkenntnisse einen wichtigen Beitrag zur Debatte über die Wirkung und Relevanz

von Trainingsmaßnahmen für die Nutzung von *Concept Maps*. Zudem konnten im Forschungsfeld des *Emotional Designs* neue Impulse gesetzt und wichtige Aspekte künftiger Forschung identifiziert werden.

Insgesamt liefert diese Dissertation einen wichtigen Erkenntnisgewinn an der Schnittstelle zwischen kognitionspsychologischer Forschung und anwendungsorientierter biologiepädagogischer Forschung.

Literaturverzeichnis

- Aben, B., Stapert, S., & Blokland, A. (2013). About the distinction between working memory and short-term memory. *Frontiers in Psychology, 3*(301), 1-20. <https://doi.org/10.3389/fpsyg.2012.00301>
- Abrams, R., Kothe, D., & Iuli, R. (2006). *Meaningful learning: A collaborative literature review of concept mapping*. California Consortium for Teacher Development Crown College. <http://www2.ucsc.edu/mlrg/clr-conceptmapping.html>
- Adesope, O. O., & Nesbit, J. C. (2013). Animated and static concept maps enhance learning from spoken narration. *Learning and Instruction, 27*, 1-10. <https://doi.org/10.1016/j.learninstruc.2013.02.002>
- Aguiar, J. G., & Correia, P. R. (2017). From representing to modelling knowledge: Proposing a two-step training for excellence in concept mapping. *Knowledge Management & E-Learning: An International Journal, 9*(3), 366-379. <https://doi.org/10.34105/j.kmel.2017.09.022>
- Akinsanya, C., & Williams, M. (2004). Concept mapping for meaningful learning. *Nurse Education Today, 24*(1), 41-46. [https://doi.org/10.1016/S0260-6917\(03\)00120-5](https://doi.org/10.1016/S0260-6917(03)00120-5)
- Alpert, S., & Grueneberg, K. (2001). *Multimedia in concept maps: A design rationale and web-based application*. Association for the Advancement of Computing in Education (AACE).
- Amadiou, F., van Gog, T., Paas, F., Tricot, A., & Mariné, C. (2009). Effects of prior knowledge and concept-map structure on disorientation, cognitive load, and learning. *Learning and Instruction 19*, 376-386.
- Anohina-Naumeca, A. (2015). Justifying the usage of concept mapping as a tool for the formative assessment of the structural knowledge of engineering students. *Knowledge Management & E-Learning: An International Journal, 7*(1), 56-72.
- Artelt, C. (2000). *Strategisches Lernen*. Waxmann.
- Artelt, C. (2006). Lernstrategien in der Schule. *Handbuch Lernstrategien*, 337-351.
- Aspinwall, L. G. (1998). Rethinking the role of positive affect in self-regulation. *Motivation and Emotion, 22*(1), 1-32.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Hrsg.), *Psychology of Learning and Motivation, 2*, 89-195. Elsevier.
- Austin, L. B., & Shore, B. M. (1995). Using concept mapping for assessment in physics. *Physics Education, 30*(1), 41.
- Ausubel, D. P. (1968). *Educational psychology - A cognitive view*. Holt, Rinehart and Winston.
- Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

- Baddeley, A. (2001). The concept of episodic memory. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1413), 1345-1350.
- Baddeley, A. (2006). Working memory: An overview. *Working Memory and Education*, 1-31.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, 63, 1-29.
- Ballstaedt, S.-P. (2007). Unterrichtsmaterial lerneffektiv gestalten. *Pädagogik*, 59(11), 22-25.
- Bannert, M., & Schnotz, W. (2006). Vorstellungsbilder und Imagery-Strategien. *Handbuch Lernstrategien*, 72-88.
- Becker-Carus, C., & Wendt, M. (2017). *Allgemeine Psychologie: Eine Einführung*. Springer-Verlag.
- Becker, L. B., Welter, V. D. E., Aschermann, E., & Großschedl, J. (2021). Comprehension-oriented learning of cell biology: Do different training conditions affect students' learning success differentially? *Education Sciences*, 11(8), 438.
- Beedie, C. J., Terry, P. C., Lane, A. M., & Devonport, T. J. (2011). Differential assessment of emotions and moods: Development and validation of the emotion and mood components of anxiety questionnaire. *Personality and Individual Differences*, 50(2), 228-233.
- Blankenship, J., & Dansereau, D. F. (2000). The effect of animated node-link displays on information recall. *The Journal of Experimental Education*, 68(4), 293-308.
- Bless, H., Clore, G. L., Schwarz, N., Golisano, V., Rabe, C., & Wölk, M. (1996). Mood and the use of scripts: Does a happy mood really lead to mindlessness? *Journal of Personality and Social Psychology*, 71(4), 665.
- Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology*, 106(3), 849-858.
- Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and toward structured representations. *Journal of Vision*, 11(5), 4-16.
- Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). *Visual long-term memory has a massive storage capacity for object details* [Konferenzbeitrag]. National Academy of Sciences, 105(38), 14325-14329.
- Brand, M., & Markowitsch, H. J. (2006). Was weiß die Hirnforschung über Lernen?. In Scheunpflug, A., & Wulf, C. (Hrsg.), *Zeitschrift für Erziehungswissenschaft* (S. 21-42). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90607-2_3
- Brandstätter, V., Schüler, J., Puca, R. M., & Lozo, L. (2018). *Motivation und Emotion*. Springer.
- Brom, C., Hannemann, T., Stárková, T., Bromová, E., & Děchtěrenko, F. (2016). *Anthropomorphic faces and funny graphics in an instructional animation may*

- improve superficial rather than deep learning: A quasi-experimental study* [Konferenzbeitrag]. European Conference on E-Learning (S. 89-97).
- Brom, C., Stárková, T., & D'Mello, S. K. (2018). How effective is emotional design? A meta-analysis on facial anthropomorphisms and pleasant colors during multimedia learning. *Educational Research Review*, 25, 100-119.
- Bryan, T., Mathur, S., & Sullivan, K. (1996). The impact of positive mood on learning. *Learning Disability Quarterly*, 19(3), 153-162.
- Buchner, A., & Brandt, M. (2017). Gedächtniskonzeptionen und Wissensrepräsentationen. In J. Müsseler & M. Rieger (Hrsg.), *Allgemeine Psychologie* (S. 401-434). Springer.
- Bülbül, A. H., & Abdullah, K. (2021). Emotional design of educational animations: Effects on emotion, learning, motivation and interest. *Participatory Educational Research*, 8(3), 344-355.
- Capota, K., van Hout, M., & van der Geest, T. (2007). *Measuring the emotional impact of websites: A study on combining a dimensional and discrete emotion approach in measuring visual appeal of university websites* [Konferenzbeitrag]. 2007 conference on Designing pleasurable products and interfaces (S. 135-147).
- Caria, A., de Falco, S., Venuti, P., Lee, S., Esposito, G., Rigo, P., Birbaumer, N., & Bornstein, M. H. (2012). Species-specific response to human infant faces in the premotor cortex. *NeuroImage*, 60(2), 884-893.
- Carver, C. (2003). Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect. *Cognition and Emotion*, 17(2), 241-261.
- Case, R. (1985). *Intellectual development: Birth to adulthood*. Academic Press.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293-332.
- Chang, K.-E., Sung, Y.-T., & Chen, I.-D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education*, 71(1), 5-23.
- Chiu, C. H. (2004). Evaluating system-based strategies for managing conflict in collaborative concept mapping. *Journal of Computer Assisted Learning*, 20(2), 124-132.
- Clark, R. C., & Mayer, R. E. (2008). Learning by viewing versus learning by doing: Evidence-based guidelines for principled learning environments. *Performance Improvement*, 47(9), 5-13.
- Clausen, S., & Christian, A. (2012). Concept Mapping als Messverfahren für den außerschulischen Bereich. *Journal für Didaktik der Biowissenschaften*, 3, 18-31.
- Collins, A., Brown, J. S., & Newman, S. E. (1988). Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics. *Thinking: The Journal of Philosophy for Children*, 8(1), 2-10.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407-428.

- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 240-247.
- Cowan, N. (1998). *Attention and memory: An integrated framework*. Oxford University Press.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Daley, B. J., Durning, S. J., & Torre, D. M. (2016). Using concept maps to create meaningful learning in medical education. *Mededpublish*. 5, 1-29.
- Damasio, A. (2012). Neuroscience and psychoanalysis: A natural alliance. *The Psychoanalytic Review*, 99(4), 591-594.
- Damasio, A. R. (2000). A second chance for emotion. *Cognitive Neuroscience of Emotion*, 12-23.
- Davidson, R. J., Sherer, K. R., & Goldsmith, H. H. (2009). *Handbook of Affective Sciences*. Oxford University Press.
- DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, 100(1), 223.
- Demmerling, C. (2013). Gefühle, Intentionalität, Leiblichkeit. In I. Günzler & K. Mertens (Hrsg.), *Wahrnehmen, Fühlen, Handeln* (S. 147-166). Phänomenologie im Wettstreit der Methoden.
- den Elzen-Rump, V., & Leutner, D. (2007). Naturwissenschaftliche Sachtexte verstehen - Ein computerbasiertes Trainingsprogramm für Schüler der 10. Jahrgangsstufe zum selbstregulierten Lernen mit einer Mapping-Strategie. In M. Landmann & B. Schmitz (Hrsg.), *Selbstregulation erfolgreich fördern* (S. 251-268). Kohlhammer.
- Dever, D. A., Wiedbusch, M. D., Cloude, E. B., Lester, J., & Azevedo, R. (2021). Emotions and the comprehension of single versus multiple texts during game-based learning. *Discourse Processes*. Online-Vorabpublikation. <https://doi.org/10.1080/0163853X.2021.1938878>
- Ekman, P. (1992). Facial expressions of emotion: An old controversy and new findings. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 335(1273), 63-69.
- Ekman, P., Friesen, W. V., & Tomkins, S. S. (1971). Facial affect scoring technique: A first validity study. *Semiotica*, 3, 37-58. <https://doi:10.1515/semi.1971.3.1.37>
- Ekman, P. E., & Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*. Oxford University Press.
- Fahr, A., & Hofer, M. (2013). Psychophysiologische Messmethoden. In W. Möhring & D. Schlütz (Hrsg.), *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft* (S. 347-365). Springer VS. https://doi.org/10.1007/978-3-531-18776-1_19

- Feldman Barrett, L., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, 74(4), 967–984. <https://doi.org/10.1037/0022-3514.74.4.967>
- Feller-Länzlinger, R., & Greder-Specht, C. (2002). *Fehler als Sprungbrett für Höhenflüge. Wenn Schülerinnen und Schüler aus Fehlern lernen: Eine Untersuchung über die Bedingungen eines lernfördernden Umgangs mit Fehlern in der Schule*. Soziothek.
- Ferrandiz, P. (2004). Classical conditioning. *Encyclopedia of Applied Psychology*, 1, 345.
- Fischer, F., & Mandl, H. (2000). Strategiemodellierung mit Expertenmaps. In H. Mandl & F. Fischer (Hrsg.), *Wissen sichtbar machen. Wissensmanagement mit Mapping-Techniken* (S. 37-54). Hogrefe-Verlag.
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Hrsg.), *The Nature of Intelligence* (S. 231-235). Erlbaum.
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, 56(3), 218–226. <https://doi.org/10.1037//0003-066x.56.3.218>
- Frenzel, A. C., Götz, T., & Pekrun, R. (2009). Emotionen. In E. Wild, & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 205-231). Springer. https://doi.org/10.1007/978-3-540-88573-3_9
- Friedrich, H. F., & Mandl, H. (1997). Analyse und Förderung selbstgesteuerten Lernens. In F. E. Weinert (Hrsg.), *Pädagogische Psychologie* (S. 254-293). Hogrefe.
- Frijda, N. H. (1986). *The emotions*. Cambridge University Press.
- Funke, J. & Spring, M. (2006). Methoden der Denk- und Problemlöseforschung. In J. Funke (Hrsg.), *Denken und Problemlösen* (S. S. 647-744). Hogrefe.
- Fürstenau, B. (2011). Concept Maps im Lehr-Lern-Kontext. *DIE Zeitschrift für Erwachsenenbildung*(1), 46-48.
- Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, 23(4), 523-552.
- Gehl, D. (2013). *Vom Betrachten zum Verstehen*. Springer.
- Gerrig, R. J., & Zimbardo, P. G. (2008). *Psychologie*. Pearson Deutschland GmbH.
- Gillet, N., Vallerand, R. J., Lafreniere, M.-A. K., & Bureau, J. S. (2013). The mediating role of positive and negative affect in the situational motivation-performance relationship. *Motivation and Emotion*, 37(3), 465-479.
- Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, 15(4), 313-331.
- Glocker, M. L., Langleben, D. D., Ruparel, K., Loughhead, J. W., Gur, R. C., & Sachser, N. (2009). Baby schema in infant faces induces cuteness perception and motivation for caretaking in adults. *Ethology*, 115(3), 257-263.
- Goetz, T., Zirngibl, A., Pekrun, R., & Hall, N. (2003). Emotions, learning and achievement from an educational-psychological perspective. In P. Mayring & C.

- von Rhoeneck (Hrsg.), *Learning emotions: The influence of affective factors on classroom learning* (S. 9-28). Peter Lang.
- Gouli, E., Gogoulou, A., & Grigoriadou, M. (2003). A coherent and integrated framework using concept maps for various educational assessment functions. *Journal of Information Technology Education: Research*, 2(1), 215-240.
- Graf, D. (2014): Concept Mapping als Diagnosewerkzeug. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 325–340). Springer Spektrum.
- Gross, J. J., & Feldman Barrett, L. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review*, 3(1), 8-16.
- Großschedl, J. (2010). *Einfluss ausgewählter instruktionaler Maßnahmen auf Struktur und Niveau zellbiologischen Wissens* [Dissertation]. Logos Verlag Berlin GmbH.
- Großschedl, J., & Harms, U. (2011). Concept mapping: Förderung der Metakognition oder metakognitiver Förderungsbedarf? [Tagungsband]. In H. Bayrhuber, U. Harms, B. Muszynski, B. Ralle, M. Rothgangel, L.-H. Schön, H. J. Vollmer & H.-G. Weigand (Hrsg.), *Empirische Fundierung in den Fachdidaktiken* (Band 1, S. 115-130). Waxmann-Verlag.
- Großschedl, J., & Tröbst, S. (2018). Biologie lernen durch Concept Mapping: Bedeutung eines Lernstrategietrainings für kognitive Belastung, kognitive Prozesse und Lernleistung–Kurzdarstellung des DFG–Projekts. *Zeitschrift für Didaktik der Biologie (ZDB)-Biologie Lehren und Lernen*, 22(1), 20-30.
- Hahn-Laudenberg, K. (2017). *Konzepte von Demokratie bei Schülerinnen und Schülern*. Springer.
- Hardy, I., & Stadelhofer, B. (2006). Concept Maps wirkungsvoll als Strukturierungshilfen einsetzen: Welche Rolle spielt die Selbstkonstruktion? *Zeitschrift für Pädagogische Psychologie*, 20(3), 175-187.
- Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, 89(1), 92.
- Harrison, A., Sullivan, S., Tchanturia, K., & Treasure, J. (2010). Emotional functioning in eating disorders: Attentional bias, emotion recognition and emotion regulation. *Psychological Medicine*, 40(11), 1887-1897.
- Hasselhorn, M., & Gold, A. (2006). *Pädagogische Psychologie: Erfolgreiches Lernen und Lehren*. Kohlhammer.
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Haudeck, H. (2008). *Fremdsprachliche Wortschatzarbeit außerhalb des Klassenzimmers: Eine qualitative Studie zu Lernstrategien und Lerntechniken in den Klassenstufen 5 und 8*. Gunter Narr Verlag Tübingen.
- Haugwitz, M., Nesbit, J. C., & Sandmann, A. (2010). Cognitive ability and the instructional efficacy of collaborative concept mapping. *Learning and individual differences*, 20(5), 536-543.

- Haugwitz, M., & Sandmann, A. (2009). Kooperatives Concept Mapping in Biologie: Effekte auf den Wissenserwerb und die Behaltensleistung. *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 89-107.
- Hay, D., Kinchin, I., & Lygo-Baker, S. (2008). Making learning visible: The role of concept mapping in higher education. *Studies in Higher Education*, 33(3), 295-311.
- Heckhausen, J., & Heckhausen, H. (2006). *Motivation und Handeln*. Springer-Verlag.
- Heidig, S., Müller, J., & Reichelt, M. (2015). Emotional design in multimedia learning: Differentiation on relevant design features and their effects on emotions and learning. *Computers in Human Behavior*, 44, 81-95.
- Hilbert, T. S., & Renkl, A. (2008). Concept mapping as a follow-up strategy to learning from texts: What characterizes good and poor mappers? *Instructional Science*, 36(1), 53-73.
- Hilge, C. (1999). *Schülervorstellungen und fachliche Vorstellungen zu Mikroorganismen und mikrobiellen Prozessen: Ein Beitrag zur didaktischen Rekonstruktion*. Didaktisches Zentrum.
- Holley, C. D., & Dansereau, D. F. (1984). The development of spatial learning strategies. In C. D. Holley & D. F. Dansereau (Hrsg.), *Spatial learning strategies. Techniques, applications, and related issues* (S. 3-19). Academic Press.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- Hörsch, C., & Kattmann, U. (2005). Schülervorstellungen zu Mikroorganismen und mikrobiellen Prozessen im Menschen I. *Erkenntnisweg Biologiedidaktik*, 7-19.
- Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, 77(1), 95-111.
- Ifenthaler, D. (2011). Identifying cross-domain distinguishing features of cognitive structure. *Educational Technology Research and Development*, 59(6), 817-840.
- Irrazabal, N., & Burin, D. (2021). Effects of emotional valence and arousal on comprehension and assembly of instructions. *Trends in Psychology*, 29(1), 104-122.
- Isen, A. M. (1987). Positive affect, cognitive processes, and social behavior. *Advances in Experimental Social Psychology*, 20, 203-253.
- Jaafarpour, M., Aazami, S., & Mozafari, M. (2016). Does concept mapping enhance learning outcome of nursing students? *Nurse Education Today*, 36, 129-132.
- Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge*. Lawrence Erlbaum Associates.

- Jonassen, D. H., Reeves, T. C., Hong, N., Harvey, D., & Peters, K. (1997). Concept mapping as cognitive learning and assessment tools. *Journal of Interactive Learning Research*, 8(3), 289.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329.
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, 38(1), 23-31.
- Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications: Introduction to the special issue. *Instructional Science*, 38(3), 209-215.
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772-775.
- Kelleher, J. D., & Dobnik, S. (2019). *Referring to the recently seen: Reference and perceptual memory in situated dialog*. ArXiv preprint.
- Killermann, W. (1995). *Biologieunterricht heute: Eine moderne Fachdidaktik*. Auer Verlag.
- Killermann, W. (2000). Ganzheitliche Naturschutz- und Umwelterziehung. *Bayerische Akademie für Naturschutz und Landschaftspflege (ANL)*, 24, 21-26.
- Kinchin, I. M. (2001). If concept mapping is so helpful to learning biology, why aren't we all doing it? *International Journal of Science Education*, 23(12), 1257-1269.
- Kirsh, D. (1997). Interactivity and multimedia interfaces. *Instructional Science*, 25(2), 79-96.
- Klein, D. C., Chung, G., Osmundson, E., Herl, H. E., & O'Neil, H. (2002). Examining the validity of knowledge mapping as a measure of elementary students' scientific understanding. *CSE Report*.
- Klepsch, M., & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science*, 48(1), 45-77.
- Knörzer, L., Brünken, R., & Park, B. (2016). Facilitators or suppressors: Effects of experimentally induced emotions on multimedia learning. *Learning and instruction*, 44, 97-107.
- Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: Why emotion matters. *Journal of Experimental Psychology: General*, 140(1), 14.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121.
- Lambiotte, J. G., Dansereau, D. F., Cross, D. R., & Reynolds, S. B. (1989). Multirelational semantic maps. *Educational Psychology Review*, 1(4), 331-367.
- Le, Y., Liu, J., Deng, C., & Dai, D. Y. (2018). Heart rate variability reflects the effects of emotional design principle on mental effort in multimedia learning. *Computers in Human Behavior*, 89, 40-47.

- Lea, R. G., Qualter, P., Davis, S. K., Pérez-González, J.-C., & Bangee, M. (2018). Trait emotional intelligence and attentional bias for positive emotion: An eye tracking study. *Personality and Individual Differences, 128*, 88-93.
- Lenski, S., Mustafa, M., & Großchedl, J. (in Vorbereitung). Concept mapping as effective retrieval-based learning opportunity in biology.
- Lenski, S., & Großschedl, J. (2021). Concept Maps im Unterricht: Eine Trainingseinheit für Schüler*innen der Sekundarstufe I.
<https://doi.org/https://doi.org/10.17605/OSF.IO/48A5W>
- Lenzner, A., Schnotz, W., & Müller, A. (2013). The role of decorative pictures in learning. *Instructional Science, 41*(5), 811-831.
- Leopold, C., & Leutner, D. (2002). Der Einsatz von Lernstrategien in einer konkreten Lernsituation bei Schülern unterschiedlicher Jahrgangsstufen. *Zeitschrift für Pädagogik, Beiheft 45*, 240-258.
- Lopez, E., Kim, J., Nandagopal, K., Cardin, N., Shavelson, R. J., & Penn, J. H. (2011). Validating the use of concept-mapping as a diagnostic assessment tool in organic chemistry: Implications for teaching. *Chemistry Education Research and Practice, 12*, 133-141.
- Lorenz, K., & Martin, R. (1971). *Studies in animal and human behaviour*. Harvard University Press.
- Mandl, H., & Fischer, F. (2000). *Wissen sichtbar machen*. Hogrefe.
- Mandl, H., & Friedrich, H. F. (2006). *Handbuch Lernstrategien*. Hogrefe.
- Mayer, R. E. (2001). A cognitive theory of multimedia learning In R. D. Mayer (Hrsg.), *Multimedia Learning* (S. 41-62). Cambridge University Press.
- Mayer, R. E. (2002). *Learning and instruction*. Merrill Prentice Hall.
- Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology, 93*(2), 390.
- Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction, 33*, 12-18.
- Mayer, R., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. Mayer (Hrsg.), *The Cambridge Handbook of Multimedia Learning* (S. 279-315). Cambridge: Cambridge University Press.
[doi:10.1017/CBO9781139547369.015](https://doi.org/10.1017/CBO9781139547369.015)
- Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology, 82*(4), 715.
- Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43-52.
- McAfoose, J., & Baune, B. (2009). Exploring visual – spatial working memory: A critical review of concepts and models. *Neuropsychology Review, 19*(1), 130-142.

- McCagg, E. C., & Dansereau, D. F. (1991). A convergent paradigm for examining knowledge mapping as a learning strategy. *The Journal of Educational Research*, 84(6), 317-324.
- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36(4), 475-492.
- McGuinness, C. (1990). Talking about thinking: The role of metacognition in teaching thinking. *Lines of Thinking*, 2, 310-312.
- Mees, U. (2006). Zum Forschungsstand der Emotionspsychologie – eine Skizze. In R. Schützeichel (Hrsg.), *Emotionen und Sozialtheorie. Disziplinäre Ansätze*, S. 104-123.
- Menz, C., & Seifried, E. (2022). Woher kommen pädagogisch-psychologische Fehlvorstellungen und wie kann man ihnen wirksam begegnen?. In G. Steins, B. Spinath, S. Dutke, M. Roth & M. Limbourg (Hrsg.), *Mythen, Fehlvorstellungen, Fehlkonzepte und Irrtümer in Schule und Unterricht* (S. 27-44). Psychologie in Bildung und Erziehung: Vom Wissen zum Handeln. Springer.
https://doi.org/10.1007/978-3-658-36260-7_2
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24(1), 167-202.
- Miller, N. L., Cañas, A. J., & Novak, J. D. (2008). Use of the Cmaptools recorder to explore acquisition of skill in concept mapping. In A. J. Cañas, P. Reiska, M. K. Åhlberg & J. D. Novak (Hrsg.), *3rd International Conference on Concept Mapping* (S. 674-681). Tallinn University.
- Mintzes, J. J., Canas, A., Coffey, J., Gorman, J., Gurley, L., Hoffman, R., Y., M. S., Miller, N., Moon, B., Trifone, J., & Wandersee, J. H. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping [Technischer Kommentar]. *Science*, 334, 453- 453.
- Moon, B. M., Hoffman, R. R., Eskridge, T. C., & Coffey, J. W. (2011). Skills in applied concept mapping. *Applied Concept Mapping: Capturing, Analyzing, and Organizing Knowledge*, 23-46.
- Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*, 22(3), 149-158.
- Moreno, R., & Mayer, R. E. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology*, 97(1), 117.
- Münchow, H., & Bannert, M. (2019). Feeling good, learning better? Effectivity of an emotional design procedure in multimedia learning. *Educational Psychology*, 39(4), 530-549.
- Münchow, H., Mengelkamp, C., & Bannert, M. (2017). The better you feel the better you learn: Do warm colours and rounded shapes enhance learning outcome in multimedia learning? *Education Research International*, 2017.
<https://doi.org/10.1155/2017/2148139>

- Myers, D. G., & Wilson, J. (2014). Gedächtnis. In D. G. Myers (Hrsg.), *Psychologie* (S.327-365). Springer.
- Navratil, S. D., Köhl, T., & Heidig, S. (2018). Why the cells look like that – the influence of learning with emotional design and elaborative interrogations. *Frontiers in Psychology, 9*, 1653.
- Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413-448.
- Neuner-Anfindsen, S. (2005). *Fremdsprachenlernen und Lernerautonomie: Sprachlernbewusstsein, Lernprozessorganisation und Lernstrategien zum Wortschatzlernen in Deutsch als Fremdsprache: Band 1*. Schneider Verlag Hohengehren.
- Neuroth, J. (2007). *Concept-mapping als Lernstrategie: Eine Interventionsstudie zum Chemielernen aus Texten*. Logos.
- Noetel, M., Griffith, S., Delaney, O., Harris, N. R., Sanders, T., Parker, P., del Pozo Cruz, B., & Lonsdale, C. (2021). Multimedia design for learning: An overview of reviews with meta-meta-analysis. *Review of Educational Research*. doi:10.3102/00346543211052329
- Novak, J., & Cañas, A. (2006). The origins of the concept mapping tool and the continuing evolution of the tool. *Information Visualization, 5*(3), 175-184. doi:10.1057/palgrave.ivs.9500126
- Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27*(10), 937-949.
- Novak, J. D. (1990). Concept maps and Vee diagrams: Two metacognitive tools to facilitate meaningful learning. *Instructional Science, 19*(1), 29-52.
- Novak, J. D., & Cañas, A. J. (2004). Building on new constructivist ideas and cmaptools to create a new model for education [Konferenzbeitrag]. First International Conference on Concept Mapping, Pamplona, Spain.
- Novak, J. D., & Cañas, A. J. (2006). The theory underlying concept maps and how to construct them. *Florida Institute for Human and Machine Cognition, 1*(1), 1-31. <http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf>
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press. doi:10.1017/CBO9781139173469
- Nückles, M., Gurlitt, J., Pabst, T., & Renkl, A. (2004). *Mind Maps & Concept Maps: Visualisieren, Organisieren, Kommunizieren*. Deutscher Taschenbuch Verlag.
- Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion, 6*(2), 257.
- O'Donnell, A. M., Dansereau, D. F., & Hall, R. H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educational Psychology Review, 14*(1), 71-86.
- Oberfoell, A., & Correia, A. (2016). Understanding the role of the modality principle in multimedia learning environments. *Journal of Computer Assisted Learning, 32*(6), 607-617.

- OECD. (2016). *PISA 2015 Assessment and analytical framework: Science, reading, mathematics and financial literacy*. OECD publishing.
- OECD. (2019). *PISA 2018 results (Volume I): What students know and can do*. OECD Publishing. <https://doi.org/doi:https://doi.org/10.1787/5f07c754-en>
- Oser, F., & Spychiger, M. (2005). *Lernen ist schmerzhaft: Zur Theorie des negativen Wissens und zur Praxis der Fehlerkultur*. Beltz.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Paivio, A. (2014). *Mind and its evolution: A dual coding theoretical approach*. Psychology Press.
- Paivio, A., & Clark, J. M. (2006). *Dual coding theory and education. Pathways to Literacy Achievement for High Poverty Children at the University of Michigan School of Education*. Citeseer.
- Palmer, S. E., & Schloss, K. B. (2010). *An ecological valence theory of human color preference* [Konferenzbeitrag]. *The National Academy of Sciences*, 107(19), 8877-8882.
- Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). Emotional design and positive emotions in multimedia learning: An eyetracking study on the use of anthropomorphisms. *Computers & Education*, 86, 30-42.
- Park, B., Moreno, R., Seufert, T., & Brünken, R. (2011). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior*, 27(1), 5-10.
- Pekrun, R. (2005). Progress and open problems in educational emotion research. *Learning and Instruction*, 15(5), 497-506.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315-341.
- Pekrun, R. (2014). Emotions and learning. *Educational Practices Series*, 24(1), 1-31.
- Pekrun, R. (2018). Emotion, Lernen und Leistung. In R. Pekrun (Hrsg.), *Bildung und Emotion* (S. 215-231). Springer.
- Pekrun, R. (2021). Emotions in reading and learning from texts: Progress and open problems. *Discourse Processes*, 1-10.
- Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology*, 98(3), 583.
- Pekrun, R., & Stephens, E. J. (2012). Academic emotions. In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Hrsg.), *APA educational psychology handbook, Vol 2: Individual differences and cultural and contextual factors* (S. 3-31). American Psychological Association.

- Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The Epistemically-Related Emotion Scales. *Cognition and Emotion, 31*(6), 1268-1276.
- Peng, X., Xu, Q., Chen, Y., Zhou, C., Ge, Y., & Li, N. (2021). An eye tracking study: Positive emotional interface design facilitates learning outcomes in multimedia learning? *International Journal of Educational Technology in Higher Education, 18*(1), 1-18.
- Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research, 31*, 459-470.
- Pintrich, P. R. (2003). Motivation and classroom learning. In W. M. Reynolds & G. E. Miller (Hrsg.), *Handbook of psychology: Educational Psychology, Vol. 7* (S. 103–122). John Wiley & Sons Inc.
- Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction, 29*, 128-140.
- Plass, J. L., & Kalyuga, S. (2019). Four ways of considering emotion in cognitive load theory. *Educational Psychology Review, 1-21*.
- Plass, J. L., & Kaplan, U. (2016). Emotional design in digital media for learning. In S. Y. Tettegah & M. Gartmeier (Hrsg.), *Emotions, technology, design, and learning* (S. 131-161). Elsevier.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Hrsg.), *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion* (S. 3-33). Academic.
- Pressley, M., Borkowski, J. G., & Schneider, W. (1989). Good information processing: What it is and how education can promote it. *International Journal of Educational Research, 13*(8), 857-867.
- Quinn, H. J., Mintzes, J. J., & Laws, R. A. (2003). Successive concept mapping: Assessing understanding in college science classes. *Journal of College Science Teaching, 33*(3), 12-17.
- Reiss, K., Weis, M., Klieme, E., & Köller, O. (2019). *PISA 2018: Grundbildung im internationalen Vergleich*. Waxmann Verlag.
- Renkl, A., & Nückles, M. (2006). Lernstrategien der externen Visualisierung. *Handbuch Lernstrategien, 135-147*.
- Rey, G. D. (2009). *E-Learning: Theorien, Gestaltungsempfehlungen und Forschung*. Huber Bern.
- Royer, J. M. (1979). Theories of the transfer of learning. *Educational Psychologist, 14*(1), 53-69.
- Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching, 38*(2), 260-278.

- Ruiz-Primo, M. A., & Shavelson, R. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569-600.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161.
- Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, 76(5), 805.
- Schaal, S., Bogner, F. X., & Girwidz, R. (2010). Concept mapping assessment of media assisted learning in interdisciplinary science education. *Research in Science Education*, 40(3), 339-352.
- Scherer, K. (2002). Emotion. In W. Stroebe, K. Jonas & M. Hewstone (Hrsg.), *Sozialpsychologie* (S. 165-213). Springer. https://doi.org/10.1007/978-3-662-08008-5_6
- Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44(4), 695-729.
- Schirp, H. (2007). Zur Entwicklung demokratischer Kompetenzen. Ein pädagogisch-schulpraktisches Konzept. In D. Lange & G. Himmelmann (Hrsg.), *Demokratiebewusstsein* (S. 164-178). Springer.
- Schmid, C., Zoelch, C., & Roebers, C. M. (2008). Das Arbeitsgedächtnis von 4-bis 5-jährigen Kindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40(1), 2-12.
- Schmid, R. F., & Telaro, G. (1990). Concept mapping as an instructional strategy for high school biology. *The Journal of Educational Research*, 84(2), 78-85. <https://doi.org/10.1080/00220671.1990.10885996>
- Schmidt-Atzert, L. (1996). *Lehrbuch der Emotionspsychologie*. Kohlhammer.
- Schneider, S., Häßler, A., Habermeyer, T., Beege, M., & Rey, G. D. (2019). The more human, the higher the performance? Examining the effects of anthropomorphism on learning with media. *Journal of Educational Psychology*, 111(1), 57.
- Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, 14(1), 101-120.
- Schnotz, W. (2005). An integrated model of text and picture comprehension. *The Cambridge Handbook of Multimedia Learning*, 49, 69.
- Schnotz, W., & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychology Review*, 19(4), 469-508.
- Schwarz, N. (1987). Stimmung als Information: Zum Einfluß von Stimmungen und Emotionen auf evaluative Urteile. *ZUMA-Arbeitsbericht*, 1987/12. <https://nbnresolving.org/urn:nbn:de:0168-ssoar-66496>
- Shangguan, C., Gong, S., Guo, Y., Wang, X., & Lu, J. (2020). The effects of emotional design on middle school students' multimedia learning: The role of learners' prior knowledge. *Educational Psychology*, 1-18.

- Shavelson, R. J., Ruiz-Primo, M. A., & Wiley, E. W. (2005). Windows into the mind. *Higher Education*, 49(4), 413-430.
- Shuman, V., & Scherer, K. R. (2014). Concepts and structures of emotions. In R. Pekrun & L. Linnenbrink-Garcia (Hrsg.), *International handbook of emotions in education* (S. 23-45). Routledge.
- Sommer, C., & Lücken, M. (2010). System competence – Are elementary students able to deal with a biological system. *Nordic Studies in Science Education*, 2(6), 125-143.
- Sreenivas, S., Boehm, S., & Linden, D. E. J. (2012). Emotional faces and the default mode network. *Neuroscience Letters*, 506(2), 229-234.
- Stark, L., Park, B., & Brünken, R. (2018). Emotionen beim Lernen mit Multimedia. In S. Ladel, J. Knopf & A. Weinberger (Hrsg.), *Digitalisierung und Bildung* (S. 141-158). Springer.
- Stárková, T., Lukavský, J., Javora, O., & Brom, C. (2019). Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning? *Journal of Computer Assisted Learning*, 35(4), 555-568.
- Steiner, G. (2006). Lernen und Wissenserwerb. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 137-202). Beltz Verlag.
- Steuer, L. (2014). *Gender und Diversity in MINT-Fächern: Eine Analyse der Ursachen des Diversity-Mangels*. Springer-Verlag.
- Siftung Jugend und Bildung in Zusammenarbeit mit dem Bundesamt für Bevölkerungsschutz und Katastrophenhilfe (BBK) (n.d.). *Max und Flocke im Helferland. Gesundheit! – Baustein zum Thema Grippe und Viren*. http://www.max-und-flocke-helferland.de/SharedDocs/Downloads/DE/Worksheets/Gesundheit.pdf;jsessionid=C6BBCD7683505B04895CBD261AF8BC16.2_cid330?_blob=publicationFile&v=4
- Stracke, I. (2004). *Einsatz computerbasierter Concept Maps zur Wissensdiagnose in der Chemie. Empirische Untersuchungen am Beispiel des Chemischen Gleichgewichts*. Waxmann.
- Strobach, T. (2020). *Kognitive Psychologie*. Kohlhammer Verlag.
- Sumfleth, E., Neuroth, J., & Leutner, D. (2010). Concept Mapping – eine Lernstrategie muss man lernen. *Chemkon*, 17(2), 66-70.
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. Mayer (Hrsg.), *Cambridge handbook of multimedia learning* (S. 19-30). Cambridge University Press. doi:10.1017/CBO9780511816819.003
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296.
- Tergan, S.-O. (1989). Psychologische Grundlagen der Erfassung individueller Wissensrepräsentationen. Teil I: Grundlagen der Wissensmodellierung. *Sprache und Kognition*, 8(3), 152-165.

- Tien, L.-C., Chiou, C.-C., & Lee, Y.-S. (2018). Emotional design in multimedia learning: Effects of multidimensional concept maps and animation on affect and learning. *EURASIA Journal of Mathematics, Science and Technology Education*, 14, 12.
- Tobinski, D. A. (2017). *Kognitive Psychologie: Problemlösen, Komplexität und Gedächtnis*. Springer-Verlag.
- Um, E., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional design in multimedia learning. *Journal of Educational Psychology*, 104(2), 485.
- Uzun, A. M., & Yildirim, Z. (2018). Exploring the effect of using different levels of emotional design features in multimedia science learning. *Computers & Education*, 119, 112-128.
- van Gog, T., Paas, F., & Sweller, J. (2010). Cognitive load theory: Advances in research on worked examples, animations, and cognitive load measurement. *Educational Psychology Review*, 22(4), 375-378.
- van Marlen, T., van Wermeskerken, M., Jarodzka, H., & van Gog, T. (2018). Effectiveness of eye movement modeling examples in problem solving: The role of verbal ambiguity and prior knowledge. *Learning and Instruction*, 58, 274-283.
- van Merriënboer, J. J., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 20(3), 343-352.
- van Merriënboer, J. J., Schuurman, J. G., De Croock, M., & Paas, F. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction*, 12(1), 11-37.
- van Velzen, J. (2015). *Metacognitive learning*. Springer.
- Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy*, 6(4), 473-482.
- von der Linden, N. (2008). *Metakognitive Überwachungsprozesse im Entwicklungsverlauf*. Verlag Dr. Kovac.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063.
- Weidenmann, B. (2002). Multicodierung und Multimodalität im Lernprozess. In L. J. Issing & P. Klimsa (Hrsg.), *Information und Lernen mit Multimedia und Internet* (S. 45-62). Beltz PVU.
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Hrsg.), *handbook of research on teaching* (Vol. 3, pp. 315-327). Macmillan.
- Wendt, H., Bos, W., Selzer, C., Köller, O., Schwippert, K., & Kasper, D. (2016). *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Waxmann.

- Wendt, L. (2021). Metakognition. In G. Kaiser (Hrsg.), *Reflexionsfähigkeit von Lehrkräften über metakognitive Schülerprozesse beim mathematischen Modellieren* (S. 9-29). Springer.
- Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, *18*(1), 411-418.
- Wild, E., Hofer, M., & Pekrun, R. (2006). Psychologie des Lernens. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 201-267). Beltz Verlag.
- Wild, K.-P., & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *15*(4), 185-200.
- Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist*, *11*(2), 87-95.
- Wong, R. M., & Adesope, O. O. (2021). Meta-analysis of emotional designs in multimedia learning: A replication and extension study. *Educational Psychology Review*, *33*(2), 357-385.
- Wulf, T. (2017). *Kooperation und Kooperation im Videospiel*. Springer.
- Zoelch, C., Berner, V.-D., & Thomas, J. (2019). Gedächtnis und Wissenserwerb. In D. Urhahne, M. Dresel & F. Fischer (Hrsg.), *Psychologie für den Lehrberuf Psychologie für den Lehrberuf* (S. 23-52). Springer.

Danksagung

Für mich war diese Dissertation eine herausfordernde, aber auch sehr gewinnbringende Reise. All denen, die mich beim Zustandekommen dieser Arbeit im Großen wie auch im Kleinen unterstützt haben, möchte ich herzlich danken. Ich lade an dieser Stelle jeden/jede, der/die mich während der Promotionsphase begleitet hat ein, sich angesprochen zu fühlen. Zudem hoffe ich, dass es mir abgesehen von diesen Zeilen auch persönlich möglich sein wird, meinem Dank Ausdruck zu verleihen.

Mein besonderer Dank gilt meinem Betreuer Herrn **Prof. Dr. Jörg Großschedl**, der diese Dissertation wie niemand sonst unterstützt hat. Das konstruktive Feedback, die aufwändigen Rückmeldungen zu sämtlichen Schriftstücken, die enge Betreuung sowie die unerschütterliche Zuversicht (besonders während der Publikationsprozesse) haben diese Dissertation erst möglich gemacht. Auch das hohe Maß an Verständnis und Weitsicht gepaart mit dem Bestreben, seinen Doktorand:innen ein bestmögliches Arbeitsumfeld zu bieten, spiegeln das beispiellose Engagement wieder.

Bei Frau **Prof. Dr. Kirsten Schlüter** möchte ich mich an dieser Stelle für die jahrelange Unterstützung und die Wertschätzung bedanken.

Herrn **Prof. Dr. Werner Rieß** danke ich für den netten Austausch und für seine Bereitschaft, die Zweitbetreuung und -begutachtung im Rahmen meines Promotionsvorhabens zu übernehmen.

Bei Herrn **Prof. Dr. Frank Schäbitz** möchte ich mich für die Übernahme des Vorsitzes in meinem Disputationsverfahren bedanken.

In diesem Zusammenhang gilt mein Dank auch Herrn **Dr. Fabian Seredszus** für die Übernahme des Beisitzes im Rahmen der Disputation.

Frau **Dr. Gabriele Schwager-Büschges** und den **Mitarbeiter:innen der Graduiertenschule für Lehrer:innenbildung der Universität zu Köln** danke ich für die finanzielle und ideelle Förderung sowie ihr unablässiges Engagement. Die Graduiertenschule war stets darauf bedacht, allen Doktorand:innen bedarfsorientierte Hilfestellungen und Unterstützung auf ihrem Weg zur Verfügung zu stellen.

Ein herzliches Dankeschön gilt auch allen **Teilnehmer:innen des Doktoranden-Kolloquiums**, die zahlreichen Anregungen haben mir immer wieder neue Blickwinkel eröffnet und den Fortgang dieser Arbeit gefördert. Bei Frau **Prof. Dr. Kirsten Schlüter** und Herrn **Prof. Dr. Jörg Großschedl** möchte ich mich für die Gründung und Organisation des Kolloquiums sowie das beständige Interesse an den vorgestellten Projekten bedanken. Dieser Einsatz ist angesichts ihrer zahlreichen anderweitigen Verpflichtungen nicht

selbstverständlich. Das Kolloquium hat einen großen Stellenwert in der Vernetzung der Doktorand:innen untereinander und in meinem eigenen Wachstumsprozess eingenommen.

Mein Dank geht auch an alle (wissenschaftlichen) **Mitarbeiter:innen** und **Hilfskräfte** für die Unterstützung während der gesamten Promotionszeit und insbesondere der intensiven Zeit der Erhebungen. Danke für Rat und Tat, für helfende Hände, leckeres Essen und offene Ohren. Danke auch für schnelle und unkomplizierte Lösungen und die gute Zusammenarbeit.

Insbesondere möchte ich meinen (ehemaligen) **Kolleg:innen/Mit-Doktorand:innen** für die freundschaftliche Arbeitsatmosphäre, die unterhaltsamen Mittagspausen und den konstruktiven Austausch danken. So viele schöne Momente konnte ich mit Euch sammeln. Ihr habt dafür gesorgt, dass mir die Zeit sehr positiv in Erinnerung bleiben wird.

Ohne die **Lehrer:innen** und **Schüler:innen**, die an meinen Studien teilgenommen und somit die Grundlage meiner Forschungsarbeit geschaffen haben, wäre die vorliegende Arbeit nicht möglich gewesen. Dafür möchte ich mich an dieser Stelle ganz herzlich bedanken.

Diesbezüglich gilt mein Dank auch den **Studierenden des Moduls „Didaktische Forschungsprojekte“** und meinen **Abschlussarbeitskandidat:innen**, welche bei der Weiterentwicklung von Erhebungsmaterialien und/oder der Erhebung von Daten unterstützt haben.

Natürlich geht ein maßgeblicher Teil meines Danks an meine **Familie** und meine **Freunde**, die jeden meiner Schritte begleiten und immer unterstützend für mich da sind. Das bedeutet mir viel.

Danke.

Eidesstattliche Versicherung zu Dissertation

gemäß der Promotionsordnung vom 12. März 2020

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

Bonn, den 04.04.2022

Sina Lenski

Erklärung zur Verfügbarkeit von Primärdaten

Die Primärdaten, welche im Rahmen dieser Arbeit erhoben und ausgewertet wurden, wurden am Institut für Biologiedidaktik der Universität zu Köln archiviert. Die Einsicht des Erhebungsmaterials und der Zugang zu digitalen Dateien kann bei Herrn **Prof. Dr. Jörg Grosschedl** unter folgende E-Mail-Adresse angefragt werden:

j.grossc@uni-koeln.de

Überblick über relevante Teilpublikationen

Die vorgelegte Dissertation umfasst eine, in einer internationalen wissenschaftlichen Fachzeitschrift erschienene Publikation (**Publikation III**), einen Beitrag für ein begutachtetes Sammelwerk, welcher sich im Druck befindet (**Publikation II**) sowie zwei bei internationalen wissenschaftlichen Fachzeitschriften eingereichte Publikationen (**Publikationen I und IV**).

Publikation I*

Lenski, S., Elsner, S., & Großschedl, J. (eingereicht). Comparing construction and study of concept maps – An intervention study on cognitive, metacognitive and emotional effects of training & learning. *Frontiers in Education*.

Publikation II*†

Lenski, S., & Großschedl, J. (im Druck). Biologie lernen mit Concept Maps: Lässt sich die Expertise im Umgang mit Concept Maps von den Augen ablesen? In P. Klein, M. Schindler, N. Graulich & J. Kuhn (Hrsg.), *Eye Tracking als Methode in der Mathematik- und Naturwissenschaftsdidaktik: Forschung und Praxis*. Springer.

Publikation III*†

Lenski, S., & Großschedl, J. (2022). Emotional design in concept maps – No support but also no burdens. *Frontiers in Education*. 7, 807627. doi: 10.3389/educ.2022.807627

Publikation IV

Lenski, S., and Großschedl J. (eingereicht). Emotional design pictures –pleasant but too weak to evoke arousal and attract attention? *BMC Psychology*.

*Ermöglicht wurde die Durchführung dieser Studien durch eine Finanzierung über das Projekt „Biologie lernen durch Concept Mapping: Bedeutung eines Lernstrategietrainings für kognitive Belastung, kognitive Prozesse und Lernleistung“, welches von der Deutschen Forschungsgemeinschaft (DFG) gefördert wurde (Projektnummer GR 4763/2).

†Für die Aufnahme dieser Beiträge wurde das schriftliche Einverständnis der Herausgeber:innen eingeholt.

Bonn, den 04.04.2022

Sina Lenski