

Unterstützen Erklär-Videos beim Experimentieren an Lernstationen?

Die hier vorgestellte Feldstudie untersucht die Auswirkungen einer Scaffolding-Maßnahme durch digitale Medien (Lehrvideos als multimedial aufbereitete Darstellungen) auf die Handlungsregulation sowie das Niveau der begrifflichen Repräsentation von Erklärungen für Phänomene, die von Dritt- und Viertklässlern in naturwissenschaftlichen Experimenten an halboffenen Experimentierstationen (im außerschulischen Bereich) untersucht wurden. Die Ergebnisse zeigen keinen Effekt auf die Handlungsregulation, einen negativen Effekt auf den Lernerfolg, keinen Moderatoreffekt der Handlungsregulation und einige differentielle Effekte bezüglich der Klassenstufe, des Geschlechts und der Durchführung der Experimente.

1 Einleitung

Die hier präsentierte Feld-Studie untersucht Effekte instruktorischer Stützung durch digitale Medien (Erklär-Videos im Sinne multimedial aufbereiteter externer analoger Repräsentationen¹) auf die Handlungsregulation sowie das Niveau der begrifflichen Repräsentation von Erklärungen für Phänomene, die im Rahmen naturwissenschaftlicher Experimente von Dritt- und Viertklässler*innen an halboffenen Experimentierstationen (im außerschulischen Bereich) untersucht wurden.

Die Studie steht im Zusammenhang mit der Evaluation eines durch die Deutsche Telekom Stiftung geförderten Projektes (Digitales Lernen Grundschule). Es handelt sich dabei um Praxis-, Feld- bzw. Entwicklungsforschung, bei der die Praxisbedingungen beachtet werden sollten. Es ist hier im Gegensatz zu Laborstudien oder auch empirisch-analytischen quasi-experimentellen Studien nicht möglich, über Variablenkontrolle alle wesentlichen Bedingungen der untersuchten Effekte zu kontrollieren, noch weitgehende Anforderungen an die Randomisierung des Samples zu realisieren. Daher sind keine Aussagen darüber zu erwarten, ob und unter welchen sauber kontrollierten Bedingungen bestimmte Effekte (z.B. des Lernens) ggf. zu beobachten sind. Im Zentrum der Studie steht dagegen die Frage, welche Effekte unter „normalen“ Feld- oder Praxisbedingungen beobachtbar und in der (gegebenen) Praxis beeinflussbar sind (vgl. auch Giest 2022a). Sie wirft daher auch ein Licht auf Praxisprobleme bei der Digitalisierung schulischen Lernens.

2 Problemdiskussion

2.1 Naturwissenschaftliches Lernen im Sachunterricht

Scientific Literacy ist ein wichtiger Zielhorizont des Unterrichts in der naturwissenschaftlichen Perspektive im Sachunterricht und unumstritten eng mit dem Experimentieren als Lernhandlung verknüpft. Vor allem quasiexperimentelle Studien haben gezeigt, dass die damit verbundenen hohen Anforderungen an

¹ Zum Problem der Unschärfe des Begriffs vgl. Matthes, Siegel und Heiland (2021).

- entsprechendes Lernhandeln (zur *selbstständigen Handlungsregulation* und zu Variablenkontrollstrategien siehe Schwichow et al. 2016 sowie Laufs und Kempert 2021; zu Analogien/Modellen siehe Böschl et al. 2018, 2019 sowie Gogolin et al. 2017) als auch
- die Kognition (zur *Begriffsbildung* – Conceptual Change – siehe Vosniadou 2015; zur Aneignung adäquaten Wissens und zum Wissenschaftsverständnis/zur Erkenntnishaltung siehe Koerber et al. 2015 sowie Grygier et al. 2008; zur Repräsentationskompetenz, kognitiven Strukturierung und Aktivierung siehe Minnameier, Hermkes & Mach 2015)

bei geeigneter instruktionaler Stützung (Scaffolding) sowohl aus kognitions-, entwicklungs-, lernpsychologischer und didaktischer Sicht auch in der Grundschule erreichbar erscheinen (Alferi et al. 2011; Kalthoff, Theyßen & Schreiber 2016; Haslbeck et al. 2019; Bohrman, Todorova & Möller 2017; Eschrich 2019; Pahl & Lück 2016; Windt, Scheuer & Melle 2011). Unklar ist jedoch, ob dies auch im Rahmen einer Feldstudie gelingen kann, wo sich Bedingungen instruktionaler Stützung nicht detailliert kontrollieren lassen (Giest 2022c).

Dies gilt besonders hinsichtlich des Erreichens eines angemessenen Niveaus der *begrifflichen Repräsentation* des Wissens. Hier bestehen besondere Lernprobleme darin, dass auf einer neuen gedanklichen Ebene die zunächst auf dem Hintergrund von Alltagsanforderungen konstruierten Begriffe und ihre Relationen neu gebildet, umgebildet oder parallel zu den Alltagsbegriffen neu kontextuiert und repräsentiert werden müssen (Vosniadou a.a.O.). Psychologisch bedeutet dies, dass die begrifflich gefassten Sachverhalte mit neuen (naturwissenschaftlichen) Merkmalen (neuer Merkmalsatz) und Relationen verbunden werden und andererseits eine neue (wissenschaftliche) Bedeutung/einen neuen Sinn erlangen müssen. Kognitionspsychologisch bedeutet das, neben der Repräsentation der Phänomene auf einer Oberflächenebene eine Repräsentation auf Tiefenebene aufzubauen.

Beide unterscheiden sich durch unterschiedliche Bezüge zum Kontext (Alltag vs. Wissenschaft/Fach), zur Anschaulichkeit (anschaulich-konkret, unanschaulich-abstrakt), zum praktisch-gegenständlichen Handeln (hand-on- vs. minds-on-Aktivität, vgl. Kaltman 2010; Marquardt-Mau & Hoffmann 2009) und zur Art der Begriffsbildung (empirische vs. theoretische Begriffsbildung, vgl. Vygotskij 2002; Giest 2020). Um diesen Conceptual Change abzubilden, werden in einschlägigen Untersuchungen (nicht nur) zum Sachunterricht drei Niveaus der begrifflichen Repräsentation unterschieden und empirisch untersucht: 1. naive Vorstellungen, 2. Zwischenvorstellungen sowie 3. wissenschaftliche Vorstellungen (vgl. Vosniadou 2015; Hardy et al. 2010; Koerber et al. 2011, 2015; Edelsbrunner et al. 2018). Weitgehend unklar ist jedoch, wie ein diesen begrifflichen Wandel beförderndes, vor allem gegenstands- bzw. domänenspezifisches adäquates Maß instruktionaler Unterstützung beschaffen ist (siehe 2.3).

2.2 Experimentieren als Lernhandlung

Das Experimentieren ist eine wesentliche naturwissenschaftliche Lernhandlung und Erkenntnisweise bzw. -methode, von deren Beherrschung sowohl das Verstehen bzw. Erklären naturwissenschaftlicher Phänomene als auch die emotional-motivationale Haltung gegenüber den Naturwissenschaften wesentlich abhängt (GDSU 2013, vgl. auch Anders et al. 2013).

Unterscheidet man zwischen hypothesenprüfendem Experimentieren und freiem Explorieren (vgl. etwa Köster 2018, Schütte 2019) so ist aus theoretischer Sicht unstrittig, dass naturwissenschaftliches (hypothesentestendes) Experimentieren, welches dem Ziel des Erwerbs fachlichen Wissens dienen soll, instruktionaler Unterstützung bedarf (Ramseger 2013) während freies Explorieren im Sinne der Selbstbildung ohne explizite Instruktion auskommt. In Theorie und Praxis wird hinsichtlich des hypothesentestenden Experimentierens einerseits kritisiert, dass Kinder zu häufig ausschließlich manuell aktiv sind und andererseits, dass sie zu „Handlangern“ (Elschenbroich 2005) gemacht werden und selbständiges Handeln zu kurz kommt (vgl. insgesamt zum Problem Murmann et al. 2019; Zahdeh & Peschel 2018; Michalik 2010; Waldenmaier, Köster & Müller 2013; Waldenmaier et al. 2015).

Unklar ist daher weiterhin, wie in Praxissituationen beim Experimentieren ein angemessenes Maß an instruktionaler Stützung (u.a. kognitive Strukturierung und Lerner*innenaktivierung) gesichert werden kann, ohne dass die Eigenregulation beim Lernen (Sinnbildung) zu kurz kommt. Das gilt in besonderem Maße für das Experimentieren an außerschulischen Lernorten. Die dort erzielten Lerneffekte sind klein (Guderian 2007) u.a., weil die Einbindung in den Unterricht und eine inhaltlich abgestimmte Aktivierung der Lernenden nicht gelingen (Itzek-Greulich et al. 2015; Alferi et al. 2011; Lewalter & Geyer 2009). Dennoch werden sie durch Politik, Öffentlichkeit und Wirtschaft unterstützt und auch finanziell gefördert (Freericks 2011; Scharfenberg et al. 2019).

2.3 Multimediale Lernstützung (Erklär-Videos)

Beim Erklären experimentell erzeugter Phänomene muss die Phänomen- von der Theorieebene getrennt werden, um Oberflächen- und Tiefenstrukturen zu erfassen (Kori et al. 2014).

Dazu sind im Rahmen forschenden Lernens (inquiry-based learning, siehe Pedaste et al. 2015; siehe auch Alifieri et al. 2011) instruktional gestützte Reflexion (Kori et al. a.a.O.) sowie instruktionale Hilfen erforderlich (individualisiert, Reflexion unterstützend und dazu anregend, Ideen zu unterscheiden), wobei computerbasierte Lernumgebungen und Multimedia erfolgreiche Lernhilfen sein können (Gerard et al. 2015; van Joolingen et al. 2005; Becker et al. 2020; Matthes, Siegel & Heiland 2021).

Allerdings weisen die meisten Studien (z.B. hinsichtlich der Lernunterstützung beim Konstruieren mentaler Modelle) bezüglich der Lerneffekte nur geringe Effektstärken auf (Mayer 2021; Leutner & Brünken 2017), weshalb Lowe und Boucheix (2016) vorschlagen, den Schwerpunkt der Forschung nicht auf die Verbesserung der (äußeren) Repräsentation, sondern des Lernens (d.h. der Konstruktionsprozesse) zu legen. Vor allem soll der Prozess der Dekomposition und Komposition durch Herausarbeitung überschaubarer funktionaler Einheiten in den Animationen erleichtert werden. Merkmale für eine lernunterstützende Wirkung sind u.a. die Elementarisierung von Inhalten (Einbezug von Alltagsvorstellungen und schrittweises Vorgehen, siehe Reinhold 2010), die Beachtung von Gestaltungsprinzipien multimedialen Lernens (Mayer 2014, 2021) sowie die sinnvolle Einbettung in den Unterricht (Cognitive Apprenticeship, siehe Kulgemeyer & Wolf 2016).

Eine besondere Form computerbasierter Lernumgebungen sind interaktive Erklär- oder Lern-Videos. Diese folgen der Intention, bei User*innen (Betrachter*innen) ein Verständnis zu erreichen bzw. einen Lernprozess auszulösen (Wolf 2015; Krämer & Böhrs 2017; Kugelmeyer et al. 2016). Sie können das Lernen unterstützen (Findeisen, Horn & Seifried 2019; Gaubitz 2021; Scheid et al. 2014; Jebe et al. 1919; Mayer 2021; Matthes, Siegel & Heiland 2021) und werden von 75% der Schüler*innen² (genutzt siehe Jebe et al. a.a.O.). Häufig betrifft die Lernunterstützung jedoch die Übung und Prüfungsvorbereitung und weniger inhaltliches Verstehen bei anspruchsvollen Inhalten (Chuang 2015). Für die Erklärung der Effekte werden drei Ansätze verwendet (representational approach, cognitive approach und instructional approach), wobei die Zusammenhänge zwischen Videoinhalten, Instruktionsdesign und Lernergebnissen, die Interaktion von Designfaktoren und Lernstrategien sowie die Identifikation nützlicher Designfaktoren in experimentellen Studien genauer zu untersuchen sind (Bétrancourta & Benetos 2018). Unklar ist vor allem, von welchen Bedingungen positive Lerneffekte abhängen (Fiorella & Mayer 2018; Kay 2012). Hinsichtlich der Grundschule (Nicolaou & Constantinou 2014) und des Sachunterrichts (Gervé 2016) gibt es bezüglich des computerbasierten Lernens und besonders mit Blick auf Erklär-Videos (Gaubitz 2021) allerdings kaum Studien.

Bekannt sind hier lediglich die stetige Zunahme der vor allem privaten Ausstattung mit und Nutzung von digitalen Medien, die diesbezüglich hohe Motivation der Kinder, Kom-

² Die Stichprobe umfasste 818 deutschsprachige Jugendliche im Alter von 12 bis 19 Jahren.

petenzdefizite bei der ziel- und aufgabenbezogenen Nutzung sowie ein noch weitgehend ungenutztes Potenzial zur Individualisierung von Lernprozessen (Gervé a.a.O.).

3 Fragestellung

Im Rahmen unserer Untersuchung³ zur instruktionalen Wirkung von Erklär-Videos gingen wir folgenden Teilfragen nach:

- 1) Wie groß sind die Lerneffekte der Bearbeitung der Experimentierstationen?
- 2) Unterstützen Erklär-Videos die (eigenständige) Handlungsregulation?
- 3) Unterstützen Erklär-Videos die Lernergebnisse im Sinne der Orientierung auf ein angemessenes Niveau der begrifflichen Repräsentation von Erklärungen untersuchter Phänomene beim Experimentieren und welche Rolle spielt der Inhalt der Station dabei?
- 4) Welche Rolle spielt das Vorwissen mit Blick auf die Handlungsregulation, die Lernergebnisse und die Nutzung der Erklär-Videos?
- 5) Wirkt die Handlungsregulation indirekt auf die Beziehung zwischen Vorwissen und Lernwerten oder moderiert sie letztere über die Wechselwirkung mit den anderen unabhängigen Variablen?
- 6) Gibt es differenzielle Effekte bezüglich der Videonutzung hinsichtlich der Handlungsregulation und der Lernwerte bzgl. der Faktoren Gender, Klassenstufe und (Inhalt der) Station?

4 Methoden

4.1 Design der Untersuchung

Untersucht wurden Merkmale der *Handlungsregulation* und *Lerneffekte* halboffenen Experimentierens an drei Stationen. Diese im Rahmen eines Vorläuferprojektes (Giest 2022a, b; Evaluationsbericht 2016) entwickelten und erprobten Stationen sind:

- 1) *Wasserhaut* (experimentelles Erforschen der Oberflächenspannung des Wassers (Schwimmen einer Büroklammer auf Wasser: „Eisen schwimmt doch“; Löwenzahn-Experiment): Besondere Charakteristik ist hier, dass der aufzuklärende Effekt (Büroklammer liegt auf der Wasseroberfläche) selbst erzeugt wird. Ein hohes Maß an eigenreguliertem Handeln ist zu erwarten, wobei offenbleibt, ob das Verstehen des Zusammenhangs durch die Stützungsmaßnahmen erreicht wird.
- 2) *Smog* (experimentell den Effekt aufklären, dass Rauch einer Räucherkerze im Glas bleibt, wenn ein Teelicht darüber entzündet wird – Smogbildung): Der Effekt lässt sich nur nach genauem Ablauf erzeugen, wobei dieser mitunter aus technischen Gründen schwer zu beobachten ist. Dadurch sind eher geringe Werte bei der eigenständigen Handlungsregulation zu erwarten, die ggf. mit Verständnisproblemen verbunden sind, da bei der Erklärung der Widerspruch der Inversion (oben wärmer als unten) zur Alltagserfahrung (warmes Gas [Rauch/Luft] steigt immer auf) beachtet werden muss.
- 3) *Schwimmen und Sinken* (durch Umformung einer Knetrolle [zum Schwimmen bringen] den Effekt des Auftriebs experimentell erkunden): Das Experiment gestattet zwar, den Effekt experimentell relativ leicht zu erzeugen, die Untersuchung der Ursachen erfordert jedoch vorgeplante Schritte. Es wird die Wasserverdrängung von Knetboot und Knetkugel bzw. -rolle verglichen und auf den Auftrieb (Druck des verdrängten Wassers) geschlossen. Daher könnte die eigenständige Handlungsregulation insgesamt geringer ausfallen, wobei fraglich bleibt, inwiefern die Stützungsmaßnahmen ein Verständnis fördern.

³ Mitglieder der Projektgruppe waren Tobias Bach, Hartmut Giest, Helvi Koch, Riccardo Krenzel und Christian Schlossnickel.

Da sich im Rahmen der Erprobung der oben genannten Stationen die Notwendigkeit einer weitergehenden instruktionalen Stützung ergab, wurden jeweils html5-programmierte multimediale Präsentationen (Text [visuell und auditiv präsentiert], Bilder, Videos, Animationen) erarbeitet (siehe Abb. 1).

Abb. 1: Beispielseite mit interaktiv startbarer Video- und Audiodatei (für den Fall eines browserbedingten Ausfalls des Hintergrundtextes)



Bei der Erstellung der Erklär-Videos wurde beachtet, dass

- sie aktivitätsfördernd sind (z.B. durch inkludierte Arbeits- und Handlungsanregungen),
- duale Kanäle ansprechen (akustisch und visuell),
- die kognitive Kapazität nicht überlasten (gleichzeitig zu verarbeitende Informationen wurden begrenzt, nur kurze Texte und wenige Bilder zu einem Inhalt auf einer Seite platziert),
- Animationen nur dann verwenden, wenn die Passung zum Vorwissen angenommen werden kann und sie ein Maximum an Kontrolle des Lernens erlauben,
- eine anforderungsbezogene Passung zwischen Bild und Text erfolgt, wobei bei dualen Kanälen beschreibende Repräsentation (symbolisch über Text) und bebilderte Repräsentation (Bilder – analog) unterschieden werden,
- sowie berücksichtigt wird, dass statische Bilder auf Faktenlernen wirken und die Lernplanung erleichtern, dynamische Bilder (Animationen) das Verstehen fördern aber nicht das Faktenlernen, Pfeile und Rahmen aber den gleichen Effekt erreichen können (vgl. insgesamt Lowe & Boucheix 2016; Richter, Scheiter & Eitel 2016; Mayer 2014, 2021; Mayer & Fiorella 2022).

Zugänglich waren diese Erklär-Videos über Tablets (5 pro Station), wobei in der Regel ein oder zwei Kinder damit arbeiteten.

Betrachtet man die in der Literatur genannten Gestaltungs- und Einsatzprinzipien (Mayer 2021), so fehlen in unserem Projekt Vorübungen bezüglich der Schlüsselkonzepte, ein*e sichtbare*r, sprechende*r und gestikulierende*r Instrukteur*in (in unseren Videos läuft gesprochenener Text im Hintergrund bzw. kann bei technischer Inkompatibilität durch Anklicken aktiviert werden), Prompts für die Anwendung bestimmter Lernstrategien (Generative Learning) – wobei anzumerken ist, dass nicht alle Lernstrategien für Grundschüler*innen im betrachteten Alter relevant sind (Brod 2021), denn hier spielen vor allem das Vorhersa-

gen/Vermuten und praktische Prüfen/Probieren eine positive Rolle, weniger das Beantworten von Testfragen, das Formulieren eigener Antworten und Erklärungen und kaum das Concept Mapping, das Ableiten von Fragen und die Veranschaulichung durch eigene Zeichnungen. Ferner ist zu beachten, dass Grundschüler*innen Probleme bei der kognitiven und metakognitiven Kapazität sowie bei der Aktivierung von Vorwissen haben (a.a.O.).

Die Evaluation der Experimentierstationen erfolgte im Rahmen von „*Experimentiertagen*“ in der Pädagogischen Werkstatt der Universität Potsdam. Dabei wurde folgendes Forschungsdesign verwendet:

- Nach der Begrüßung und kurzen Einweisung erfolgt ein Pretest im Plenum („Quiz“ mit multiple Choice-Aufgaben zu jeder Station).
- Die Kinder besuchen und wechseln selbstbestimmt die Stationen, sehen das Erklär-Video an und beginnen mit der Arbeit, wobei ein Forscher*innenheft (enthält Beschreibungen der Experimente und Lückentexte, die beim Aufstellen von Vermutungen und beim Fixieren der Beobachtungen und Erklärungen unterstützen sollen) genutzt werden *kann*. Die Tutor*innen an der Station helfen *bei Bedarf*. Während der Arbeit an den Stationen werden die anwesenden Kinder beobachtet und z.T. videographiert.
- Am Ende erfolgt ein Posttest mit identischen Aufgaben im Plenum und z.T. wurden zur Illustration der Ergebnisse strukturierte Interviews an den Stationen durchgeführt und audiographiert. Dazu wurden festgelegte Fragen zur Erklärung des im Experiment untersuchten Phänomens und eine Anwendungsfrage gestellt (Beispiel: Kannst du mir sagen, warum der Rauch von der Räucherkerze im Glas bleibt? Was hat denn ein qualmender Ofen mit Smog zu tun?). Aus den Antworten wurden charakteristische Kategorien extrahiert (Beispiel: Benennung des Phänomens, Angabe der adäquaten bzw. einer inadäquaten Ursache, Nutzung der Modellvorstellung, ohne Antwort).

Durch das Design der Gesamtuntersuchung lassen sich im Sinne von Experimental- und Kontrollgruppen Klassen/Kinder unterscheiden, die mit bzw. ohne Video experimentierten und solche, die an den Stationen arbeiteten oder auch nicht.

4.2 Abhängige und unabhängige Variablen

Abhängige Variablen sind:

- 1) *Handlungsregulation*: Die *Kinder* wurden beim Experimentieren an den Stationen von Studierenden beobachtet, wobei die beobachteten Aktivitäten hinsichtlich nachfolgender Items der Eigenständigkeit der Handlungsregulation (Fragebogen) beurteilt wurden.
 - a) *Handlungsorientierung*: Arbeitsaufnahme (zielgerichtet=3, durch Nutzung des Forscher*innenheftes=2, instruiert=1, ohne=0); explizite Instruktion (ja=1, nein=2); Hilfe von Kindern (ohne/selbständig=3; durch Forscher*innenheft=2; durch Erwachsene=1);
 - b) *Handlungsausführung*: *Handlung* (selbstständig=4, miteinander=3, durch Nachmachen=2, durch Zusehen=1); Experimentieren (Experiment=3, Versuch=2, Exploration=1); Handlungsverlauf (sicher=4, sicher mit Instruktion=3, unsicher=2, unsicher mit Instruktion=1, ohne=0); Kooperation (allein/Paar=2, Gruppe=1, ohne=0); Hilfe (von Kindern/ohne=3; durch Forscher*innenheft=2, durch Erwachsene=1);
 - c) *Handlungskontrolle*: *Vorhandensein* (ja=1, nein=0); durch (Kinder=3, Forscher*innenheft=2, Erwachsene=1, ohne=0).

Die Prüfung der Reliabilität der Beobachtungsdaten (10 Items) ergab noch akzeptable Werte (Cronbachs Alpha zwischen 0,511 und 0,701), die stichprobenartig⁴ erhobene Interraterrelia-

⁴ Die Stichproben entstanden dadurch, dass von zwei oder mehreren Beobachter*innen das gleiche Kind beobachtet wurde, wenn zufällig die betreffende Station nur wenig besucht war. Eine systematische Analyse war aus Personalgründen nicht möglich.

bilität war leider nicht immer zufriedenstellend (Kappa zwischen 0,5 und 0,4). Die ermittelten Werte lassen sich daher nicht zu Messzwecken inhaltlich nutzen, sondern sind nur in ihrer relativen Höhe im Sinne eines mehr oder weniger eigenregulierten Handelns zu interpretieren.

2) *Posttestwert (Begriffsniveau)*: Die Antworten im Rahmen des Tests (Quiz mit in der Regel fünf Multiple-Choice-Aufgaben, z.B. Warum kann ein Insekt auf der Wasseroberfläche stehen, ein Vogel aber nicht?) wurden je nach Niveaustufe bepunktet (siehe Tab. 1).

begriffliches Niveau	Beispiel	Pkt.
sinnlich konkret (Benennung des Phänomens)	„...weil Vögel untergehen, Insekten aber nicht“	1
abstrakt	„...weil Wasser eine dünne Haut bildet, auf der Insekten gerade noch stehen können, Vögel aber nicht“	2
geistig-konkret (Erklärung des Phänomens durch Theorieausage)	„...weil die Wasserteilchen nur zum Wasser hin angezogen werden, bilden sie so etwas wie eine Haut, die Insekten trägt, Vögel aber nicht“	3
fake – Phänomen- und Theorieebene falsch verbunden [„wissenschaftlich klingender Unsinn“]	„...weil die Wasserteilchen eine Haut bilden“	0

Tab. 1: Punktbewertung der begrifflichen Niveaus

Da der Wissenstest für den Experimentiertag konstruiert war – 8 bis 10 nicht nur Niveau-Aufgaben pro Test –, existiert in der Regel nur eine Niveau-Aufgabe pro Station, was lediglich Trendaussagen gestattet, die näher untersucht werden müssten.

3) *Lerneffekt*: Dieser errechnet *sich* aus der Differenz zwischen Post- und Pretestwerten.

Unabhängige Variablen sind:

- *Besuch der Station*,
- *Nutzung des Erklär-Videos*,
- *Vorwissen* (Wert des Pretests – analog Posttest ermittelt),
- *Station* (bzw. Inhalt der Station),
- *Geschlecht und Klassenstufe* (3 bzw. 4).

5 Ergebnisse

Die *Stichprobe* (N=229 Kinder; 107 Mädchen und 122 Jungen) besteht aus 390 Datensätzen aus den Klassenstufen 3 (N=308) und 4 (N=82). 304 Beobachtungsdaten (Bobachtungsrates = 77,9%) und 355-360 (ca. 81%) Lerndaten liegen vor. Die Datensätze aus den Klassenstufen (3, 4) verteilen sich wie folgt auf die Stationen: Wasserhaut (164, 11); Smog (51, 71); Schwimmen und Sinken (93, 0). Damit fällt die Station Schwimmen und Sinken hinsichtlich der Analyse des Faktors Klassenstufe heraus und die Station Wasserhaut ist diesbezüglich mit Vorsicht zu analysieren. Videonutzer sind 77 Jungen und 60 Mädchen aus Klasse 3 (N=137), daher sind hier Aussagen zu Klasse 4 nicht möglich.

Mit den statistischen Analysen werden geprüft: a) der Vergleich von Pre- und Posttest, b) die Nullhypothesen bezogen auf Unterschiede bei den Variablen Handlungsregulation, Vorwissen, Posttest und Lerneffekt bezüglich der geprüften Faktoren, c) Zusammenhänge zwischen den Variablen, d) moderierende (Vorwissen*Videonutzung/ Station/ Geschlecht/ Klassenstufe) und mediatisierende Wirkungen auf Pretest → *Variable* → Posttest/Lerneffekt sowie Pretest → *Variable* → Handlungsregulation, wobei Handlungsregulation auch als unabhängige Variable (Moderator-Variable) geprüft wird.

Die Nullhypothesen werden mit nichtparametrischen Verfahren geprüft (aufgrund der fehlenden Normalverteilung der Daten), Korrelationsanalysen nach Spearman, Moderationsanalysen erfolgen mit linearer Regression (Residuen normalverteilt, ansonsten Bootstrapping) und Mediationsanalysen mit einem entsprechenden Verfahren JASP 0.14.1. Moderation bedeutet hier, dass der Effekt einer unabhängigen Variablen X auf eine abhängige Variable Y für verschiedene Werte einer dritten Moderator-Variable unterschiedlich ist. Mediation bedeutet, dass der Effekt einer Variablen X auf eine Variable Y (partiell) indirekt ist: er wird (auch) durch eine dritte Variable M verursacht.

5.1 Wie groß sind die Lerneffekte der Bearbeitung der Experimentierstationen?

In der Literatur wird über kleine Effekte berichtet, daher wurden im Projekt zusätzlich verbale Instruktion und Erklär-Videos (Multimedia) angeboten. Zu vermuten ist, dass Kinder, die an den Stationen gearbeitet hatten, höhere Lernwerte erreichen als jene, bei denen das nicht der Fall war. Die Ergebnisse können Tabelle 2 entnommen werden.

Station	N (Jungen/ Mädchen)	Pretest	Posttest	Lerneffekt
alle Stationen	390 (204/186)	1,49 (1,46/1,51)	1,44 (1,54/1,34)	-0,06 (0,07/-0,19)
bearbeitet	304 (163/141)	1,48 (1,45/1,50)	1,41 (1,48/1,34)	-0,08 (0,01/-0,17)
nicht bearbeitet	86 (41/45)	1,53 (1,50/1,55)	1,54 (1,75/1,32)	0,02 (0,29/-0,26)
Wasserhaut	175 (91/84)	1,54 (1,70/1,38)	1,27 (1,10/1,44)	-0,27 (-0,6/0,06)
bearbeitet	148 (76/72)	1,55 (1,43/1,66)	1,36 (1,46/1,25)	-0,19 (0,03/-0,40)
nicht bearbeitet	27 (15/12)	1,56 (1,20/1,92)	0,89 (1,40/0,38)	-0,67 (0,20/-1,54)
Smog	122 (64/58)	1,86 (1,86/1,86)	1,99 (1,82/2,15)	0,11 (-0,06/0,28)
bearbeitet	82 (44/38)	1,89 (1,88/1,89)	1,95 (2,15/1,74)	0,03 (0,23/-0,18)
nicht bearbeitet	40 (20/20)	1,82 (1,82/1,81)	2,07 (2,16/1,98)	0,28 (0,39/0,17)
Schwimmen	93 (49/44)	0,93 (0,77/1,08)	0,98 (1,07/0,89)	0,05 (0,29/-0,19)
bearbeitet	74 (43/31)	0,91 (1,07/0,74)	0,93 (0,83/1,02)	0,02 (-0,24/0,27)
nicht bearbeitet	19 (6/13)	1,01 (1,17/0,85)	1,26 (1,33/1,19)	0,26 (0,17/0,35)

Tab. 2: Lernwerte mit und ohne Bearbeitung der Stationen und insgesamt (gerundete Mittelwerte – der rechnerische Skalennitelpunkt beträgt bei den Tests 1,5)⁵

Die folgenden signifikanten Unterschiede konnten festgestellt werden. (Es kann allerdings nicht ausgeschlossen werden, dass der eigentlich bei identischen Tests zu erwartende Lerneffekt im Posttest durch ein etwas geringeres Konzentrationsvermögen am Ende des Experimentiertages negativ beeinflusst wurde.)

a) alle Stationen

Die Klasse 4 erreicht signifikant höhere Werte in Pre- ($z_{(1,272)}=-2,715$; $p=0,007$; $d=0,334$) und Posttest ($z_{(1,275)}=-3,772$; $p<0,001$; $d=0,467$). Bei Kindern 4. Klassen, die nicht an den Stationen waren, liegen die Posttestwerte ($z_{(1,85)}=-2,631$; $p=0,009$; $d=0,596$) und Lerneffekte ($z_{(1,83)}=-1,980$; $p=0,048$; $d=0,445$) signifikant höher als bei denen in Klasse 3 (beides nicht in der Tabelle).

b) Station Wasserhaut

Nur bei Mädchen liegen die Posttestwerte mit ($z_{(1,72)}=-2,049$; $p=0,040$; $d=0,498$) und ohne Stationsbesuch ($z_{(1,12)}=-2,585$; $p=0,010$; $d=2,242$) signifikant (Wilcoxon-Test) unter denen des Pretests (negativer Lerneffekt – siehe Tabelle 2). Mädchen, die die Station besuchten, erreichen aber signifikant (Mann-Whitney-Test) höhere Posttestwerte als jene, die das nicht taten ($z_{(1,84)}=-2,315$; $p=0,021$; $d=0,522$), daher unterscheiden sich die (wenn auch negativen) Lerneffekte beider Gruppen signifikant ($z_{(1,84)}=-2,359$; $p=0,018$; $d=0,533$). Jungen, die nicht an der

⁵ Um die Orientierung zu erleichtern, werden in den Tabellen einige deskriptive Daten kursiv hervorgehoben. Diese Daten werden bei den statistischen Berechnungen in den folgenden Absätzen verwendet.

Station waren, erreichen höhere Posttestwerte ($z_{(1,27)}=-2,55$; $p=0,012$, $d=1,126$) und erzielen einen höheren Lerneffekt als Mädchen ohne Stationsbesuch ($z_{(1,27)}=-2,673$; $p=0,008$, $d=1,2$).

c) Station Smog

In Klasse 4 liegen bei Kindern, die nicht an der Station waren, die Posttestwerte signifikant über denen des Pretests ($z_{(1,26)}=-1,802$; $p=0,072$; $d=0,756$ – nicht in der Tabelle).

Jungen erreichen bei Bearbeitung der Station signifikant höhere Posttestwerte als Mädchen ($z_{(1,82)}=-2,084$, $p=0,037$; $d=0,473$).

Regressionsanalysen erbrachten folgende Ergebnisse (Tab. 3):

Interaktion	abh. Variable	Statistik
alle Probanden mit Bearbeitung der Stationen		
Station * Vorwissen	Posttest/ Lerneffekt	$\Delta R^2=1,9\%$ / $\Delta R^2=1,4\%$ $f_{(1,268)}=5,713$; $p<0,018$; 95% CI[0,023; 0,237]
Klasse 3	Posttest/ Lerneffekt	$\Delta R^2=3,2\%$ / $\Delta R^2=2,3\%$ $f_{(1,214)}=7,773$; $p=0,006$; 95% CI[0,045; 0,261]
alle Probanden ohne Bearbeitung		
Gender * Vorwissen	Posttest/ Lerneffekt	$\Delta R^2=6,1\%$ / $\Delta R^2=4,1\%$ $f_{(1,79)}=5,567$; $p=0,021$; 95% CI[0,086; 1,016]
alle Jungen mit Bearbeitung		
Station * Vorwissen	Posttest/ Lerneffekt	$\Delta R^2=4,1\%$ / $\Delta R^2=3,2\%$; $f_{(1,143)}=7,187$; $p=0,008$; 95% CI[0,048; 0,318]
Jungen Klasse 3 mit Bearbeitung		
Station * Vorwissen	Posttest/ Lerneffekt	$\Delta R^2=6,5\%$ / $\Delta R^2=5,1\%$; $f_{(1,118)}=9,843$; $p=0,002$; 95% CI[0,078; 0,347].

Tab. 3: Festgestellte signifikante Interaktionen

Generell beeinflusst die Interaktion von Station und Vorwissen die Lerneffekte, was vor allem auf die Klasse 3 (geringeres Vorwissen) zurückgeht. Werden die Stationen nicht bearbeitet, so interagieren Gender und Vorwissen (bei Mädchen in der Summe größer) und beeinflussen die Lerneffekte. Nur bei Jungen, welche die Stationen bearbeitet hatten, beeinflusst die Interaktion von Station und Vorwissen die Lerneffekte, was besonders für die Klasse 3 charakteristisch ist (Smog negativ, ansonsten positiv). Die Ergebnisse der Mediationsanalyse (indirekte Effekte des Faktors *Klassenstufe*) sind der Tabelle 4 zu entnehmen.

direkter Effekt	indirekter Effekt	Statistik	totaler Effekt
Vorwissen → Posttest – alle Stationen und Probanden			
0,231	0,035	$z=2,406$; $p=0,016$	0,266
alle Stationen mit Bearbeitung			
0,245	0,035	$z=2,156$; $p=0,031$	0,279
Vorwissen → Lerneffekt – alle Stationen und Probanden			
-0,769	0,035	$z=2,406$; $p=0,016$	-0,734
alle Stationen mit Bearbeitung			
-0,755	0,035	$z=2,156$; $p=0,031$	-0,721

Tab. 4: Festgestellte signifikante indirekte Effekte des Faktors Klassenstufe

Für alle Stationen und alle Probanden ist eine signifikante indirekte Wirkung des Faktors Klassenstufe (höhere Lernwerte der Klasse 4) auf die Beziehung Vorwissen und Posttestwert beobachtbar. Hier wirkt sich auch der Faktor Stationsbesuch (Bearbeitung) entsprechend posi-

tiv aus. Allerdings ist der geringe Stichprobenumfang (Klasse 4) und die Tatsache zu beachten, dass diese Klassenstufe die inhaltlich schwierige Station Schwimmen nicht bearbeitete. Wegen der bei hohem Vorwissen durch die Niveaustufen begrenzten Lernwerte fällt die Analyse für die Beziehung Vorwissen und Lerneffekt hier negativ aus: Bei höherem Vorwissen kann weniger dazugelernt werden.

Erwartungsgemäß sind, entsprechend der Erkenntnislage zum Lernen an außerschulischen Lernorten (vgl. Alfieri et al. 2011; Graf & Naschka-Roos 2009; Roth et al. 1997; Lewalter & Geyer 2009; Itzek-Greulich et al. 2015), die Effekte der Bearbeitung der Stationen bezüglich des Vergleichs von Pre- und Posttest sowie der Testwerte mit und ohne Bearbeitung nicht signifikant. Es gibt lediglich einige differenzielle Effekte bezogen auf die Faktoren (Inhalt der Station, Gender und Klassenstufe, welche den zu erwartenden Lerneffekt bei identischem Pre- und Posttest bestätigen, zugleich aber auf überraschende Wissensunterschiede zwischen dritten und vierten Klassen (5.3) sowie auf negative Wirkungen bei Bearbeitung der Stationen verweisen.

Unsere Hypothese wird zwar hinsichtlich der Erkenntnislage aus der Literatur bestätigt, ist aber mit Blick auf die erwartete Instruktionwirkung zurückzuweisen: die Lerneffekte sind klein und unterscheiden sich nur gering, z.T. erwartungswidrig.

5.2 Unterstützen Multimedia-Präsentationen (Erklär-Videos) die Handlungsregulation?

Zu vermuten ist, dass zusätzliche Informationen über die Inhalte der Experimente handlungsunterstützend wirken können, wengleich die Erklär-Videos vorrangig auf die Vermutungsbildung und Erklärung der zu untersuchenden Phänomene ausgerichtet waren. Die Ergebnisse können Tabelle 5 entnommen werden:

Station	N (Jungen/Mädchen)	Handlungsregulation
alle Stationen	304 (163/141)	1,70 (1,69/1,70)
mit Video	137 (77/60)	1,73 (1,72/1,74)
ohne Video	167 (86/81)	1,67 (1,67/1,67)
Wasserhaut	148 (76/72)	1,70 (1,69/1,71)
mit Video	63 (34/29)	1,81 (1,81/1,81)
ohne Video	85 (42/43)	1,62 (1,59/1,65)
Smog	82 (64/58)	1,71 (1,71/1,71)
mit Video	29 (16/13)	1,67 (1,63/1,70)
ohne Video	53 (28/25)	1,74 (1,76/1,72)
Schwimmen	74 (43/31)	1,67 (1,67/1,66)
mit Video	45 (27/18)	1,66 (1,66/1,66)
ohne Video	29 (16/13)	1,67 (1,69/1,65)

Tab. 5: Werte der Handlungsregulation mit und ohne Videonutzung (Mittelwerte – der rechnerische Skalenmittelpunkt beträgt 1,65)

Die Gruppen mit und ohne Videonutzung unterscheiden sich hinsichtlich der Handlungsregulation nicht signifikant. Die Videonutzung hilft insgesamt nicht dabei, selbständiger zu experimentieren. Das gilt sowohl im Vergleich der Geschlechtergruppen als auch für Mädchen und Jungen und die Klassenstufe 3 (4 arbeitet nur ohne Videonutzung) und im Vergleich der Stationen, außer bei Wasserhaut: hier unterstützt die Videonutzung die Handlungsregulation ($z_{(1,148)}=-2,611$; $p=0,009$; $d=0,439$), da erstens in der Präsentation ein handlungsanleitendes Video eingebaut war und zweitens eine starke Orientierung auf Hands-on-Aktivitäten beobachtet wurde (wettbewerbsartige Versuche, die Büroklammer nicht untergehen zu lassen). Daher lässt die Interaktion zwischen Station und Videonutzung die Handlungsregulation signifikant voraussagen: $\Delta R^2=1,3\%$; $f_{(1,300)}=4,013$; $p=0,046$; 95% CI[0,133; -0,001], was wegfällt, wenn die Station Wasserhaut aus der Analyse ausgeschlossen wird. Indirekte Effekte

(Mediation) der Videonutzung auf die Beziehung zwischen Vorwissen/Station und Handlungsregulation konnten nicht festgestellt werden.

Der Faktor Video hat insgesamt, wie vermutet, keinen nennenswerten Einfluss auf die Handlungsregulation, interessanterweise – mit Ausnahme der Station Wasserhaut – auch nicht der Inhalt der hier untersuchten Stationen.

5.3 Unterstützen Erklär-Videos die Lernergebnisse im Sinne der Orientierung auf ein angemessenes Niveau der begrifflichen Repräsentation von Erklärungen untersuchter Phänomene beim Experimentieren und welche Rolle spielt der Inhalt der Station dabei?

Die Erklär-Videos sollten die instruktionale Wirkung verbessern, indem die Zusammenhänge noch einmal interaktiv und multimedial gestützt veranschaulicht und erklärt werden. Lernergebnisse beziehen sich auf die Posttestwerte und den Lerneffekt, da Posttestwerte allein keinen Lernzuwachs verdeutlichen. Zu beachten ist aber, dass auch der Lerneffekt nur beschränkt aussagefähig ist, denn bei hohem Vorwissen können die Posttestwerte (wegen der drei Niveaustufen) nicht weiter steigen. Da die Wissensentwicklung vom Inhalt abhängt (domänenspezifisch verläuft), sollten sich stationsabhängige Unterschiede zeigen. Die Ergebnisse können Tabelle 6 entnommen werden.

Station	N	Pretest	Posttest	Lerneffekt
alle Stationen	304	1,49	1,44	-0,06
mit Video	137	1,34	1,09	-0,25
ohne Video	167	1,62	1,73	0,09
Wasserhaut	148	1,54	1,27	-0,27
mit Video	63	1,39	1,26	-0,23
ohne Video	85	1,71	1,46	-0,25
Smog	82	1,86	1,99	0,11
mit Video	29	1,88	1,60	-0,29
ohne Video	53	1,89	2,13	0,20
Schwimmen	74	0,93	0,98	0,05
mit Video	45	0,88	0,52	-0,36
ohne Video	29	0,95	1,51	0,57

Tab. 6: Lernwerte an den Stationen und insgesamt mit/ohne Videonutzung (gerundete Mittelwerte; Skalenmittelpunkt bei Pre- und Posttest ist 1,5)

Die Posttestwerte sind insgesamt (alle Stationen) bei Videonutzung tendenziell kleiner als die im Pretest ($z_{(1,137)}=-1,869$; $p=0,062$; $d=0,324$); bei der Station Schwimmen ist der Unterschied signifikant ($z_{(1,45)}=-2,035$; $p=0,042$; $d=0,637$).

Bei Videonutzung liegen die Posttestwerte insgesamt (alle Stationen) signifikant unter denen ohne Videonutzung ($z_{(1,304)}=-4,171$; $p<0,001$; $d=0,493$), was sich auch für den Lerneffekt zeigt ($z_{(1,304)}=-2,321$; $p=0,02$; $d=0,269$). Analoge Beobachtungen wurden auch bei den einzelnen Stationen (außer Wasserhaut) gemacht: für Smog beim Posttestwert ($z_{(1,82)}=-2,309$; $p=0,021$; $d=0,527$); für Schwimmen beim Posttestwert ($z_{(1,74)}=-3,732$; $p<0,001$; $d=0,963$) und Lerneffekt ($z_{(1,74)}=-2,522$; $p=0,012$; $d=0,613$).

Videonutzung führt relativ unabhängig vom Vorwissen zu negativen Lerneffekten, es konnte aber keine signifikante Moderation (Interaktionseffekt von Videonutzung und Vorwissen hinsichtlich der Lernwerte) oder Mediation (knapp n.s.) des Faktors Videonutzung hinsichtlich der Beziehung zwischen Vorwissen und Posttestwert/Lerneffekt festgestellt werden. Ohne Videonutzung lässt die Interaktion zwischen Vorwissen und (Inhalt der) Station (positive Lerneffekte bei Schwimmen und Smog) die Lernwerte signifikant voraussagen – Posttestwerte: $\Delta R^2=3,2\%$; Lerneffekt: $\Delta R^2=2,2\%$; $f_{(1,131)}=4,531$ $p=0,035$; 95% CI[0,012; 0,331].

Vergleicht man die Stationen untereinander, so zeigen sich einige Unterschiede, die auf die Bedeutung des Inhalts der Stationen hinweisen: Ohne Videonutzung liegen bei Smog die

Posttestwerte signifikant über denen bei Wasserhaut ($z_{(1,138)}=-2,663$; $p=0,008$; $d=0,465$), was auch für die Lerneffekte gilt ($z_{(1,138)}=-2,463$; $p=0,014$; $d=0,429$). Mit Videonutzung ist der Posttestwert bei Wasserhaut signifikant größer als bei Schwimmen ($z_{(1,108)}=-2,986$; $p=0,003$; $d=0,6$), was ohne Videonutzung für den Lerneffekt genau umgekehrt ist ($z_{(1,114)}=-2,156$; $p=0,031$; $d=0,412$). Mit Videonutzung liegen die Posttestwerte bei Smog über denen bei Schwimmen ($z_{(1,74)}=-4,016$; $p=0,002$; $d=1,056$), was auch ohne Videonutzung zu beobachten war ($z_{(1,82)}=-2,583$; $p=0,010$; $d=0,595$). Diese Vergleiche unterstreichen insgesamt – im Gegensatz zum Faktor Handlungsregulation – die Bedeutung des Inhalts der Stationen für den Lerneffekt.

Die nachfolgend dargestellten Ergebnisse der Interviewdaten illustrieren, dass teilweise instruktionsabhängige Wissens Elemente vorhanden sind, diese jedoch abstrakt verbleiben und eine Integration in das Vorwissen nicht gelingt.

Bei der Station *Wasserhaut* argumentieren einige Kinder in den Interviews auf abstraktem Niveau mit dem Teilchenbegriff. Aus den 73 Interviewdaten (inklusive weniger Doppelbenennungen) konnten vier wesentliche Kategorien extrahiert werden: Benennung des Phänomens, Wasserhaut, Schwimmen und Teilchenmodell. 13 Kinder benennen das Phänomen (*Weil wenn man dieses Spülmittel nimmt, dann geht sie nämlich unter, weil das so anzieht. Wenn nicht, wenn man das so ganz kurz mit den Händen anstupst, man darf aber nicht ans Wasser gehen, dann bleibt das da auf dem Wasser schwimmen.*); 26 geben die Wasserhaut als Ursache an (*Weil die Wasserhaut sie (Büroklammer) trägt, wie ein Insekt.*); 9 das Schwimmen (als Fehlinterpretation) (*Weil die Büroklammer leichter ist als das Wasser, schwimmt sie.*); 12 verbinden die Erklärung mit der Modellebene (*Dass die 'ne Haut hat und die Teilchen sich gegenseitig anziehen. Weil da so eine ganz kleine Haut ist, und wenn man die Büroklammer langsam rauf macht, kann sich das Wasser an das Gewicht gewöhnen.*). 15 Kinder geben keine Antwort.

Beim Transfer (Erklärung der Reinigung der Wasseroberfläche mit einem Kescher) wird es noch schwieriger. In den Antworten tauchen auf: 5-mal Wasserhaut (*Vielleicht weil da auch 'ne Wasserhaut ist und das Blatt genauso wie die Büroklammer, die Teilchen [die der Bademeister rausholt] könnten ja genauso liegen wie die Büroklammer*); 4-mal Schwimmen und 3-mal Teilchen (*Weil die Teilchen dann zusammenbleiben, bilden sie eine Wasserhaut.*), in 61 Fällen gibt es keine Antwort.

Die Videonutzung verfolgte das Ziel, die bekannten und weiter oben demonstrierten geringen Effekte des Lernens an außerschulischen Lernorten zu steigern. Die Vergleiche der Pre-Posttests sowie der Testwerte mit und ohne Videonutzung weisen jedoch darauf hin, dass die Videonutzung überwiegend zu negativen Lerneffekten führt. Offenbar hat sie die Kinder nicht dabei unterstützt, neues – u.a. Modellwissen – konsistent in die vorhandenen Wissensstrukturen einzubauen oder diese sinnstiftend zu verändern. Damit ist unsere Hypothese zurückzuweisen. Allerdings müsste genauer untersucht werden (z.B. im Rahmen von Interviews als Pre- und Posttest), inwiefern diese Verwirrung einen Schritt des Conceptual Change darstellt und wie man diesen ggf. (z.B. durch Unterricht) produktiv machen könnte. Ferner ist zu prüfen, welche über die aus der Literatur bekannten, die medienspezifischen Informationsverarbeitungsprozesse betreffenden Konstruktionsmerkmale für multimediale Präsentationen bzw. Erklär-Videos hinausgehenden lernprozessbezogenen Merkmale (Sinnbildung, bewusstseinspflichtige mentale Konstruktionsprozesse (Lerntätigkeit)) stärker Beachtung finden müssen.

5.4 Welche Rolle spielt das Vorwissen mit Blick auf die Handlungsregulation, die Lernergebnisse und die Nutzung der Erklär-Videos?

Vorwissen dürfte die Handlungsregulation als auch die Lernresultate unterstützen. Kinder mit mehr Vorwissen dürften ggf. dazu tendieren, auf die Videonutzung zu verzichten, sich diesen Zwischenschritt zu ersparen, um schnell mit den Hands-on Aktivitäten beginnen zu können.

Station	Stichprobe	r ⁶
alle	alle	-0,166**
Wasserhaut	mit Videonutzung	-0,290*

Tab. 7: Korrelationsanalyse Vorwissen und Handlungsregulation

Aus Tabelle 7 wird ersichtlich, dass höheres Vorwissen nicht mit höherer Handlungsregulation zusammenzuhängen scheint, denn diese spiegelt vor allem Hands-on-Aktivitäten wider.

Variablen	Station	Stichprobe	r
Pre- * Posttest	alle	alle	0,254**
		Videonutzung	0,242**
		ohne Videonutzung	0,230**
Pre- * Lerneffekt	alle	alle	-0,525***
		Videonutzung	-0,595***
		ohne Videonutzung	-0,514***

Tab. 8: Korrelationsanalyse Vorwissen und Posttestwert/Lerneffekt

Für die gesamte sowie die Teilstichproben (Videonutzung) korrelieren Vorwissen und Posttestwert schwach, aber signifikant positiv. Vorwissen und Lerneffekt korrelieren wegen der Testkonstruktion negativ (bei höherem Vorwissen ist wegen der gedeckelten Niveaustufen der Lerneffekt geringer). Ein höheres Vorwissen hängt erwartungsgemäß mit höheren Lernergebnissen zusammen.

Videonutzer verfügen bezogen auf alle Stationen über ein knapp nicht signifikant geringeres Vorwissen als jene, die das Video nicht nutzen ($z_{(1,304)}=-1,899$; $p=0,058$; $d=0,219$ – vgl. Tab. 6). Kinder ohne Videonutzung verfügen bei der Station Wasserhaut über ein signifikant höheres Vorwissen als bei Schwimmen ($z_{(1,114)}=-2,695$; $p=0,007$, $d=0,552$); was auch für die Station Smog für Kinder mit ($z_{(1,74)}=-3,862$; $p<0,001$; $d=1,005$) und ohne Videonutzung gilt ($z_{(1,82)}=-3,704$; $p<0,001$; $d=0,897$). Zum Inhalt der Station Schwimmen ist das Vorwissen sehr gering.

Festzustellen ist, dass das Vorwissen sich in Abhängigkeit vom Inhalt der Stationen unterscheidet und positiv mit den Posttestwerten korreliert. Videonutzer verfügen über ein tendenziell geringeres Vorwissen und nutzen ggf. daher das Video.

Signifikante Moderationseffekte der Interaktion von Videonutzung mit den unabhängigen Variablen (außer Klassenstufe, da Klasse 4 nur ohne Videonutzung arbeitete) hinsichtlich der Lernwerte bzw. Handlungsregulation waren nicht festzustellen. Eine Mediationsanalyse deutete darauf hin, dass die Videonutzung sich nicht signifikant auf die Beziehung zwischen Vorwissen und Handlungsorientierung (siehe 5.5) sowie nicht signifikant ($p=0,076$) auf die zwischen Vorwissen und Lernwerten auswirkt.

Unsere Hypothese trifft nur teilweise zu, denn Vorwissen hängt zwar mit den Lernwerten zusammen, aber nicht signifikant mit der Handlungsregulation und nur tendenziell negativ mit der Videonutzung. Videonutzer verfügen auch nur tendenziell über ein geringeres Vorwissen.

⁶ * $p=0,05$; ** $p=0,01$; *** $p<0,001$

5.5 Wirkt die Handlungsregulation indirekt auf die Beziehung zwischen Vorwissen und Lernwerten oder moderiert sie letztere über die Wechselwirkung mit den anderen unabhängigen Variablen?

Handlungsregulation kann als unabhängige Variable und zugleich abhängige Variable angesehen werden. Die Handlungsregulation könnten als abhängige Variable durch (den Inhalt der) Stationen und Vorwissen beeinflusst werden, aber auch eine wichtige Bedingung für den Lernerfolg sein, daher als unabhängige Variable wirken.

Ergebnis:

Bildet man aus den Werten der Handlungsregulation anhand positiver und negativer z-Werte zwei Gruppen, so kann mit dem Mann-Whitney-Test geprüft werden, ob diese Gruppen sich bezüglich der Lernwerte signifikant unterscheiden. Das ist weder für die Gruppe mit bzw. ohne Videonutzung der Fall. Handlungsregulation weist mit und ohne Videonutzung keine signifikante indirekte Wirkung auf die Beziehung zwischen Vorwissen und Posttestwert bzw. Lerneffekt auf. Auch moderierende Effekte der Interaktion von Handlungsregulation mit den Faktoren Vorwissen, Station, Gender und Klassenstufe sind weder hinsichtlich der Posttestwerte noch der Lerneffekte bei beiden Gruppen festzustellen. Die Handlungsregulation ist daher nicht als unabhängige Variable anzusehen. Die Analysen zur Handlungsregulation als abhängige Variable sind weiter oben dargestellt worden.

Die oben genannte Hypothese bezüglich der Handlungsregulation als unabhängige Variable ist zurückzuweisen.

5.6 Gibt es differenzielle Effekte bezüglich der Videonutzung hinsichtlich der Handlungsregulation und der Lernwerte bzgl. der Faktoren Gender, Klassenstufe und (Inhalt der) Station?

Station			Pretest		Posttest		Lerneffekt	
	Jungen	Mädchen	Jungen	Mädchen	Jungen	Mädchen	Jungen	Mädchen
alle Stationen	163	141	1,46	1,51	1,54/	1,34	0,07	-0,19
mit Video	77	60	1,35	1,32	1,22	0,95	-0,13	-0,37
ohne Video	86	81	1,56	1,67	1,76	1,70	0,16	0,02
Wasserhaut	76	72	1,70	1,38	1,10	1,44	-0,6	0,06
mit Video	34	29	1,37	1,41	1,46	1,05	0,09	-0,36
ohne Video	42	43	1,50	1,91	1,46	1,46	-0,04	-0,45
Smog	44	38	1,86	1,86	1,82	2,15	-0,06	0,28
mit Video	16	13	1,84	1,92	1,81	1,38	-0,03	-0,54
ohne Video	28	25	1,90/	1,87	2,34	1,92	0,38	0,02
Schwimmen	43	31	0,77	1,08	1,07	0,89	0,29	-0,19
mit Video	27	18	1,04	0,72	0,57	0,47	-0,46	-0,25
ohne Video	16	13	1,13	0,77	1,25	1,77	0,13	1,00

Tab. 9: Lernwerte von Mädchen und Jungen an den Stationen und insgesamt mit/ohne Videonutzung (gerundete Mittelwerte; Skalenmittelpunkt bei Pre- und Posttest ist 1,5)

Mädchen und Jungen zeigen in der Regel unterschiedliche Präferenzen bei Inhalten und der Art und Weise des Lernens (z.B. bevorzugen Mädchen tendenziell Stationen, an denen gebastelt oder sorgfältig beobachtet wird, Jungen solche, an denen ein Wettbewerbscharakter auszumachen ist; vgl. Giest 2022a, b), daher könnten auch bei den Gruppen mit und ohne Videonutzung Unterschiede bei Handlungsregulation und Wissenswerten zwischen den Geschlechtern beobachtbar sein, die je nach Station variieren, wenngleich diese in den untersuchten Klassenstufen eher relativ gering ausfallen dürften, was auch für die Klassenstufen selbst zutreffen sollte.

Die Ergebnisse stellen sich wie folgt dar: Der Vergleich der *Handlungsregulation* (mit vs. ohne Videonutzung – siehe auch 5.2, Tab. 5) zeigt, dass Jungen mit Videonutzung bei der Station Wasserhaut signifikant höhere Werte bei der Handlungsregulation erreichen als solche ohne ($z_{(76)}=-2,319$; $p=0,02$; $d=0,552$). Da die Klassenstufe 4 nur ohne Videonutzung gearbeitet hat, kann dieser Faktor nur für die Gruppe ohne Videonutzung betrachtet werden: hier liegt die Klasse 4 knapp n.s. vor Klasse 3 ($z_{(1,167)}=-1,853$; $p=0,064$; $d=0,290$).

Im Vergleich der Geschlechtergruppen und Klassenstufen sind die Unterschiede bei der Handlungsregulation mit und ohne Videonutzung insgesamt nicht signifikant, was im Wesentlichen auch auf Mädchen und Jungen untereinander zutrifft. Bei den Stationen gibt es weder mit, noch ohne Videonutzung signifikante Unterschiede zwischen den Geschlechtern. Lediglich bei Wasserhaut wirkt sich bei Jungen die handlungsanleitende Wirkung des Videos signifikant positiv aus.

Vom Faktor Gender gehen bei den Gruppen mit und ohne Videonutzung keine signifikanten indirekten Wirkungen auf die Beziehung zwischen Vorwissen und Handlungsregulation aus. Diese Aussagen gelten auch für jede einzelne Station. Ein differenzierter Vergleich der Klassenstufen ist hier nicht sinnvoll (s.o).

Bezüglich der *Wissenswertes* (Pre- und Posttest/Lerneffekt, vgl. Tab. 9) wurden folgende Ergebnisse ermittelt.

Der Vergleich von Pre- und Posttestwerten (Wilcoxon-Test) zeigt einen signifikanten Abfall der Posttestwerte bei *Videonutzung* (es sind nur Daten nur für die Klasse 3 verfügbar) ($z_{(1,45)}=-2,035$; $p=0,042$; $d=0,637$ – nicht in der Tabelle).

Ohne Videonutzung ist dies nur bei der Station Wasserhaut der Fall, bei Mädchen sind die Posttestwerte sogar signifikant kleiner als die im Pretest ($z_{(1,43)}=-2,040$; $p=0,041$; $d=0,655$).

Testet man die unabhängigen Stichproben bezüglich des Faktors *Videonutzung* (Mann-Whitney-Test) so weisen videonutzende Mädchen gegenüber solchen ohne Videonutzung bei *allen Stationen* ein tendenziell geringeres Vorwissen ($z_{(1,125)}=-1,842$; $p=0,065$; $d=0,314$) und signifikant niedrigere Posttestwerte ($z_{(1,126)}=-3,441$; $p<0,001$; $d=0,606$) sowie Lerneffekte auf ($z_{(1,125)}=-1,919$; $p=0,055$, $d=0,335$). Für Jungen trifft dies auf die Posttestwerte zu ($z_{(1,149)}=-2,603$; $p=0,009$; $d=0,417$), genauso wie für die Klasse 3 ($z_{(1,137)}=-2,535$; $p=0,011$; $d=0,444$ – nicht in der Tabelle).

Bezüglich der einzelnen *Stationen* war zu beobachten, dass auch hier Videonutzende signifikant niedrigere Wissenswerte erreichen: bei Wasserhaut Mädchen im Vorwissen ($z_{(1,72)}=-1,749$; $p=0,080$; $d=0,421$); bei Smog Jungen im Posttest ($z_{(1,38)}=-2,016$; $p=0,044$; $d=0,692$) sowie bei dieser Station die Klasse 3 bei Posttest ($z_{(1,37)}=-2,753$; $p=0,006$; $d=1,015$) sowie Lerneffekt ($z_{(1,37)}=-2,199$; $p=0,029$; $d=0,775$). Analoge Resultate erbrachte die Testung bei der Station Schwimmen für Mädchen im Posttest ($z_{(1,31)}=-3,120$; $p=0,002$; $d=1,353$), Lerneffekt ($z_{(1,31)}=-2,592$; $p=0,010$; $d=1,052$) sowie bei Jungen im Posttest ($z_{(1,43)}=-2,137$; $p=0,033$; $d=0,689$) und bei der Klasse 3 in Posttest ($z_{(1,74)}=-3,732$; $p<0,001$; $d=0,963$) und Lerneffekt ($z_{(1,74)}=-2,522$; $p=0,012$; $d=0,613$).

Die Klasse 4 erreicht ohne Videonutzung insgesamt (alle Stationen) signifikant höhere Posttestwerte ($z_{(1,148)}=-1,972$; $p=0,049$; $d=0,344$), lernt also hinzu.

Nur bei Videonutzung ist für alle Stationen ein Moderationseffekt der Interaktion von Gender und Vorwissen hinsichtlich der Vorhersage der Lernwerte festzustellen Posttestwert: $\Delta R^2=3,5\%$ /Lerneffekt: $\Delta R^2=2,3\%$; $f_{(1,133)}=5,144$; $p=0,025$ 95% CI[0,049; 0,727]). Bei Jungen und Mädchen konnten für alle Stationen keine Moderationseffekte der Interaktion zwischen Videonutzung und Vorwissen hinsichtlich der Posttestwerte festgestellt werden. Nur Wasserhaut macht hier eine Ausnahme (Posttest: $\Delta R^2=17,2\%$ /Lerneffekt: $\Delta R^2=12\%$; $f_{(1,59)}=14,218$; $p<0,001$; 95% CI[0,402; 1,310]). Bei Mädchen lässt die Interaktion von Videonutzung und Vorwissen die Lernwerte vorhersagen (Posttest: $\Delta R^2=29,9\%$ /Lerneffekt: $\Delta R^2=14\%$; $f_{(1,53)}=14,328$; $p<0,001$; 95% CI[0,470; 1,529]).

Die Mediationsanalyse bei Mädchen und Jungen bleibt hinsichtlich indirekter Wirkungen der Videonutzung auf die Beziehung zwischen Vorwissen und Lernwerten unauffällig. Für die Klassenstufe 3 sind die entsprechenden Moderations- und Mediationsanalysen unauffällig. Klasse 4 fällt wegen der Untersuchungsbedingungen heraus (s.o.).

Unsere Hypothese ist nur teilweise anzunehmen, da erstens die (negative) Wirkung der Videonutzung auf den Lerneffekt sich dominant bei den untersuchten Gruppen widerspiegelt und kaum durch die hier untersuchten Faktoren moderiert wird und zweitens sich die Stationen bezüglich des Wissens und der Lerneffekte deutlich unterscheiden.

6 Fazit

Sowohl die Untersuchungsbedingungen (Feldversuch unter nur eingeschränkt kontrollierbaren Praxisbedingungen) als auch die Limitationen der Messinstrumente (geringe Reliabilität der Beobachtungen) geben dazu Anlass, die Ergebnisse mit Vorsicht zu interpretieren.

Dennoch konnte wahrscheinlich gemacht werden, dass die von uns an den drei untersuchten Stationen eingesetzten Erklär-Videos insgesamt mit einem negativen Lerneffekt zusammenhängen. Nicht erwartungswidrig ist, dass bei Kindern ohne Videonutzung bei der Wiederholung eines mit dem Pretest identischen Posttests (mit gleichen Aufgaben) teilweise ein gewisser Lerneffekt zu beobachten war. Dass dieser im Falle der Video-Gruppe ausfällt bzw. negativ wird, kann als (erwartungswidrige) Wirkung der Erklär-Videos interpretiert werden. Die Ursachen dafür sind wahrscheinlich weniger in der äußeren Präsentation, also in der medialen Qualität der Erklär-Videos zu suchen, sondern in der zu geringen Berücksichtigung von Besonderheiten des Lernens (vgl. Lowe & Boucheix 2016).

Zu beachten ist ferner, dass die Erklär-Videos aus einer traditionellen Sichtweise auf naturwissenschaftliches Lernen (vordergründig gerichtet auf den Erwerb fachlichen Wissens) heraus eingesetzt wurden und diesbezüglich die in sie gesetzten Erwartungen nicht erfüllten. Aufzuklären bleibt unter einer solchen Sichtweise, ob die ausgebliebenen Lerneffekte Ausdruck einer produktiven Verunsicherung (im Sinne der Erschütterung des Vorwissens und Anbahnung eines Conceptual Change) sind oder ob die Inhalte der Erklär-Videos einfach nicht die Zone der nächsten Entwicklung der Kinder treffen. Eine Erschütterung des Vorwissens würde ein notwendiger Schritt zum Aufbau neuen Wissen (im Sinne des Conceptual Change) darstellen. Die hier ablaufenden Lernprozesse müssten jedoch (im Rahmen von experimentellen Studien) detaillierter untersucht und bei der Konzipierung der Erklär-Videos berücksichtigt werden.

Da Erklär-Videos grundsätzlich auf der Anwendung des Zeigens basieren, ist zu beachten, dass gut gezeigt werden kann, was an sichtbare und damit visualisierbare Aktivitäten gebunden ist. Insofern wäre die Frage zu stellen, ob der Einsatz von Erklär-Videos in dieser Hinsicht produktiv und lernwirksam sein könnte.

Zu beachten ist ferner, dass Minds-on-Aktivitäten, d.h. innere, mentale Konstruktionsaktivität schwer oder gar nicht sichtbar zu machen sind. Erklär-Videos können hier nur entsprechende Aktivitäten (z.B. zur Ko-Konstruktion) anregen. Dazu ist eine Integration in Unterrichtssituationen bzw. ein sich direkt an die User*innen wendende*r (sprechende*r und gestikulierende*r) Instrukteur*in erforderlich, außerdem sind Vorübungen bezüglich der Schlüsselkonzepte und ggf. (bei Beachtung der Besonderheiten der Altersgruppe) Prompts für die Anwendung bestimmter Lernstrategien einzubauen.

Die Untersuchung verweist daher darauf, dass bei der Entwicklung von Lernhilfen in Form von Erklär-Videos die Adaptivität (Anpassung an Vorwissen und Handlungsbedingungen an den Stationen) sowie die Integration in den Unterricht wesentlich zu verbessern sind.

Bei der Interpretation und Bewertung der Untersuchung ist nochmals zu betonen, dass es sich nicht um eine sauber planbare Laborstudie handelt, sondern das eigentliche Ziel der Projekte in der Entwicklung und Evaluation der Experimentierstationen bestand. Diesem Ziel

mussten sich die wissenschaftlichen Begleituntersuchungen unterordnen, was die saubere Variablenkontrolle und Randomisierung behinderte. Insofern sind die hier mitgeteilten Ergebnisse als Hinweise auf das Erfordernis einer sauberen wissenschaftlichen Prüfung der Wirkbedingungen von lernunterstützenden Maßnahmen, hier Erklär-Videos, zu interpretieren (Bétrancourta & Benetos 2018).

Literatur

- Alfieri, Louis, Brooks, Patricia J., Aldrich, Naomi J. & Tenenbaum, Harriet R. (2011): Does discovery-based instruction enhance learning? In: *Journal of Educational Psychology*, 103, pp. 1-18. doi:10.1037/a0021017.
- Anders, Yvonne, Hardy, Ilonca, Pauen, Sabina, Ramseger, Jörg, Sodian, Beate & Steffensky, Mirjam (2013): *Wissenschaftliche Untersuchungen zur Arbeit der Stiftung „Haus der kleinen Forscher“*. Band 5, herausgegeben von der Stiftung Haus der kleinen Forscher. Schaffhausen.
- Becker, Sebastian, Klein, Pascal, Gößling, Alexander & Kuhn, Jochen (2020): Using mobile devices to enhance inquiry-based learning processes. In: *Learning and Instruction*, 69, pp. 1-14.
- Bétrancourta, Mireille & Benetos, Kalliopi (2018): Why and when does instructional video facilitate learning? A commentary to the special issue “developments and trends in learning with instructional video”. In: *Computers in Human Behavior*, 89, pp. 471-475.
- Böschl, Florian, Gogolin, Sarah, Lange-Schubert, Kim & Hartinger, Andreas (2018): Modellverstehen von Grundschüler/innen in Abhängigkeit von Kontext und Kompetenzniveau In: Ute Franz, Hartmut Giest, Andreas Hartinger, Anja Heinrich-Dönges & Bernd Reinthoffer (Hrsg.): *Handeln im Sachunterricht*. Bad Heilbrunn, S. 17-24.
- Böschl, Florian, Gogolin, Sarah, Lange-Schubert, Kim und Hartinger, Andreas (2019): Mixed-Methods-Design als Grundlage für die Entwicklung von Aufgaben zur Erfassung von Modellkompetenz von Grundschulkindern. In: Hartmut Giest, Eva Gläser & Andreas Hartinger (Hrsg.): *Methodologien der Forschungen zur Didaktik des Sachunterrichts*. Bad Heilbrunn, S. 115-138.
- Bohrmann, Mareike, Todorova, Maria & Möller, Kornelia (2017): Welcher Magnet ist stärker? – Förderung des Verständnisses der Variablenkontrolle im Sachunterricht der dritten Klasse. In: Hartmut Giest, Andreas Hartinger & Sandra Tänzer (Hrsg.): *Vielperspektivität im Sachunterricht*. Bad Heilbrunn, S. 91-99.
- Brod, Garvin (2021): Generative Learning: Which Strategies for What Age? In: *Educational Psychology Review*, 33, pp. 1295-1318. <https://doi.org/10.1007/s10648-020-09571-9>.
- Chuang, Ying-Ting (2015): SSCLS: A Smartphone-Supported Collaborative Learning System. In: *Telematics and Informatics*, 32, 3, pp. 463-474. <http://dx.doi.org/10.1016/j.tele.2014.10.004>.
- Edelsbrunner, Peter A., Schalk, Lennart, Schumacher, Ralph & Stern, Elsbeth (2018): Variable control and conceptual change. A large-scale quantitative study in elementary school. In: *Learning and Individual Differences*, 66, pp. 38-53.
- Elschenbroich, Donata (2005): *Weltwunder. Kinder als Naturforscher*. München.
- Eschrich, Ulrike (2019): Die Inszenierung sachunterrichtlicher Experimentiersituationen durch Studierende des Grundschullehramts In: Martina Knörzer, Lars Förster, Ute Franz & Andreas Hartinger (Hrsg.): *Forschendes Lernen im Sachunterricht*. Bad Heilbrunn, S. 193-200.
- Evaluationsbericht (2016): Erstellt von Giest, Hartmut, Egbert, Björn, Heiden, Saskia, Lachmann, Franziska & Wandel, Katja. Unveröffentlichtes Material. Potsdam.
- Findeisen, Stefanie, Horn, Sebastian & Seifried, Jürgen (2019): Lernen Durch Videos – Empirische Befunde Zur Gestaltung Von Erklärvideos. In: *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, (Oktober), S. 16-36. <https://doi.org/10.21240/mpaed/00/2019.10.01.X>.
- Fiorella, Logan & Mayer, Richard E. (2018): What works and doesn't work with instructional video. In: *Computers in Human Behavior*, 89, pp. 465-470. DOI: 10.1016/j.chb.2018.07.015.
- Freericks, Renate (2011): Außerschulische Lernorte: Typologie und Entwicklungsstand. In: Renate Freericks & Dieter Brinkmann (Hrsg.): *Zukunftsfähige Freizeit. Analysen – Perspektiven – Projekte*. 1. Bremer Freizeitkongress Hochschule Bremen; Dokumentation der Fachtagung 12./13. November 2010. Bremen, S. 11-22.
- Gaubitz, Sarah (2021): Sache und Sprache in Erklärvideos für den Sachunterricht – Möglichkeiten, Grenzen und Bedingungen In: Ute Franz, Hartmut Giest, Melanie Haltenberger, Andreas Hartinger, Julia Kantreiter & Kerstin Michalik (Hrsg.): *Sache und Sprache*. Bad Heilbrunn, S. 133-141.
- Gesellschaft für Didaktik des Sachunterrichts (GDSU) (Hrsg., 2013): *Perspektivrahmen Sachunterricht*. Vollst. überarb. u. erw. Ausg. Bad Heilbrunn.
- Gerard, Libby, Matuk, Camilla, McElhaney, Kevin W. & Linn, Marcia C. (2015): Automated, adaptive guidance for K-12 education. In: *Educational Research Review*, 15, pp. 41-58.

- Gervé, Friedrich (2016): ICT im Sachunterricht – Impulse für Forschung und Entwicklung. In: Markus Peschel (Hrsg.): *Mediales Lernen*. Baltmannsweiler, S. 35-52.
- Giest, Hartmut (2020): *Vorlesungen über Didaktik des Sachunterrichts. Ein Beitrag zur Konkretisierung kulturhistorischer Didaktik*. Berlin.
- Giest, Hartmut (2022a): *Handeln und Lernen an Experimentierstationen*. In: www.widerstreit-sachunterricht.de, Nr. 27, Mai 2022 (16 Seiten).
- Giest, Hartmut (2022b): *Ist der Lernerfolg beim Experimentieren von Instruktion abhängig?* In: *GDSU-Journal*, 13, S. 23-37.
- Giest, Hartmut (2022c): *Methodologie und Methodiken in der Forschung zum Sachunterricht*. In: www.widerstreit-sachunterricht.de, Nr. 27, Mai 2022.
- Gogolin, Sarah, Krell, Moritz, Lange-Schubert, Kim, Hartinger, Andreas, Upmeier zu Belzen, Annette und Krüger, Dirk (2017): *Erfassung von Modellkompetenz bei Grundschüler/innen*. In: Hartmut Giest, Sandra Tänzer & Andreas Hartinger (Hrsg.): *Vielperspektivität im Sachunterricht*. Bad Heilbrunn, S. 108-115.
- Graf, Bernhard und Noschka-Roos, Annette (2009): *Stichwort Lernen im Museum oder: Eine Kamerafahrt mit der Besucherforschung*. In: *Zeitschrift für Erziehungswissenschaft*, 12, 1, S. 7-27.
- Grygier, Patricia, Jonen, Angela, Kircher, Ernst, Sodian, Beate & Thoermer, Claudia (2008): *„Wissenschaftsverständnis“ und Erwerb von naturwissenschaftlichem Wissen und Experimentierfähigkeit in der Grundschule*. In: Hartmut Giest & Jutta Wiesemann (Hrsg.): *Kind und Wissenschaft*. Bad Heilbrunn, S. 69-82.
- Guderian, Pascal (2007): *Wirksamkeitsanalyse außerschulischer Lernorte*. Diss. Berlin: Humboldt Universität.
- Hardy, Ilonca, Kleickmann, Thilo, Koerber, Susanne, Mayer, Daniela, Möller, Kornelia, Pollmeier, Judith, Schwippert, Knut & Sodian, Beate (2010): *Die Modellierung naturwissenschaftlicher Kompetenz im Grundschulalter*. In: *Zeitschrift für Pädagogik*, 56. Beiheft, S. 115-125.
- Haslbeck, Heide, Lankes, Eva-Maria, Kohlhauf, Lucia & Neuhaus, Birgit (2019): *Wie viele Variablen darf ich beim Experimentieren variieren? Ein Training für Grundschullehrkräfte zum Einsatz der Variablenkontrollstrategie im Unterricht*. In: Martina Knörzner, Lars Förster, Ute Franz & Andreas Hartinger (Hrsg.): *Forschendes Lernen im Sachunterricht*. Bad Heilbrunn, S. 47-54.
- Itzek-Greulich, Heike, Flunger, Barbara, Vollmer, Christian, Nagengast, Benjamin, Rehm, Markus & Trautwein, Ulrich (2015): *Effects of a science center outreach lab on school students' achievement – Are student lab visits needed when they teach what students can learn at school?* In: *Learning and Instruction*, 38, pp. 43-52.
- Jebe, Frank, Konietzko, Sebastian, Lichtschlag, Margrit & Liebau, Eckard (2019): *Studie: Jugend/YouTube/Kulturelle Bildung*. In: *Horizont* 2019. Essen. https://www.bosch-stiftung.de/sites/default/files/publications/pdf/2019-06/Studie_Jugend%20Youtube%20Kulturelle%20Bildung%202019.pdf [27.03.2023].
- Kalthoff, Britta, Theyßen, Heike & Schreiber, Nico (2016): *Vergleich von expliziter und impliziter Instruktion im Experimentalpraktikum für Sachunterrichtsstudierende*. In: Hartmut Giest, Thomas Goll & Andreas Hartinger (Hrsg.): *Sachunterricht – zwischen Kompetenzorientierung, Persönlichkeitsentwicklung, Lebenswelt und Fachbezug*. Bad Heilbrunn, S. 132-139.
- Kaltman, Gwen S. (2010): *Hands-on learning*. Washington DC.
- Kay, Robin H. (2012): *Exploring the use of video podcasts in education: A comprehensive review of the literature*. In: *Computers in Human Behavior*, 28, 3, pp. 820-831. DOI: 10.1016/j.chb.2012.01.011.
- Koerber, Susanne, Sodian, Beate, Kropf, Nicola, Mayer, Daniela & Schwippert, Knut (2011): *Die Entwicklung des wissenschaftlichen Denkens im Grundschulalter. Theorieverständnis, Experimentierstrategien, Dateninterpretation*. In: *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43, 1, S. 16-21.
- Koerber, Susanne, Mayer, Daniela, Osterhaus, Christopher, Schwippert, Knut & Sodian, Beate (2015): *The development of scientific thinking in elementary school. A comprehensive inventory*. In: *Child Development*, 86, 1, pp. 327-336.
- Köster, H. (2018): *Freies Explorieren und Experimentieren. Eine Untersuchung zur selbstbestimmten Gewinnung von Erfahrungen mit physikalischen Phänomenen im Sachunterricht*. 2., unveränd. Aufl. Berlin.
- Kori, Külli, Pedaste, Margus, Leijen, Äli & Mäeots, Mario (2014): *Supporting reflection in technology-enhanced learning*. In: *Educational Research Review*, 11, pp. 45-55.
- Krämer, Andreas & Böhrs, Sanrda (2017): *How do consumers evaluate explainer videos? An empirical study on the effectiveness and efficiency of different explainer video formats*. In: *Journal of Education and Learning*, 6, 1, pp. 254-266. <https://doi.org/10.5539/jel.v6n1p254>.
- Kulgemeyer, Christoph & Wolf, Karsten D. (2016): *Lernen mit Videos? Erklärvideos im Physikunterricht*. In: *Naturwissenschaften im Unterricht. Physik*, 27, 152, S. 36-41.
- Laufs, Ann-Kathrin & Kempert, Sebastian (2021): *Außerschulische Interessen als Ressource für die individualisierte Vermittlung der Variablenkontrollstrategie in der Grundschule*. In: *Unterrichtswissenschaft*, 49, S. 31-56. <https://doi.org/10.1007/s42010-020-00092-6>.

- Leutner, Detlev & Brünken, Roland (2017): Lernen mit elektronischen Medien. In: Detlef H. Rost, Jörn R. Sparfeldt & Susanne R. Buch (Hrsg.): Handwörterbuch Pädagogische Psychologie. 5. Aufl. Weinheim, S. 450-455.
- Lewalter, Doris & Geyer, Claudia (2009): Evaluation von Schulklassenbesuchen im Museum. In: Zeitschrift für Erziehungswissenschaft, 3, S. 775-785.
- Lowe, Richard K. & Boucheix, Jean-Michel (2016): Principled animation design improves comprehension of complex dynamics. In: Learning and Instruction, 45, pp. 72-84.
- Marquardt-Mau, Brunhild & Hoffmann, Yvonne (2009): Naturwissenschaften in altersgemischten Lernsituationen. In: Barbara Berthold & Heike Hahn (Hrsg.): Altersmischung als Lernressource – Impulse aus Fachdidaktik und Grundschulpädagogik. Baltmannsweiler, S. 268-284.
- Matthes, Eva, Siegel, Stefan T. & Heiland, Thomas (Hrsg., 2021): Lehrvideos – das Bildungsmedium der Zukunft? Bad Heilbrunn.
- Mayer, Richard E. (2014): Cognitive theory of multimedia learning. In: Richard E. Mayer (Ed.): The Cambridge handbook of multimedia learning. 2nd Ed. New York, pp. 43-71.
- Mayer, Richard E. & Fiorella, Logan (2022): The Cambridge handbook of multimedia learning. 3rd Ed. New York u.a. <https://doi.org/10.1017/9781108894333>.
- Mayer, Richard E. (2021): Evidence-Based Principles for How to Design Effective Instructional Videos. In: Journal of Applied Research in Memory and Cognition, 10, 2, pp. 229-240.
- Michalik, Kerstin (2010): Didaktische Konzepte für die naturwissenschaftliche Grundbildung von Kindern im Elementarbereich. In: Hans-Joachim Fischer, Peter Gansen & Kerstin Michalik (Hrsg.): Sachunterricht und frühe Bildung. Bad Heilbrunn, S. 93-108.
- Minnameier, Gerhard, Hermkes, Rico und Mach, Hanna (2015): Kognitive Aktivierung und konstruktive Unterstützung als Prozessqualitäten des Lehrens und Lernens. In: Zeitschrift für Pädagogik, 61, 6, S. 836-855.
- Murmann, Lydia, Pech, Detlef, Schomaker, Claudia & Stiller, Jurik (2019): „Aus der Perspektive von Kindern“ – Inwiefern kann der Forschungsansatz der Phänomenographie ein Impuls für die Beschreibung kindlicher Lernentwicklung und die Formulierung von Kompetenzniveaus sein? In: GDSU-Journal, Heft 9, S. 82-94.
- Nicolaou, Christiana T. & Constantinou, Costas P. (2014): Assessment of the modeling competence: A systematic review and synthesis of empirical research. In: Educational Research Review, 13, pp. 52-73.
- Pahl, Angelika & Lück, Gisela (2016): Naturwissenschaftliches Experimentieren – Kinder frühzeitig und kompetenzorientiert. In: Hartmut Giest, Thomas Goll & Andreas Hartinger (Hrsg.): Sachunterricht – zwischen Kompetenzorientierung, Persönlichkeitsentwicklung, Lebenswelt und Fachbezug. Bad Heilbrunn, S. 58-65.
- Pedaste, Margus, Mäeots, Mario, Siiman, Leo A., de Jong, Ton, van Riesen, Siswa A.N., Kamp, Ellen T., Manoli, Constantinos C., Zacharia, Zacharias C. & Tsourlidaki, Eleftheria (2015): Phases of inquiry-based learning: Definitions and the inquiry cycle. In: Educational Research Review, 14, pp. 47-61.
- Ramseger, Jörg (2013): Prozessbezogene Qualitätskriterien für den naturwissenschaftlichen Unterricht – zehn Kriterien für wirksames didaktisches Handeln im Elementar- und Primarbereich. In: Stiftung Haus der kleinen Forscher: Wissenschaftliche Untersuchungen zur Arbeit der Stiftung „Haus der kleinen Forscher“. Bd. 5. Schaffhausen, S. 147-171.
- Reinhold, Peter (2010): Den Physikunterricht fundieren. Elementarisierung und didaktische Rekonstruktion. In: Helmut F. Mikelskis (Hrsg.): Physik-Didaktik. Praxishandbuch für die Sekundarstufe I und II. 2. Aufl. Berlin, S. 86-119.
- Richter, Juliane, Scheiter, Katharina & Eitel, Alexander (2016): Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis. In: Educational Research Review, 17, pp. 19-36.
- Roth, Wolff-Michael, McRobbie, Campbell J., Lucas, Keith B. & Boutonné, Sylvie (1997): The Logical Production of Order in Traditional Science Laboratories: A Phenomenological Analysis. In: Learning and Instruction, 7, 2, pp. 107-136.
- Scharfenberg, Franz-Josef, Möller, Andrea, Kaufmann, Kathrin & Bogner, Franz X. (2019): Schülerlabore und Lehr-Lern-Labore. In: Jorge Groß, Marcus Hammann, Philipp Schmiemann & Jörg Zabel (Hrsg.): Biologie-didaktische Forschung: Erträge für die Praxis. Berlin, S. 229-249. DOI:10.1007/978-3-662-58443-9_13.
- Scheid, Jochen, Müller, Andreas, Schnotz, Wolfgang, Hettmannsperger, Rosa, Kuhn, Jochen, Telli, Sibel & Vogt, Patrik (2014): Development of Representational Competence via Cognitively Activating Tasks for Physics Experiments Conference: GIREP 2010 – Teaching and Learning Physics Today: Challenges? Benefits? https://www.researchgate.net/publication/266234448_Development_of_Representational_Competence_via_Cognitive_Activating_Tasks_for_Physical_Experiments [27.03.2023].
- Schütte, Florian (2019): Freies Explorieren zum Thema elektrischer Stromkreis. Eine Suchraumrekonstruktion nach der dokumentarischen Methode. Wiesbaden. DOI <https://doi.org/10.1007/978-3-658-23059-32019>.
- Schwichow, Martin G., Croker, Steve, Zimmerman, Corinne, Höffler, Tim N. & Härtig, Hendrik (2016): Teaching the control of variables strategy. A meta-analysis. In: Developmental Review, 39, pp. 37-63.

- van Joolingen, Wouter R., de Jong, Ton, Lazonder, Adrianus W., Savelsbergh, Elvin R. & Manlove, Sarah (2005): Co-lab: Research and development of an on-line learning environment for collaborative scientific discovery learning. In: *Computers in Human Behavior*, 21, pp. 671-688.
- Vosniadou, Stella (Ed., 2015): *International Handbook of Research on Conceptual Change*. New York u.a.
- Vygotskij, Lev S. (2002): *Denken und Sprechen*. Weinheim & Basel.
- Waldenmaier, Christine, Köster, Hilde & Müller, Bernhard (2013): Unterschiede bezüglich der Engagiertheit von Kindergruppen bei geöffneten und geschlossenen Experimentierangeboten im naturwissenschaftsbezogenen Sachunterricht. In: Hans-Joachim Fischer, Hartmut Giest & Detlef Pech (Hrsg.): *Der Sachunterricht und seine Didaktik*. Bad Heilbrunn, S. 137-146.
- Waldenmaier, Christine, Müller, Bernhard, Köster, Hilde & Körner, Hans-Dieter (2015): Engagiertheit und Motivation in unterschiedlichen Experimentiersituationen im Sachunterricht. In: Hans-Joachim Fischer, Hartmut Giest & Kerstin Michalik (Hrsg.): *Bildung im und durch Sachunterricht*. Bad Heilbrunn, S. 87-92.
- Windt, Anna, Scheuer, Rupert & Melle, Insa (2011): Methoden beim naturwissenschaftlichen Experimentieren im Elementarbereich. In: Hartmut Giest, Astrid Kaiser & Claudia Schomaker (Hrsg.): *Sachunterricht – auf dem Weg zur Inklusion*. Bad Heilbrunn, S. 157-162.
- Wolf, Karsten D. (2015): Video-Tutorials und Erklärvideos als Gegenstand, Methode und Ziel der Medien- und Filmbildung. In: Anja Hartung-Griemberg, Thomas Ballhausen, Christine Trültzsch-Wijnen, Alessandro Barberi & Katharina Kaiser-Müller (Hrsg.): *Filmbildung im Wandel*. Wien, S. 121-131.
- Zadeh, Mahsa Vali & Peschel, Markus (2018): SelfPro-Entwicklung von Selbstkonzepten beim Offenen Experimentieren. In: Ute Franz, Hartmut Giest, Andreas Hartinger, Anja Heinrich-Dönges & Bernd Reinthoffer (Hrsg.): *Handeln im Sachunterricht*. Bad Heilbrunn, S. 183-190.