

Welche Strategien verwenden Schülerinnen und Schüler bei der Betrachtung von Geographieschulbuchseiten mit inhaltlich aufeinander bezogenen Text- und Bildkomponenten?

Eine explorative Eye-Tracking-Studie in der 5. und 10. Jahrgangsstufe

Which Strategies Do Students Use Looking at Geography Schoolbook Pages Including Corresponding Text and Image Components? An Eye Tracking Study with 5th and 10th Grade Students

¿Qué estrategias utilizan los estudiantes al mirar las páginas del libro escolar de Geografía, incluidos los componentes de texto e imagen correspondientes? Un estudio de seguimiento ocular con estudiantes de quinto y décimo grado

Dominik Conrad , Gabriele Obermaier

Zusammenfassung Gemäß der kognitiven Theorie multimedialen Lernens (MAYER, 2001) unterstützt eine simultane Verarbeitung bildhafter und verbaler Informationen den Aufbau eines kohärenten mentalen Modells. Mittels Eye-Tracking werden Strategien von Schülerinnen und Schülern der 5. und 10. Jahrgangsstufe ($n = 76$) bei der Betrachtung geographischer Schulbuchseiten in Bezug auf korrespondierende Text- und Bildkomponenten untersucht und der Frage nachgegangen, ob diese sich durch textbasiertes Signaling verändern. Bei Betrachtungen von Seiten mit Signaling werden signifikant mehr Abbildungen in direkter zeitlicher Abfolge zum korrespondierenden Text betrachtet. Im Vergleich beider Jahrgangsstufen zeigt sich, dass dies bei Betrachtungen der Schülerinnen und Schüler der 5. Jahrgangsstufe signifikant häufiger geschieht.

Schlüsselwörter Schulbuchforschung, Eye-Tracking, Strategien, kognitive Theorie des Multimedialernens (CTML), Signaling-Prinzip

Abstract According to the Cognitive Theory of Multimedia Learning (MAYER, 2001), simultaneous processing of visual and verbal information helps to build a coherent mental model. Eye tracking was used to investigate the strategies of 5th and 10th grade students when looking at geography textbook pages with corresponding text and image components and to analyze if these strategies change through text-based signaling. Participants of both grades looked at significantly more illustrations and corresponding texts located in temporal proximity on pages including signaling. The linking of corresponding text and illustrations is significantly more frequent in grade 5.

Keywords textbook research, eye tracking, strategies, Cognitive Theory of Multimedia Learning, signaling principle

Resumen Según la Teoría Cognitiva del Aprendizaje Multimedial (MAYER, 2001), el procesamiento simultáneo de información visual y verbal facilita la construcción de un modelo mental coherente. El presente trabajo utilizó el seguimiento ocular para investigar las estrategias de los estudiantes de quinto y décimo grado cuando miraban páginas de libros de texto de Geografía conteniendo componentes de texto e imagen. También se analizó si estas estrategias cambian a través de señales basadas en texto. Los participantes de ambos grados observaron muchas más ilustraciones y textos correspondientes ubicados en proximidad temporal en páginas que incluían señalización. La vinculación de textos e ilustraciones correspondientes es mucho más frecuente en el quinto grado.

Palabras clave investigación de libros de texto, seguimiento ocular, estrategias, Teoría Cognitiva del Aprendizaje Multimedial, principio de señalización

1. Einleitung

Das Schulbuch nimmt auch heute noch eine bedeutende Position unter den im Geographieunterricht zum Einsatz kommenden Medien ein (BAGOLY-SIMÓ, 2014; LATHAN, 2021) und wird von Schülerinnen und Schülern als ein sehr wichtiges Medium zum Lernen betrachtet (LATHAN, 2021). Als Verbundmedium beinhaltet es unterschiedliche Medien zu einem Themengebiet, neben Texten sind dies Abbildungen wie z.B. Photographien, Zeichnungen oder Karten, das heißt, Informationen werden verbal und bildhaft angeboten. Auf Basis einer Analyse von Geographieschulbuchseiten weist BEHNKE (2016a) darauf hin, dass Kenntnisse zum Lernen mit Medien wie beispielsweise die kognitive Theorie des multimedialen Lernens (CTML) (MAYER, 2001) bislang in unzureichender Weise bei der Gestaltung von Geographieschulbuchseiten berücksichtigt werden. Als eine Möglichkeit der

Optimierung führt sie „textuelle Verlinkungen zwischen zusammengehörigen Materialien“ (BEHNKE, 2016a, S. 28) an. Eine Möglichkeit einer solche Verlinkung ist es, verbale Hinweise im Text auf Abbildungen einzufügen, wodurch Lernende befähigt werden sollen, Informationen aus Text und Abbildung in unmittelbarer zeitlicher Nähe aufzunehmen und dadurch beim Aufbau eines mentalen Modells unterstützt werden. Allerdings fehlt es wie bei vielen Gestaltungsempfehlungen der CTML bislang an Studien mit authentischen Lernmaterialien und Schülerinnen und Schülern als Probanden (vgl. BEHNKE, 2018; CHENG ET AL., 2015; JAROZDKA ET AL., 2017). Dies gilt auch für Forschungsarbeiten, welche die Wirksamkeit eines solchen textbasierten Signalings durch verbale Hinweise im Text auf korrespondierende Abbildungen untersuchen.

2. Theorie

Der Lernprozess kann nach KLAUER und LEUTNER (2007) in drei Phasen (Vorbereitung - Durchführung - Abschluss) gegliedert werden. In dieser Studie wird das Lesen einer Schulbuchseite als eine in der Phase der Durchführung zum Einsatz kommende Strategie des Informationserwerbs betrachtet. KLAUER (2011) versteht unter einer Strategie „ein[en] Handlungsplan [...], eine Handlungsfolge oder Handlungskette, die bewusst eingesetzt werden kann oder auch nur faktisch, ohne den Handlungsplan als solchen zu kennen oder gar benennen zu können“ (S. 99). Die Anwendung einer Strategie kann also auch unbewusst erfolgen.

Gemäß MAYERS (2001) Definition des multimedialen Lernens „[...] as the presentation of material using both words and pictures“ (S. 2) kann der Erwerb von Wissen mittels eines Schulbuches dem multimedialen Lernen zugerechnet werden. Unter Wörtern werden hierbei gedruckte und gesprochene Wörter subsumiert. Unter Bildern versteht Mayer alle Arten bildlicher Darstellungen, z. B. Fotos, Karten, Illustrationen, aber auch Animationen und Filme (MAYER, 2003, 2014; MAYER & MORENO, 2003). Die von Mayer entwickelte CTML geht davon aus, dass bei der Aufnahme von Informationen beim multimedialen Lernen drei Aktivitäten zentral sind: (1) Selektion von relevantem Material, (2) Organisation des ausgewählten Materials und (3) Integration mit dem bereits vorhandenen Wissen (MAYER, 2014; MAYER & MORENO, 2003; MORENO &

MAYER, 1998). Über Augen und Ohren eingehende Informationen gelangen in den sensorischen Speicher. Aufgrund der geringen Speicherkapazität des Arbeitsgedächtnisses erfolgt dort eine Auswahl der aus dem sensorischen Speicher eingehenden Informationen.

Im nächsten Schritt, der Organisation des ausgewählten Materials, werden die selektierten Töne beziehungsweise Bilder in sinnhafte verbale Modelle beziehungsweise einfache bildhafte Modelle überführt. Bei der Transformation in kohärente Wissensstrukturen in einem integrierten Modell werden verbales Modell und bildhaftes Modell aufeinander bezogen. Hierbei spielt das im Langzeitgedächtnis gespeicherte Vorwissen eine wichtige Rolle (MAYER, 2014).

Die CTML basiert auf der Theorie der dualen Kodierung (CLARK & PAIVIO, 1991; PAIVIO, 1986), der Theorie des Arbeitsgedächtnis (BADDELEY, 1986) sowie der Cognitive Load Theory (PAAS & SWELLER, 2014; SWELLER, 1999). Es werden drei Arten von kognitiver Belastung (*Cognitive Load* [CL]) beim Lernen unterschieden. Der *intrinsic CL* bezieht sich auf die kognitive Belastung durch den Schwierigkeitsgrad des Lerngegenstandes (NIEDING ET AL., 2015). Dieser ist gebunden an die Lernaufgabe, er kann durch die Gestaltung des Lernmaterials nicht verändert werden, allerdings kann ein großes Vorwissen des Lernenden den *intrinsic CL* mindern (PASS & SWELLER, 2014; SWELLER, 2010). Der *extra-*

neous CL wird durch unsachgemäße Gestaltung des Lernmaterials erzeugt, die zu einer Überlastung des Arbeitsgedächtnisses führen kann (PASS & SWELLER, 2014; SWELLER, 2010). Der *germane CL* errechnet sich aus dem *intrinsic CL* minus dem *extraneous CL* und stellt somit die effektive kognitive Belastung des Arbeitsgedächtnisses dar.

Viele Gestaltungsempfehlungen zum Multimedialernen zielen darauf ab, den *extraneous CL* zu reduzieren. Eine Verringerung des *extraneous CL* bewirkt, dass Kapazitäten des Arbeitsgedächtnisses für den *intrinsic CL* frei werden. Hieraus ergibt sich eine Erhöhung des *germane CL* (PASS & SWELLER, 2014). Für die vorliegende Studie sind folgende auf Basis der CTML entwickelte Gestaltungsempfehlungen grundlegend: das Multimedialprinzip, das Prinzip der räumlichen Nähe sowie insbesondere das Signalingprinzip.

Das Multimedialprinzip besagt, dass Lernen mithilfe von Wörtern und Bildern einfacher erfolgen kann als durch Wörter alleine. Durch die Darbietung in Wort- und Bildform erhalten die Lernenden die Möglichkeit, ein verbales sowie ein bildhaftes Modell zu konstruieren, diese aufeinander zu beziehen und unter Einbezug des Vorwissens in ein kohärentes Modell zu integrieren (MAYER, 2001).

Das Prinzip der räumlichen Nähe beinhaltet, dass Lernen vereinfacht wird, wenn Wörter und Bilder in räumlicher Nähe präsentiert werden. Das Arbeitsgedächtnis wird nicht durch das Absuchen einer Seite oder des Bildschirms zusätzlich belastet, der *extraneous CL* wird reduziert. Verbale und bildhafte Informationen können so leichter zusammen im Arbeitsgedächtnis gehalten und aufeinander bezogen werden. Demnach ist die gleichzeitige Bearbeitung von Bild- und Textinformationen effektiver als die aufeinanderfolgende Bearbeitung von Text- und Bildinformationen (MAYER, 2001).

3. Forschungsstand

Zwar ist die Wirksamkeit der aus der CTML abgeleiteten Gestaltungsempfehlungen zum Multimedialernen empirisch belegt (MAYER, 2008), aber der Großteil der Studien wurde bislang mit Studentinnen und Studenten durchgeführt und es besteht zudem weiterer Forschungsbedarf mit authentischen Lernmaterialien, die von der jeweiligen Zielgruppe getestet werden (BEHNKE, 2018, 2021; CHENG ET AL., 2013; JARODZKA ET AL., 2017).

JARODZKA ET AL. (2017) weisen zudem darauf hin, dass insbesondere der Einsatz von Eye-Tracking-Verfahren dazu beitragen kann, die der CTML zugrunde liegenden Annahmen zu spezifizieren. Betrachtet man die bislang zum Multimedialernen durchgeführten Eye-Tracking-Studien im

Das Signalingprinzip (MAUTONE & MAYER, 2001; MAYER, 2008; MAYER & FIORELLA, 2014; VAN GOG, 2014) postuliert, dass Lernende von Multimediaangeboten stärker profitieren, wenn die Aufmerksamkeit durch Hinweise/Signale gesteuert wird.

Bei den angesprochenen Signalen handelt es sich entweder um bildbasierte (z.B. Hervorhebungen durch Veränderung von Farben) oder textbasierte Signale (z.B. im Text befindliche Hinweise zu korrespondierenden Bildern) oder eine Kombination aus beiden (z.B. gleiche farbliche Hervorhebung zusammengehöriger Text- und Bildelemente). Bezogen auf die CTML soll Signaling den ersten Schritt multimedialen Lernens, die Informationsauswahl, vereinfachen. Zudem soll durch Signaling die erfolgreiche Integration verbaler und bildlicher Informationen in ein kohärentes mentales Modell unterstützt werden (SCHNEIDER ET AL., 2018; VAN GOG, 2014).

Kritik hat Mayers Theorie aufgrund fehlenden Aufgreifens emotionaler und motivationaler Faktoren erfahren (NIEDING ET AL., 2015; SCHNEIDER ET AL., 2018). WESTELINCK ET AL. (2005) postulieren eine Einschränkung des Geltungsbereiches des Multimedialprinzips und des Prinzips der räumlichen Nähe auf intuitiv verständliche bildhafte Darstellungen. KALYUGA (2014) verweist darauf, dass viele Gestaltungsempfehlungen zum multimedialen Lernen bei Lernenden mit hohem Vorwissen zu keinem höheren Lernerfolg beitragen, den Lernerfolg sogar mindern können. Dieses als Expertise-Reversal-Effekt bezeichnete Phänomen führt Kalyuga darauf zurück, dass für Lernende mit hohem Vorwissen aufgrund nicht notwendiger Hilfestellungen das Arbeitsgedächtnis durch eine Erhöhung des *extraneous CL* unnötig belastet werde.

Speziellen, wird deutlich, dass die Probanden auch hier größtenteils Studentinnen und Studenten gewesen sind (MASON ET AL., 2013).

In der Geographiedidaktik wurden bislang drei Eye-Tracking-Studien zur visuellen Wahrnehmung von Geographieschulbuchseiten durchgeführt (BEHNKE, 2016a, 2016b, 2022, 2023). BEHNKE (2022) untersuchte mögliche Einflussfaktoren auf die visuelle Aufmerksamkeit beim Lernen mit Geographieschulbüchern. In einer weiteren Studie widmete sich BEHNKE (2023) den visuellen Informationsselektionsprozessen der Probanden beim Bearbeiten von Aufgaben auf Fotos, Karten und Graphiken in Geographieschulbüchern. In der Studie von BEHNKE (2016a, 2016b) wurde der Einfluss des De-

signs von Geographieschulbuchseiten auf die visuelle Aufmerksamkeit untersucht. Bei allen betrachteten Schulbuchseiten widmeten die Probanden, zwölf Studentinnen und Studenten und acht Schülerinnen und Schüler, den Textelementen mit weitem Abstand die größte visuelle Aufmerksamkeit. HANNUS und HYÖNÄ (1999) kamen in einer Studie mit Schülerinnen und Schülern der 4. Jahrgangsstufe ($n = 24$) mit Lernmaterialien im Fach Biologie zu einem ähnlichen Ergebnis. Sie stellten zudem fest, dass es kaum zu Blickbewegungen zwischen Text und korrespondierenden Abbildungen kommt. Dieser Befund deckt sich mit den Ergebnissen der Eye-Tracking-Studie von HOLSANOVA ET AL. (2008) bei erwachsenen Probanden beim Zeitunglesen.

Hinsichtlich der Wirksamkeit von Signaling konstatieren drei Metaanalysen (MAYER & FIORELLA, 2014; RICHTER ET AL., 2016; VAN GOG, 2014), dass dieses zumeist den Lernerfolg erhöht. Das Vorwissen ist nach RICHTER ET AL. (2016) hierbei die wichtigste moderierende Einflussgröße. So profitieren Probanden mit geringem oder mittlerem Vorwissen stärker von Signaling als Probanden mit großem Vorwissen. Als eine mögliche Erklärung hierfür wird von RICHTER ET AL. (2016) der Expertise-Reversal-Effekt betrachtet. Die Ursache für diesen wird (siehe Theorieteil) in einem erhöhten *extraneous CL* durch für Probanden mit großem Vorwissen nicht notwendige Signale gesehen. Eine andere mögliche Erklärung dafür, dass Lernende mit hohem Vorwissen nicht von Signaling profitieren, könnte nach Meinung von RICHTER ET AL. (2016) aber auch die Informations-Reduktions-Hypothese (HAIDER & FRENCH, 1999) bieten. Diese besagt, dass mit steigender Expertise in einem Gebiet die Fähigkeit wächst, unnötige Informationen zu ignorieren. Dies würde bedeuten, dass Probanden mit hohem Vorwissen die Signale gar nicht beachten. Hier sehen RICHTER ET AL. (2016) weiteren Forschungsbedarf. Auch VAN GOG (2014) bilanziert, dass erheblicher Forschungsbedarf bezüglich der Frage besteht, unter welchen Bedingungen Signaling wirkungsvoll ist. Einflussfaktoren können hierbei die unterschiedlichen Arten des Signalings, aber auch Unterschiede in den Probandengruppen darstellen. Sie sieht in Eye-Tracking-Studien eine gute Option, die Wirkungsweise des Signalingprinzips besser zu verstehen (VAN GOG, 2014).

FOLKER ET AL. (2005), OZCELIK ET AL. (2009), DÉSION ET AL. (2018) sowie SCHEITER und EITEL (2015) untersuchten mittels Eye-Tracking die Auf-

merksamkeitssteuerung durch gleiche Farbgebung korrespondierender Elemente in Text und Abbildung bei Studierenden. Bei SCHEITER und EITEL (2015) kamen zudem als Signale noch kursiv gedruckte Ausdrücke im Text, die zudem auch als kursiv gedruckte Bildbeschriftung auftauchen, sowie verbale Hinweise im Text auf Teile einer Abbildung zum Einsatz. Zur Aufmerksamkeitslenkung durch Signaling bei Schülerinnen und Schülern unter den Bedingungen kontinuierlich dargebotener Informationen mittels Text und Abbildungen (keine Animationen) und selbstgesteuerter Betrachtung liegen unseres Wissens bislang zwei Studien vor. RICHTER und SCHEITER (2019) untersuchten die visuelle Aufmerksamkeit von Schülerinnen und Schülern im Alter von 13-17 Jahren mit und ohne Signaling in Abhängigkeit vom Vorwissen der Probanden. Signaling erfolgte größtenteils durch gleiche Farbgebung korrespondierender Text- und Abbildungselemente und Abbildungsbeschriftungen sowie einmal durch Hinweise im Text auf die Abbildung. Eine getrennte Auswertung der Signalingarten erfolgte nicht. Sie stellen fest, dass die visuelle Aufmerksamkeit der Probanden mit hohem Vorwissen durch Signaling nicht beeinflusst wird. Ihre Ergebnisse stützen die Annahme, dass Schülerinnen und Schüler mit hohem Vorwissen gemäß der Informations-Reduktions-Hypothese erkennen, ob es für sie notwendig ist, Signale überhaupt zu beachten (RICHTER & SCHEITER, 2019).

Bei einer Eye-Tracking-Studie mit Schülerinnen und Schülern der 6. Jahrgangsstufe ($n = 56$) untersuchten MASON ET AL. (2013) sowohl den Lernzuwachs als auch das Blickverhalten unter drei Bedingungen. Eine Probandengruppe erhielt nur einen Text, eine andere einen Text mit drei Abbildungen, eine andere Gruppe einen Text mit drei Abbildungen, die im Gegensatz zur anderen Gruppe Bildbeschriftungen enthielten. Direkte Verweise vom Text auf die Abbildungen wurden keine gezogen. Sie stellten fest, dass Signaling mittels Bildbeschriftungen bei Schülerinnen und Schülern nicht nur den Lernerfolg erhöht, sondern auch dazu führt, dass Schülerinnen und Schüler, die den Text gelesen haben und nach der Betrachtung der Abbildung ihren Blick wieder auf den Text richten, längere Fixationen relevanter Textsegmente vornehmen. Die Aufmerksamkeitslenkung ausschließlich durch verbale Hinweise im Text auf Abbildungen oder Teile der Abbildungen bei Schülerinnen und Schülern wurde nach unserem Wissen bislang noch nicht systematisch untersucht.

4. Fragestellung

Die vorliegende explorative Untersuchung geht der Frage nach, mit welchen Strategien Schülerin-

nen und Schüler Geographieschulbuchseiten betrachten, einmal ohne Signaling, einmal bei textba-

siertem Signaling durch verbale Hinweise im Text auf Abbildungen oder Teile der Abbildungen. Zwei Fragestellungen sind leitend:

(1) Welche Strategien verwenden Schülerinnen und Schüler bei der Betrachtung von Text-Bild-Komponenten auf Schulbuchseiten ohne Signaling? Gibt es Unterschiede zwischen 5. und 10. Jahrgangsstufe?

(2) Verändert sich die Strategie beim Betrachten von Schulbuchseiten durch das Einfügen von Hinweisen im Text auf die Abbildungen, und wenn ja, wie? Zeigen sich Unterschiede zwischen 5. und 10. Jahrgangsstufe?

Aufbauend auf die Forschungsergebnisse von HANNUS und HYÖNÄ (1999) sowie HOLSANOVA ET AL. (2008) ist zu vermuten, dass es beim Betrachten der Schulbuchseiten ohne Signaling eher selten dazu kommt, dass auch Schülerinnen und Schüler der Sekundarstufe 1 Text und dazugehörige Abbildungen in direkter zeitlicher Abfolge anschauen, und wenn sie es tun, dann nicht systematisch.

Eye-Tracking-Studien (z. B. FOLKER ET AL., 2005) konnten zeigen, dass gleiche farbliche Hervorhebungen in Abbildung und korrespondierendem Text Einfluss auf die Verteilung der visuellen Aufmerksamkeit von Studentinnen und Studenten haben. Bei Studien mit Schülerinnen und Schülern unter Einsatz gleicher Beschriftungen von Texten und Bildern (MASON ET AL., 2013) und hauptsächlich gleicher farblicher Hervorhebung (RICHTER & SCHEITER, 2019) zeigte sich Ähnliches, bei letztgenannter Studie allerdings nur bei Lernenden mit geringem oder mittlerem Vorwissen. Es ist anzunehmen, dass sich die Strategie der Probanden bei der Jahrgangsstufe durch verbale Hinweise im Text auf korrespondierende Abbildungen dahingehend unterscheidet, dass von Schülerinnen und Schülern der 5. Jahrgangsstufe Abbildungen und korrespondierende Textstellen häufiger in direkter zeitlicher Abfolge betrachtet werden. Diese haben vermutlich ein geringeres Vorwissen zu den Themen der Schulbuchseiten als Schülerinnen und Schüler der 10. Jahrgangsstufe. Bei diesen ist aufgrund der Ergebnisse der Studie von RICHTER ET AL. (2019) im Einklang mit der Informations-Reduktions-Hypothese (HAIDER & FRENSCH, 1999) anzunehmen, dass Signaling, wenn überhaupt, einen geringeren Einfluss auf das Blickverhalten hat.

5. Forschungsdesign

5.1 Eye-Tracking

Die Studie wurde mittels Eye-Tracking-Technologie durchgeführt. Bei Eye-Tracking handelt es sich um ein Verfahren zur Registrierung von Blickbewegun-

Hypothesen:

Hypothese 1a: Ohne Signaling kommt es bei mehr als der Hälfte der Seitenbetrachtungen zur Anwendung einer Strategie, bei der Text und korrespondierende Abbildung überhaupt nicht direkt aufeinanderfolgend betrachtet werden.

Hypothese 1b: Ohne Signaling ist die Häufigkeit der Verwendung einer Strategie mit mindestens einer direkt aufeinanderfolgenden Betrachtung korrespondierender Text- und Abbildungselemente unabhängig von der Jahrgangsstufe des Betrachters.

Hypothese 2a: Mit Signaling kommt es bei Betrachtungen in beiden Jahrgangsstufen häufiger zu mindestens einer direkt aufeinanderfolgenden Betrachtung korrespondierender Text- und Abbildungselemente als bei der Betrachtung ohne Signaling.

Hypothese 2b: Mit Signaling und bei einer Seitenbetrachtung mit mindestens einer direkt aufeinanderfolgenden Betrachtung korrespondierender Text- und Abbildungselemente werden pro Seitenbetrachtung in beiden Jahrgangsstufen Abbildungen häufiger in direkter zeitlicher Abfolge zum korrespondierenden Text betrachtet als ohne Signaling.

Hypothese 3a: Unter der Bedingung Signaling kommt es bei Seitenbetrachtungen von Schülerinnen und Schülern der 5. Jahrgangsstufe häufiger zur Nutzung einer Strategie, bei der mindestens einmal eine direkt aufeinanderfolgende Betrachtung korrespondierender Text- und Bildelemente erfolgt, als bei Betrachtungen von Schülerinnen und Schülern der 10. Jahrgangsstufe.

Hypothese 3b: Bei Seitenbetrachtungen von Schülerinnen und Schülern in der 5. Jahrgangsstufe unter Nutzung einer Strategie, bei der es mindestens einmal zur direkt aufeinanderfolgenden Betrachtung korrespondierender Text- und Bildelemente kommt, liegt unter der Bedingung Signaling die Anzahl in direkter zeitlicher Abfolge betrachteter korrespondierender Text-Bild-Komponenten höher als bei Schülerinnen und Schülern der 10. Jahrgangsstufe.

gen im Genauigkeitsbereich von Millisekunden (BLAKE, 2013; GEISE, 2011). Grob betrachtet bestehen Blickbewegungen aus Sakkaden und Fixationen. Sakkaden bezeichnen die Bewegungen des Auges von einer Fixation zur nächsten. Sie dauern

in der Regel zwischen 30-80 ms (HOLMQVIST & ANDERSON, 2017). Während der Dauer einer Sakkade ist die Wahrnehmung stark eingeschränkt (HOLMQVIST & ANDERSON, 2017; RAYNER, 1998). Von einer Fixation spricht man, wenn sich das Auge eines Rezipienten in einer „Phase des relativen Stillstands der Augen“ (BLAKE, 2013, S. 278) befindet. Der Stillstand der Augen wird als relativ charakterisiert, weil das Auge auch bei einer Fixation Mikrobewegungen vollführt (DUCHOWSKI, 2007; RÖTTING & SEIFERT, 1999). Die Dauer einer Fixation beträgt mindestens 100 ms (RÖTTING & SEIFERT, 1999), die durchschnittliche Länge liegt bei 200-400 ms (HOLMQVIST & ANDERSON, 2017). Grundlegend für Eye-Tracking-Studien sind die von JUST und CARPENTER (1980) formulierte *Immediacy Assumption*, der zufolge eine kognitive Verarbeitung unmittelbar nach der Fixation erfolgt, sowie die *Eye-Mind-Assumption*, der zufolge die Dauer der Fixation der Dauer der mentalen Verarbeitung entspricht. Auch wenn die Allgemeingültigkeit beider Annahmen wissenschaftlich kontrovers diskutiert wird (BLAKE, 2013; GEISE, 2011), bilden Fixationen die Voraussetzung zur kognitiven Verarbeitung (HOLMQVIST & ANDERSON, 2017) und es ist davon auszugehen, dass visuelle Aufmerksamkeit und Fixationen bei komplexeren Aufgaben des Informationserwerbs sehr stark miteinander verknüpft sind (RAYNER, 1998).

In dieser Studie wurde ein monitorbasierter Eye-Tracker (Tobii X2-60) mit einer Samplingrate von 60 Hz verwendet. Die Schulbuchseiten wurden auf einem 17-Zoll-Monitor präsentiert. Als Software bei der Darbietung und zur Datenanalyse wurde Tobii Studio 3.2.0 verwendet. Der richtige Abstand zwischen Probanden und Bildschirm sowie die richtige Sitzhöhe wird im Rahmen der Kalibrierung in Tobiistudio 3.2.0 sichergestellt, für die Auswertung wurde der Tobii Fixation Filter (Velocity Threshold 0,42 pixels/ms; Distance Threshold 35 pixels) verwendet.

Der Einsatz eines monitorbasierten Eye-Trackers ermöglicht den Probanden im Gegensatz zu Geräten, bei denen der Kopf fixiert wird, eine gewisse Freiheit der Kopfbewegung. Gerade im Hinblick auf Schülerinnen und Schüler als Probanden erschien dies ein wichtiger Faktor bei der Gerätewahl. Zudem ist das Gerät während der Studie weniger präsent als mobile *Head-Mounted-Eye-Tracking-Systeme* (HOLMQVIST & ANDERSON, 2017). Nachteil der freien Kopfbewegung ist eine etwas schlechtere Datenqualität, die auch in der vorliegenden Studie den Ausschluss einiger Aufnahmen notwendig machte (vgl. hierzu Kap. 5.5).

5.2 Durchführung

Als erster Schritt der Durchführung der Untersuchung wurde jeweils der Eye-Tracker kalibriert. Hierbei mussten fünf Punkte, die über den gesamten Bildschirm verteilt waren, fixiert werden. „Daruf aufbauend ist es dem System dann möglich [...] den fixierten Bereich des Sichtfeldes zu schätzen“ (BLAKE, 2013, S. 371-372).

Im Anschluss bekamen die Probanden ihre erste Aufgabe gezeigt: „Du wirst gleich eine Schulbuchseite sehen. Schau dir die Schulbuchseite genau an. Wenn du damit fertig bist, gib mir bitte Bescheid“. Nachdem sie mit der Betrachtung der ersten Schulbuchseite fertig waren, wurde die zweite Aufgabe eingeblendet: „Jetzt kommt gleich eine weitere Schulbuchseite. Schau dir die Seite wieder genau an. Gib mir Bescheid, wenn du fertig bist“. Anschließend erschien die zweite Schulbuchseite. Aufgabenstellung und Schulbuchseiten konnten ohne Zeitlimit betrachtet werden.

5.3 Materialien

Für unsere Untersuchung wurden von uns zwei Schulbuchseiten ausgewählt, eine mit dem Thema Landwirtschaft in Spanien, die andere mit dem Thema Erdöl aus der Nordsee. Die Themen sollten lehrplanrelevant, aber noch nicht in der 5. Jahrgangsstufe behandelt worden sein. Außerdem sollten die Themen weder aus dem oberen noch dem unteren Interessenbereich (HEMMER & HEMMER, 2010) stammen, da individuelles Interesse einen Einflussfaktor auf die Betrachtung von Schulbuchseiten darstellt (BEHNKE, 2022). Die Texte und Abbildungen stammen dabei aus verschiedenen Schulbüchern der 6. Jahrgangsstufe für die Realschule in Bayern (BACIGALUPO ET AL., 2008; FRANK ET AL., 2009; KOZEL ET AL., 2009). Beide Schulbuchseiten enthalten ähnlich viel Text und drei Materialien (vgl. Fig. 1). Dabei wurde jeweils eine Abbildung mit viel Inhalt ausgewählt, die verschiedene Informationen aus dem Text bildhaft veranschaulicht (z.B. Aufbau und Funktionsweise einer Erdölraffinerie) und zwei Abbildungen mit vergleichsweise weniger Inhalt. Es wurden in Bezug auf WESTELINCK ET AL. (2005) nur intuitiv verständliche Abbildungen verwendet. Die Anordnung (rechte und linke Spalte) von Text und Abbildungen sind unterschiedlich. Bei allen verwendeten Abbildungen handelt es sich um Originalschulbuchabbildungen, allerdings wurden die Texte der Schulbuchseiten gekürzt, um jeweils problemlos in Originalgröße auf eine Bildschirmseite zu passen.

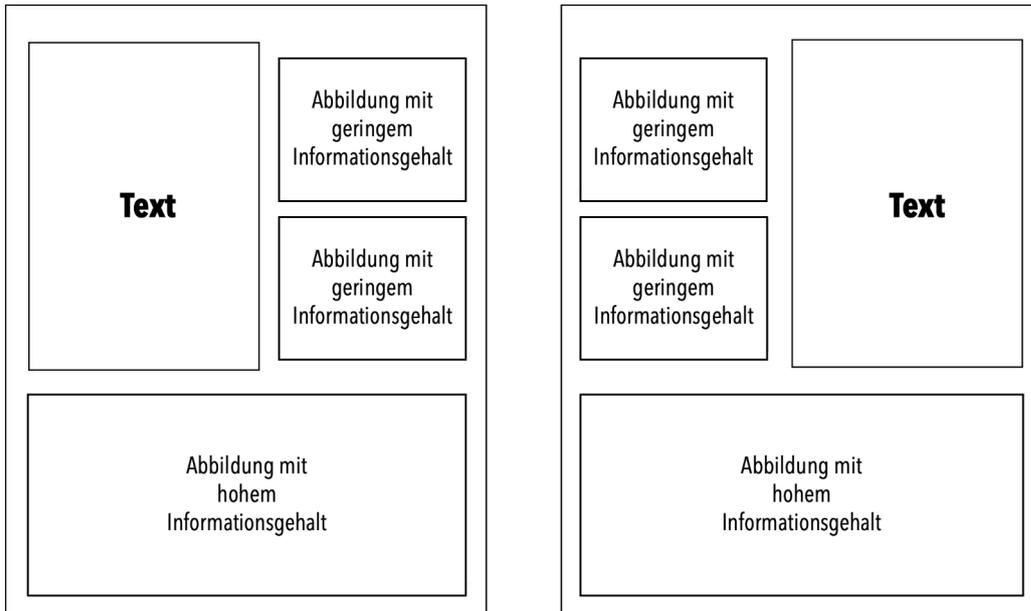


Fig. 1. Aufbau der untersuchten Schulbuchseiten (Quelle: Autorin und Autor)

Die beiden Schulbuchseiten wurden anschließend durch textbasiertes Signaling verändert. Bei den Abbildungen mit geringerem Informationsgehalt geschah dies durch einen Verweis im Text auf die betreffende Abbildung. Dieser Verweis wurde in den Text in Klammern integriert (Fig. 2), wie dies bei dem vorliegenden Verweis auf Fig. 2 gerade geschehen ist. Bei der Abbildung mit viel Inhalt wurden Hinweise im Text und Markierungen in der

Abbildung (vgl. Fig. 2) eingefügt. Hierbei erfolgte beim Verweis auf die Abbildung mit viel Inhalt der Hinweis integriert in den Textfluss (z. B. „In Abbildung 1 ist zu sehen“, „In Abbildung 3 kannst du erkennen“). Außerdem wurden in den Text Einzelhinweise als Zahl in einem kleinen Kreis auf korrespondierende Elemente der Abbildung eingefügt. Das gleiche Symbol findet sich an entsprechender Stelle der Abbildung wieder (vgl. Fig. 2).

In Abbildung 1 ist zu sehen, dass das Wasser aus Staueeen in den Bergen ① kommt. Mit modernen Wasserleitungen ② gelangt dieses im Winter gesammelte Regenwasser ③ zu den Grundstücken der Landwirte ④.

Dort wurde es früher durch Kanäle (Abb. 2) zu den Pflanzen geführt. Dabei verdunstete allerdings viel Wasser. Heute arbeiten dagegen viel Obst- und Gemüsebauern mit der modernen Tröpfchenbewässerung (Abb. 3). Bei dieser gelangt das Wasser in Plastikschräuchen zu den Pflanzen. Feine Löchlein geben ganz zielgerichtet Wasser ab.

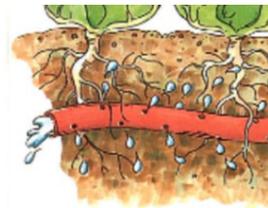


Abb. 3: Moderne Tröpfchenbewässerung

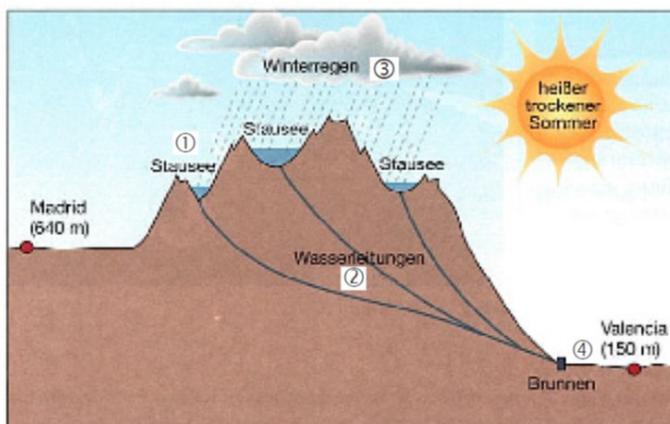


Abb. 1: Woher kommt das Wasser für die Landwirtschaft?

Fig. 2. Schulbuchseite zur Bewässerungslandwirtschaft mit textbasiertem Signaling (Quelle: Autorin und Autor, basierend auf BACIGALUPO ET AL., 2008, S. 71 und KRONFELDNER ET AL., 2009, S. 80)

5.4 Stichprobe

Die Stichprobe bestand aus 76 Probanden zweier Gymnasien und jeweils zwei unterschiedlichen Jahrgangsstufen (Jgst. 5: $n=31$, Jgst. 10: $n=45$). So waren für den einen Teil der Probanden die vorgelegten Materialien nahezu passgenau, für den anderen Teil der Probanden nicht altersgemäß und der Inhalt aus dem vorangegangenen Unterricht bereits bekannt. Der Anteil an Mädchen (37) und Jungen (39) war ungefähr ausgeglichen. Die eine Hälfte der Probanden bekam die Schulbuchseiten ohne Signaling, die andere Hälfte die Schulbuchseiten mit Signaling. Die Zuteilung zu den Probandengruppen erfolgte zufällig. Es wurde einzig auf eine ausgeglichene Geschlechterverteilung geachtet. Die Untersuchungen fanden in den Schulen in einem Besprechungsraum statt. Die Schülerinnen und Schüler stammten jeweils aus mehreren Klassen pro Jahrgang und konnten sich freiwillig für die Studie melden. Über ihre genauen Inhalte wurden im Vorfeld keine Informationen gegeben, um eine besondere Auswahl und Vorbereitung der Probanden zu verhindern.

5.5 Auswertung

Da die Daten teilweise eine schlechte Accuracy aufwiesen, wurden zur Analyse der Betrachtung der Schulbuchseiten die Scanpaths (Blickpfade) der Probanden von zwei Forschern ausgewertet. Ein Scanpath ist ein Video, auf dem die Fixationen der Probanden punktförmig (Punktgröße in Abhängigkeit zur Fixationsdauer) und Sakkaden linienförmig dargestellt werden. Auf diese Weise erhält man einen visuellen Überblick davon, was in welcher Reihenfolge von den Probanden betrachtet wird. Die Auswertung verlief mehrschrittig. Zunächst erfolgte die Ermittlung genutzter Strategien

6. Ergebnisse

6.1 Strategien beim Betrachten der Schulbuchseiten ohne Signaling

Zur Untersuchung der ersten Forschungsfrage wurden die Strategien beim Betrachten der Schulbuchseiten ohne Signaling ausgewertet. In der 5. Jahrgangsstufe wurde die Betrachtung von 34 Schulbuchseiten untersucht, in der 10. Jahrgangsstufe waren es 44. Im Zuge der Qualitätssicherung wurden jeweils ca. 30% der Aufnahmen beider Jahrgangsstufen (Jgst. 5: 10 Seiten, Jgst. 10: 13 Seiten) nicht in die Auswertung einbezogen, da diese in der Regel minimale Aufnahmelücken aufwiesen. Die Ergebnisse sind in [Fig. 3](#) dargestellt.

bei der Seitenbetrachtung im Hinblick auf die zeitnahe Betrachtung korrespondierender Text- und Bildkomponenten in einem induktiv-deduktiven Verfahren, basierend auf der in Kapitel 2 beschriebenen CTML.

Die Scanpaths wurden dahingehend ausgewertet, in welcher Abfolge Schülerinnen und Schüler Text- und Bildelemente anschauen und ob hierbei zusammengehörige Text- und Bildpassagen in unmittelbarer zeitlicher Nähe betrachtet werden, d. h. ob es zu einem Springen der Blickbewegungen von Textelementen zu korrespondierenden Bildelementen bzw. von Bildelementen zu korrespondierenden, bereits gelesenen Textelementen kommt. Eine Zuordnung zu den Strategien erfolgte sowohl für die Bedingung mit als auch ohne Signaling. Bei den Seitenbetrachtungen, bei denen es zu mindestens einer Betrachtung zusammengehöriger Text- und Bildkomponenten in direkter zeitlicher Abfolge kam, wurde zudem anhand der Scanpaths analysiert, bei wie vielen der dargebotenen Abbildungen dies der Fall gewesen ist.

Zur Sicherung der Datenqualität wurden Scanpaths, die Mängel aufgrund kurzer Aufnahmeaussetzer oder Kalibrierungsungenauigkeiten aufwiesen, von der Analyse ausgeschlossen. Die Zuordnung zu einer Strategie durch den Erstcodierer erfolgte durch mindestens dreifache Betrachtung der Scanpaths in reduzierter Abspielgeschwindigkeit. Auch ein Zweitcodierer wurde zur Auswertung herangezogen. Der Cohens-Kappa betrug bei der Zuordnung zu Strategien .88, bei der Fragestellung, ob Bilder und korrespondierende Textelemente in unmittelbarer zeitlicher Abfolge betrachtet worden sind, .94 und zeigte damit insgesamt eine fast perfekte Übereinstimmung an.

Die in [Fig. 3](#) angegebenen absoluten Häufigkeiten beziehen sich ebenso wie die relativen Häufigkeiten auf die Schulbuchseiten, die auf diese Weise betrachtet wurden. Unter Verknüpfung korrespondierender Text- und Bildelemente wird hier wie auch im Folgenden die Betrachtung korrespondierender Text- und Bildelemente in unmittelbarer zeitlicher Abfolge verstanden, auf mentale Prozesse kann hierbei nicht verlässlich geschlossen werden.

Es zeigt sich, dass bei der Betrachtung der Schulbuchseiten sowohl Text- als auch Bildkomponenten in der überwiegenden Mehrheit der Seitenbetrachtungen angeschaut werden. Lediglich in 8% der Fälle (5. Jgst.) bzw. 3% der Fälle (10.

Strategien der Betrachtung von Schulbuchseiten	5. Jgst. (n=24 Seiten)	10. Jgst. (n=31 Seiten)
A Text wird vollständig oder teilweise betrachtet, darauf folgt die Betrachtung einiger oder aller Bilder.	8 S. (33,33%)	6 S. (19,36%)
B Es werden alle Bilder oder ein Teil der Bilder betrachtet und dann der gesamte Text oder ein Teil des Textes gelesen.	0 S. (0%)	0 S. (0%)
C Text- und Bildelemente werden im Wechsel betrachtet, hierbei erfolgt keine Verknüpfung korrespondierender Text- und Bildabschnitte.	7 S. (29,17%)	17 S. (54,84%)
D Text- und Bildelemente werden im Wechsel betrachtet, hierbei erfolgt mindestens einmal eine Verknüpfung korrespondierender Text- und Bildabschnitte.	7 S. (29,17%)	7 S. (22,58%)
E Es werden nur alle Bilder oder ein Teil der Bilder betrachtet, der Text wird nicht betrachtet.	2 S. (8,33%)	1 S. (3,23%)

Fig. 3. Strategien der Betrachtung von Text- und Bildelementen auf Schulbuchseiten ohne Signaling (Quelle: Autorin und Autor)

Jgst.) erfolgt eine ausschließliche Betrachtung von Bildkomponenten der Seite. In keiner der beiden Jahrgangsstufen wird bei einer Seite ausschließlich der Text betrachtet. In 33% der Fälle wird in der 5. Jgst. (Jgst. 10: 19%) zunächst der Text vollständig oder teilweise betrachtet und dann der Blick auf eine oder mehrere Abbildungen geworfen. Bei der Mehrheit der Seitenbetrachtungen erfolgt bei Schülerinnen und Schülern der fünften (58%) und der zehnten Jahrgangsstufe (77%) eine wechselnde Betrachtung von Text und Bild. Allerdings wandert der Blick in unmittelbarer zeitlicher Abfolge zwischen Text und mindestens einer korrespondierenden Abbildung (Strategie D) in der 5. Jahrgangsstufe nur bei 29% der Seitenbetrachtungen, in der 10. Jahrgangsstufe bei 23% der Seitenbetrachtungen. Gemäß einseitigem exaktem Binomialtest wurde bei den Seitenbetrachtungen signifikant häufiger eine Strategie eingesetzt, bei der es zu keiner einzigen unmittelbaren Betrachtung korrespondierender Text- und Bildelemente kommt ($n = 55$, $p < .001$). Gemäß zweiseitigem nach Yates korrigiertem Chi-Quadrat-Test gibt es keinen signifikanten Unterschied betreffend der Häufigkeit der Verwendung von Strategie D in Abhängigkeit von der Jahrgangsstufe, $\chi^2(1, n = 55) = 0.00$, $p = 1.000$.

6.2 Ergebnisse bei der Betrachtung der Schulbuchseiten mit Signaling

Auch unter der Bedingung Signaling wurden die Scanpaths hinsichtlich der angewandten Betrachtungsweise ausgewertet. Insgesamt konnten 60 Seitenbetrachtungen in die Auswertung einbezogen werden. In Jahrgangsstufe 5 wurden aufgrund

mangelhafter Qualität sechs Seitenbetrachtungen (21%), in der 10. Jahrgangsstufe acht Seitenbetrachtungen (18%) von der Analyse ausgeschlossen. Fig. 4 zeigt die absoluten und relativen Häufigkeiten der angewandten Strategien der Betrachtung unter der Bedingung Signaling.

Wird Strategie D bei den Betrachtungen ohne Signaling in der 5. Jahrgangsstufe in 29% der Fälle und in Jahrgangsstufe 10 in 23% der Fälle angewendet, so werden die Betrachtungen der Seiten mit Signaling in Jahrgangsstufe 5 in 91% und in Jahrgangsstufe 10 in 84% der Fälle dieser Strategie zugeordnet. Mittels nach Yates korrigiertem einseitigem Chi-Quadrat-Test wurde für beide Jahrgangsstufen ein hoch signifikanter Zusammenhang zwischen Anwendung/Nichtanwendung der Strategie D und der Bedingung mit oder ohne Signaling festgestellt [Jgst 5: $\chi^2(1, n = 50) = 15.59$, $p < .001$, Jgst. 10: $\chi^2(1, n = 65) = 23.94$, $p < .001$]. Es besteht jeweils ein starker Zusammenhang (Jahrgangsstufe 5: $\phi = .63$; Jahrgangsstufe 10: $\phi = .62$). Entgegen der Annahme ist die Häufigkeit der Nutzung der Strategie D gemäß einseitigem Fisher-Yates-Test in Jahrgangsstufe 5 nicht signifikant größer als in Jahrgangsstufe 10 ($p = .377$).

Fig. 5 stellt dar, wie viele Abbildungen bei Anwendung von Strategie D unter den Bedingungen mit und ohne Signaling jeweils bei einer Betrachtung in direkter zeitlicher Abfolge zum korrespondierenden Text angeschaut werden.

Betrachtet man die 5. Jahrgangsstufe, werden unter der Bedingung ohne Signaling bei Anwendung von Strategie D eine Abbildung in 71%, zwei Abbildungen in 29% und drei Abbildungen in 0% der Fälle in unmittelbarer zeitlicher Nähe zum kor-

Strategien der Betrachtung von Schulbuchseiten	5. Jgst. (n=22 Seiten)	10. Jgst. (n=38 Seiten)
A Text wird vollständig oder teilweise betrachtet, darauf folgt die Betrachtung einiger oder aller Bilder.	1 S. (4,55%)	2 S. (5,26%)
B Es werden alle Bilder oder ein Teil der Bilder betrachtet und dann der gesamte Text oder ein Teil des Textes gelesen.	0 S. (0%)	0 S. (0%)
C Text- und Bildelemente werden im Wechsel betrachtet, hierbei erfolgt keine Verknüpfung korrespondierender Text- und Bildabschnitte.	1 S. (4,55%)	4 S. (10,53%)
D Text- und Bildelemente werden im Wechsel betrachtet, hierbei erfolgt mindestens einmal eine Verknüpfung korrespondierender Text- und Bildabschnitte.	20 S. (90,9%)	32 S. (84,21%)
E Es werden nur alle Bilder oder ein Teil der Bilder betrachtet, der Text wird nicht betrachtet.	0 S. (0%)	0 S. (0%)

Fig. 4. Strategien der Betrachtung von Text- und Bildelementen auf Schulbuchseiten mit Signaling (Quelle: Autorin und Autor)

respondierenden Text angeschaut ($M=1.29$, $SD=0.49$), wohingegen bei Betrachtungen mit Signaling in dieser Jahrgangsstufe in 75% der Fälle alle drei Abbildungen, in 10% der Fälle zwei Abbildungen und in 15% der Fälle nur eine Abbildung verknüpft werden ($M=2.60$, $SD=0.75$). In Jahrgangsstufe 10 werden ohne Signaling pro Betrachtung unter Anwendung von Strategie D durchschnittlich 1.43 Abbildungen ($SD=0.54$) in direkter zeitlicher Abfolge zum korrespondierenden Text betrachtet, unter der Bedingung Signaling sind es 1.97 ($SD=0.74$). Gemäß einseitigem T-Test für unabhängige Stichproben werden in beiden Jahrgangsstufen unter der Bedingung Signaling signifikant mehr Abbildungen in direkter zeitlicher Nähe zum korrespondierenden Text betrachtet, in der 5. Jahrgangsstufe ist das Ergebnis hochsignifikant, [Jgst. 5: $t(27)=4.28$, $p<0.001$, Jgst. 10: $t(39)=1.822$, $p=0.04$]. Aufgrund der geringen Fall-

zahlen wurde jeweils ein einseitiger Mann-Whitney-U-Test durchgeführt, der das signifikante Ergebnis für die 10. Jahrgangsstufe und das hochsignifikante Ergebnis in der 5. Jahrgangsstufe bestätigte, [Jgst. 5: $U(n=27)=124.5$ $z=3.371$, $p<0.001$, Jgst. 10: $U(n=39)=156.5$ $z=-1.760$, $p=.039$]. Vergleicht man die Mittelwerte der 5. Jahrgangsstufe ($M=2.6$, $SD=0.75$) und der 10. Jahrgangsstufe ($M=1.97$, $SD=.74$), ist gemäß einseitigem T-Test für unabhängige Stichproben die durchschnittliche Anzahl der verknüpften Abbildungen in Jahrgangsstufe 5 unter der Bedingung Signaling signifikant höher als in Jahrgangsstufe 10, $t(52)=2.98$, $p=.025$. Dieser signifikante Unterschied wurde in einem aufgrund der kleinen Stichprobe durchgeführten einseitigen Mann-Whitney-U-Test bestätigt, $U(n=52)=173.5$, $z=-2.960$, $p=.001$.

Vergleicht man die Anzahl der Verknüpfungen pro Seitenbetrachtung unter der Bedingung kein

Material/verknüpfte Text-Bildkomponenten pro Seitenbetrachtung	Eine Abbildung	Zwei Abbildungen	Drei Abbildungen
5. Jahrgangsstufe			
Seiten ohne Signaling	5 (71,4%)	2 (28,6%)	0 (0%)
Seiten mit Signaling	3 (15%)	2 (10%)	15 (75%)
10. Jahrgangsstufe			
Seiten ohne Signaling	4 (57,1%)	3 (42,9%)	0 (0%)
Seiten mit Signaling	9 (28,1%)	15 (46,8%)	8 (25%)

Fig. 5. Häufigkeit in direkter zeitlicher Abfolge betrachteter korrespondierender Abbildungen bei Seiten mit und ohne Signaling bei Nutzung von Strategie D (Quelle: Autorin und Autor)

Signaling, so werden mit $M = 1.43$ ($SD = 0.54$) in Jahrgangsstufe 10 etwas mehr Abbildungen pro Betrachtung verknüpft als in Jahrgangsstufe 5

($M = 1.29$, $SD = 0.49$). Der Unterschied ist gemäß einseitigem T-Test aber nicht signifikant, $t(14) = -.522$, $p = .306$.

7. Diskussion

Bei der Betrachtung der Schulbuchseiten ohne Signaling zeigte sich wie in Hypothese 1a vermutet, dass Abbildungen und korrespondierende Textstellen von Schülerinnen und Schülern der 5. und 10. Jahrgangsstufe nicht in unmittelbarer zeitlicher Abfolge betrachtet werden. Der Anteil der Betrachtungen, bei denen mindestens bei einer Abbildung in direkter zeitlicher Abfolge inhaltlich korrespondierende Text- und Bildkomponenten angeschaut werden, liegt unter 30% in beiden Jahrgangsstufen. Hierbei unterscheiden sich die Jahrgangsstufen nicht, wie in Hypothese 1b vermutet, in der Häufigkeit der Anwendung der Strategie, bei der es zumindest einmal zur Betrachtung korrespondierender Text- und Bildelemente in unmittelbarer zeitlicher Abfolge kommt. Wird eine solche Strategie angewandt, so bleibt es in beiden Jahrgängen in mehr als der Hälfte der Fälle (Jgst. 5: 74%, Jgst. 10: 51%) bei der Betrachtung nur einer Abbildung in unmittelbarer zeitlicher Abfolge zur korrespondierenden Textstelle. Die Ergebnisse ähneln denen der Studie von HANNUS und HYÖNÄ (1999) bei Grundschülerinnen und Grundschülern und denen von HOLSANOVA ET AL. (2008) bei erwachsenen Zeitungslesern. Sie zeigen, dass auch Schülerinnen und Schüler in der Sekundarstufe 1 Schulbuchtexte nicht mit einer Strategie lesen, bei der es zu einer systematischen zeitnahen Betrachtung korrespondierender Text- und Bildelemente kommt. Hierbei gibt es, betrachtet man die durchschnittlich verknüpften Text- und Bildkomponenten, keine signifikanten Unterschiede zwischen der 5. und der 10. Jahrgangsstufe. Die Schülerinnen und Schüler scheinen Text und Abbildungen als separate Informationsquellen zu betrachten. Es ist anzunehmen, dass eine Strategie der Verknüpfung von Text- und Bildelementen im Unterricht nicht erlernt wird. Einschränkend ist zu sagen, dass die Ergebnisse anders sein können, wenn Schülerinnen und Schüler spezifischere Aufgabenstellungen vor dem Betrachten einer Seite erhalten (YARBUS, 1967).

Signaling wird das Potential zugeschrieben, die Aufmerksamkeit der Lernenden zu steuern, um eine Integration verbaler und bildhafter Informationen in ein mentales Modell zu erleichtern. Zur Aufmerksamkeitslenkung von Schülerinnen und Schülern bei kontinuierlich dargebotenen Informationen, die selbstgesteuert erfasst werden, ausschließlich durch verbale Hinweise im Text auf statische Abbil-

dungen, lagen bislang unseres Wissens noch keine Forschungsergebnisse vor. In der vorliegenden Studie lag bei der Verwendung von verbalen Hinweisen im Text auf inhaltlich korrespondierende Abbildungen der Anteil der Betrachtungen, bei denen es mindestens einmal zur zeitlich unmittelbar aufeinanderfolgenden Betrachtung korrespondierender Text- und Bildkomponenten kommt, signifikant höher als bei Seitenbetrachtungen ohne Signaling (Jgst. 5: 91% mit Signaling und 29% ohne Signaling, Jgst. 10: 84% mit Signaling und 23% ohne Signaling). Zu ähnlichen Ergebnissen kommen Studien, bei denen die Aufmerksamkeitssteuerung mit anderen Formen des Signalings untersucht wurde (z. B. FOLKER ET AL., 2005; MASON ET AL., 2013). Die Daten erlauben allerdings keine verlässliche Aussage darüber, ob Signaling in dieser Studie tatsächlich zu einem höheren inhaltlichen Verständnis führte. Zudem lassen sich signifikante Unterschiede der Aufmerksamkeitssteuerung durch Signaling zwischen den Jahrgangsstufen erkennen. Bei Schülerinnen und Schülern der 5. Jahrgangsstufe werden durchschnittlich signifikant mehr Abbildungen in direkter zeitlicher Abfolge zum Text betrachtet, wie in Hypothese 3b vermutet. Die Ergebnisse können mit Vorsicht in der Weise gedeutet werden, dass die Probanden der 10. Jahrgangsstufe einen Blick vom Text auf die Abbildungen für das Verständnis des Inhalts nicht für notwendig hielten, da sie vermutlich über ein größeres Vorwissen zu diesen bereits in den vorangehenden Jahrgangsstufen behandelten Themen verfügen. So interpretiert unterstützen die Daten die Informations-Reduktions-Hypothese (HAIDER & FRENSCH, 1999) und decken sich mit den Ergebnissen von RICHTER und SCHEITER (2019) beim Einsatz von Signaling überwiegend durch gleiche Farbgebung korrespondierender Text- und Bildelemente. Kritisch ist anzumerken, dass das Vorwissen in der vorliegenden Studie nicht explizit erhoben wurde. Hierzu bedarf es weiterer Forschung.

BEHNKE (2016a) weist darauf hin, dass Erkenntnisse zum Lernen mit Medien noch nicht hinreichend in die Gestaltung von Geographieschulbüchern eingeflossen sind und in Zukunft stärker beachtet werden sollten. Die Autoren des vorliegenden Artikels unterstützen diese Forderung, allerdings verbunden mit dem Postulat, die Wirksamkeit der Gestaltungsempfehlungen zum Multimedialernen in weiteren Forschungsprojekten auch

bei Schülerinnen und Schülern zu überprüfen. Die vorliegende Studie konnte trotz der bereits skizzierten Einschränkungen einen Beitrag hierzu leisten, indem aufgezeigt wurde, dass textbasiertes Signaling auf Abbildungen auf Schulbuchseiten auch bei Schülerinnen und Schülern der 5. und 10. Jahrgangsstufe das Blickverhalten signifikant verändern kann und es hierdurch zur signifikant häufigeren Betrachtung korrespondierender Text- und Bildkomponenten in direkter zeitlicher Abfolge kommt. Die qualitative Analyse der Blickpfade sollte in auf dieser Studie aufbauenden Studien durch quantitative Analysen (z. B. einer Transitionmatrix) ergänzt werden. Zudem sollten künftige Studien innerhalb der Geographiedidaktik neben der Untersuchung der visuellen Aufmerksamkeit von Schülerinnen und Schülern auch die Effekte auf den Lernerfolg untersuchen. Hierbei sollten gezielt originär geographische Darstellungen (z. B. Karten, Klimadiagramme) in den Fokus der Forschung zum multimedialen Lernen innerhalb der Geographiedidaktik gerückt werden.

Für die Gestaltung von Unterrichtsmaterialien wie Schulbuchseiten oder Arbeitsblättern im Geographieunterricht scheint im Hinblick auf die Ergebnisse der vorliegenden Studie und vor dem Hintergrund des aktuellen Forschungsstandes zur Wirksamkeit des Signaling-Prinzips (vgl. MAYER & FIORELLA, 2014; RICHTER ET AL., 2016; VAN GOG, 2014) das Einfügen verbaler Hinweise im Text auf Abbildungen oder Teile der Abbildungen insbesondere dann sinnvoll, wenn Schülerinnen und Schüler über kein großes Vorwissen zu einer Thematik verfügen. Dies sollte auch bei der Konzeption binnendifferenzierender Materialien Berücksichtigung finden. Vorwissen im Unterricht zu diagnostizieren erscheint hierbei von großer Bedeutung. Zudem sollten Schülerinnen und Schüler im Geographieunterricht befähigt werden, Strategien der Verknüpfung von Text- und Bildelementen selbstständig anzuwenden.

Literatur

- BACIGALUPO, S., HEINDL, C., KRUG, S., SCHATZ, A., SCHREIEGG, J., & THALMAIER, J. (2008). *Seydlitz Erdkunde 6 Realschule Bayern*. Bildungshaus Schulbuchverlage.
- BAGOLY-SIMÓ, P. (2014). Traditionen geographiedidaktischer Schulbuchforschung – eine Metaanalyse. In P. KNECHT, E. MATTHES, S. SCHÜTZE, & B. AAMOTSBAKKEN (Hg.), *Methodologie und Methoden der Schulbuch- und Lehrmittelforschung* (S. 111–123). Julius Klinkhardt.
- BEHNKE, Y. (2016a). [Wie betrachten und bewerten Lernende Geographieschulbuchseiten?](#) *Zeitschrift für Geographiedidaktik | Journal of Geography Education*, 44(3), 5–34.
- BEHNKE, Y. (2016b). How Textbook Design May Influence Learning with Geography Textbooks? *Nordidactica – Journal of Humanities and Social Science Education*, 2016(1), 38–62.
- BEHNKE, Y. (2018). Textbook Effects and Efficacy. In E. FUCHS & A. BOCK (Hg.), *The Palgrave Handbook of Textbook Studies* (S. 383–398). Palgrave.
- BEHNKE, Y. (2022). [Welche Faktoren beeinflussen die visuelle Aufmerksamkeit Lernender für Abbildungen in geographischen Lernmedien?](#) *Zeitschrift für Geographiedidaktik | Journal of Geography Education (ZGD)*, 50(3), 159–179.
- BEHNKE, Y. (2023). [Wie interagieren Lernende visuell mit Fotos, Grafiken und Karten in Geographieschulbüchern während der Bearbeitung einer Lernaufgabe?](#) *Zeitschrift für Geographiedidaktik | Journal of Geography Education (ZGD)*, 51(1), 14–47.
- BLAKE, C. (2013). Eye-Tracking. Grundlagen und Anwendungsfelder. In W. MÖHRING & D. SCHLÜTZ (Hg.), *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft* (S. 367–387). Springer.
- CHENG, M.-C., CHOU, P.-I., WANG, Y.-T., & LIN, C.-H. (2015). [Learning Effects of a Science Textbook Designed with Adapted Cognitive Process Principles on Grade 5 Students](#). *International Journal of Science and Mathematics Education*, 13(3), 467–488.
- CLARK, J. M., & PAIVIO, A. (1991). [Dual Coding Theory and Education](#). *Educational Psychology Review*, 3(3), 149–210.
- DÉSIRON, J. C., BÉTRANCOURT, M., & DE VRIES, E. (2018). How Cross-Representational Signaling Affects Learning from Text and Picture: An Eye-Tracking Study. In P. CHAPMAN, G. STAPLETON, A. MOKTEFI, S. PEREZ-KRIZ, & F. BELLUCCI (Hg.), *Diagrammatic Representation and Inference. 10th international Conference, Diagrams 2018, UK, June 18-22, Proceedings (Band 5)* (S. 725–728). Springer.
- DUCHOWSKI, A. (2007). *Eye Tracking Methodology. Theory and Practice*. Springer.
- FOLKER, S., RITTER, H., & SICHELSCHEMIDT, L. (2005). Processing and Integrating Multimodal Material:

- The Influence of Color Coding. In B. G. BARA, L. BARSALOU, & M. BUCCIARLLI (Hg.), *Proceedings of 27th Annual Conference of the Cognitive Science Society* (S. 690–695). Erlbaum.
- FRANK, S., HAUNSTETTER, A., KÜHNE, K., PYRITZ, E., SALOMON, U., & WOSTE, R. (2009). *Terra Erdkunde 6 Realschule Bayern*. Ernst Klett Verlag.
- GEISE, S. (2011). [Eyetracking in der Kommunikations- und Medienwissenschaft: Theorie, Methode und kritische Reflexion](#). *Studies in Communication Media*, 2011(2), 149–263.
- HAIDER, H., & FRENSCH, P. A. (1999). [Information Reduction During Skill Acquisition: The Influence of Task Instruction](#). *Journal of Experimental Psychology*, 5(2), 129–151.
- HANNUS, M., & HYÖNÄ J. (1999). [Utilization of Illustrations during Learning of Science Textbooks Passages among Low- and High-Ability Children](#). *Contemporary Educational Psychology*, 24(2), 95–123.
- HEMMER, I., & HEMMER M. (Hg.). (2010). *Schülerinteresse an Themen, Regionen und Arbeitsweisen des Geographieunterrichts*. HGD.
- HOLMQVIST, K., & ANDERSON, R. (2017). *Eye Tracking. A Comprehensive Guide to Methods, Paradigms, and Measures*. Lund Eye-Tracking Research Institute.
- HOLSANOVA, J., HOLMQVIST, K., & HOLMBERG, N. (2008). [Reading Information Graphics: The Role of Spatial Proximity and Dual Attentional Guidance](#). *Applied Cognitive Psychology*, 23, 1215–1226.
- JARODZKA, H., HOLMQVIST, K., & GRUBER, H. (2017). [Eye Tracking in Educational Science. Theoretical Frameworks and Research Agendas](#). *Journal of Eye Movement Research*, 10(1), 1–18.
- JUST, M. A., & CARPENTER, P. A. (1980). [A Theory of Reading. From Eye Fixations to Comprehension](#). *Psychological Review*, 87(4), 329–354.
- KALYUGA, S. (2014). The Expertise Reversal Principle in Multimedia Learning. In R. E. MAYER (Hg.), *The Cambridge Handbook of Multimedia Learning* (S. 576–597). Cambridge University Press.
- KLAUER, K. J., & LEUTNER, D. (2007). *Lehren und Lernen. Einführung in die Instruktionspsychologie*. Beltz.
- KLAUER, K. J. (2011). *Transfer des Lernens. Warum wir oft mehr Lernen als gelehrt wird*. Kohlhammer.
- KRONFELDNER, H., LAST, S., PETERHOFF, F., PETZOLD, W., SCHAAL, C., STADLER, K., & VOSSEN, J. (2009). *Diercke Erdkunde RS Bayern 6 Jgst*. Westermann.
- LATHAN, H. (2021). [Das geographische Schulbuch aus der Schülerperspektive: Ein Beitrag zur Anwenderorientierung in Lehr-Lern-Materialien](#) [Dissertation, Universität Vechta].
- MASON, L., PLUCHINO, P., & TORNATORA, M. C. (2013). [Effects of Picture Labeling on Science Text Processing and Learning: Evidence From Eye Movements](#). *Reading Research Quarterly*, 48(2), 199–214.
- MAUTONE, P. D., & MAYER, R. E. (2001). [Signaling as a Cognitive Guide in Multimedia Learning](#). *Journal of Educational Psychology*, 93(2), 377–389.
- MAYER, R. E. (2001). *Multimedia Learning*. Cambridge University Press.
- MAYER, R. E. (2003). [The Promise of Multimedia Learning: Using the Same Instructional Design Methods Across Different Media](#). *Learning and Instruction*, 13(2), 125–139.
- MAYER, R. E. (2008). [Applying the Science of Learning: Evidence-Based Principles for the Design of Multimedia Instruction](#). *American Psychologist*, 63(8), 760–769.
- MAYER, R. E. (2014). Cognitive Theory of Multimedia Learning. In R. E. MAYER (Hg.), *The Cambridge Handbook of Multimedia Learning* (S. 43–71). Cambridge University Press.
- MAYER, R. E., & FIORELLA, L. (2014). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles. In R. E. MAYER (Hg.), *The Cambridge Handbook of Multimedia Learning* (S. 279–315). Cambridge University Press.
- MAYER, R. E., & MORENO, R. (2003). [Nine Ways to Reduce Cognitive Load in Multimedia Learning](#). *Educational Psychologist*, 38(1), 43–52.
- MORENO, R., & MAYER, R. E. (1998). [A Split-Attention in Multimedia Learning: Evidence for Dual Processing Systems in Working Memory](#). *Journal of Educational Psychology*, 90(2), 312–320.
- NIEDING, G., OHLER, P., & REY, G. D. (2015). *Lernen mit Medien*. Ferdinand Schöningh.
- OZCELIK, E., KARAKUS, T., KURSUN, E., & CAGILTAY, K. (2009). [An Eye-Tracking Study of How Color Coding Affects Multimedia Learning](#). *Computers and Education*, 53(2), 445–453.
- PAAS, F., & SWELLER, J. (2014). Implications of Cognitive Load Theory for Multimedia Learning. In R. E. MAYER (Hg.), *The Cambridge Handbook of Multimedia Learning* (S. 27–42). Cambridge University Press.
- PAIVIO, A. (1986). *Mental Representations: A Dual Coding Approach*. Oxford University Press.
- RAYNER, K. (1998). [Eye Movements in Reading and Information Processing: 20 Years of Research](#). *Psychological Bulletin*, 124(3), 372–422.
- RICHTER, J., & SCHEITER, K. (2019). [Studying the Expertise Reversal of the Multimedia Signaling](#)

- [Effect at a Process Level: Evidence from Eye Tracking](#). *Instructional Science*, 47, 627–658.
- RICHTER, J., SCHEITER, K., & EITEL, A. (2016). [Signaling Text-Picture Relations in Multimedia Learning: A Comprehensive Meta-Analysis](#). *Educational Research Review*, 17, 19–36.
- RÖTTING, M., & SEIFERT, K. (Hg.). (1999). *Blickbewegungen in der Mensch-Maschine-Systemtechnik*. Pro Universitate Verlag.
- SCHEITER, K., & EITEL, A. (2015). [Signals Foster Multimedia Learning by Supporting Integration of Highlighted Text and Diagram Elements](#). *Learning and Instruction*, 36, 11–26.
- SCHNEIDER, S., BEEGE, M., NEBEL, S., & REY, G. D. (2018). [A Meta-Analysis of How Signaling Affects Learning with Media](#). *Educational Research Review*, 23, 1–24.
- SWELLER, J. (1999). *Instructional Design in Technical Areas*. ACER Press.
- SWELLER, J. (2010). [Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load](#). *Educational Psychology Review*, 22(2), 123–138.
- VAN GOG, T. (2014). The Signaling (or Cueing) Principle in Multimedia Learning. In R. E. MAYER (Hg.), *The Cambridge Handbook of Multimedia Learning* (S. 263–278). Cambridge University Press.
- WESTELINCK, K. DE, VALCKE, M., CRAENE, B. DE, & KIRSCHNER, P. (2005). [Multimedia Learning in Social Sciences: Limitations of External Graphical Representations](#). *Computers in Human Behavior*, 21(4), 555–573.
- YARBUS, A. L. (1967). *Eye Movements and Vision*. Plenum Press.

Autorin und Autor

Prof. em. Dr. Gabriele Obermaier

Universität Bayreuth
 Universitätsstraße 30
 95447 Bayreuth
 gabriele.obermaier@uni-bayreuth.de

✉ Dr. Dominik Conrad

Pädagogische Hochschule Ludwigsburg
 Reuteallee 46
 71634 Ludwigsburg
 dominik.conrad@ph-ludwigsburg.de