

Analoges Enkodieren und die Fähigkeit zum Transfer im naturwissenschaftlichen Sachunterricht

Katrin Lohrmann · Jana Groß Ophoff · Andreas Hartinger

Zusammenfassung Forschungsbefunde zeigen, dass der Vergleich multipler Konkretionen, orientiert am Ansatz des Analoges Enkodierens, den Wissenserwerb positiv unterstützt. Offen ist, inwieweit dies auch für die Fähigkeit zum Transfer des erworbenen Wissens gültig ist und inwieweit die Ausgestaltung der im Unterricht verwendeten Konkretionen (bezogen auf ihre oberflächliche Ähnlichkeit) von Bedeutung ist. Zur Beantwortung dieser Frage wurden im Rahmen einer quasi-experimentellen Unterrichtsstudie 367 Schülerinnen und Schüler der 3. Jahrgangsstufe in ihrer Fähigkeit zum nahen Transfer untersucht. Die Analyse latenter Regressionsmodelle zeigt, dass sich die untersuchten Merkmale (Versuchsgruppen, Intelligenz, prozedurales bzw. konzeptuelles Wissen) unterschiedlich auf die Fähigkeit zum Transfer auswirken. Mögliche Gründe für die Ergebnisse sowie Konsequenzen für den Sachunterricht der Grundschule werden diskutiert.

Schlüsselwörter Transfer · Exemplarisches Lernen · Sachunterricht · Naturwissenschaftliches Lernen · Lehr-Lernforschung

Prof. Dr. K. Lohrmann (✉) · Dr. J. Groß Ophoff
Institut für Erziehungswissenschaft, Pädagogische Hochschule Freiburg,
Kunzenweg 21, 79117 Freiburg, Deutschland
E-Mail: katrin.lohrmann@ph-freiburg.de

Dr. J. Groß Ophoff
E-Mail: jana.grossophoff@ph-freiburg.de

Prof. Dr. A. Hartinger
Lehrstuhl für Grundschulpädagogik und -didaktik, Universität Augsburg,
Universitätsstraße 10, 86159 Augsburg, Deutschland
E-Mail: andreas.hartinger@phil.uni-augsburg.de

Analogical encoding and transfer ability in science education in primary school

Abstract According to the state of research, comparing multiple concrete cases aligned with the approach of Analogical Encoding can support learning. But it is still open, whether this applies to the ability to transfer acquired knowledge, too, and to what extent the design of the concrete cases (with regard to superficial similarity) used in instruction has an impact thereon. To answer this question, the ability of near transfer of 367 third-graders was investigated in a quasi-experimental study. The results of latent regression analysis indicate that the examined predictors (experimental group, intelligence, procedural vs. conceptual knowledge) have different effects on transfer ability. In conclusion, possible reasons and consequences for science education in primary school are discussed.

Keywords Transfer · Exemplary Learning · Science Learning · Teaching and Learning Research

1 Einleitung

Schulisches Lernen ist in seiner Sinnhaftigkeit letztlich nur legitimierbar, wenn man davon ausgeht, dass Schülerinnen und Schüler das im Unterricht erworbene Wissen auf neue Situationen übertragen können, sodass es ihnen auch bei der Bewältigung von Alltagssituationen dienlich ist. Deshalb ist es ein zentrales Ziel von Unterricht, konzeptuelles Wissen auf- und auszubauen, das auf neue Kontexte transferierbar ist. Dies gilt auch für den Sachunterricht der Grundschule, der im Sinne „Grundlegender Bildung“ (Köhnlein 2015) und vor dem Hintergrund von „Scientific Literacy“ (Prenzel et al. 2003) Schülerinnen und Schüler darin unterstützen soll, ihre Umwelt sachbezogen zu verstehen und darin zu handeln. Der Aufbau anwendungsfähigen, „intelligenten“ Wissens (Weinert 1998, S. 111) stellt eine besondere Herausforderung für Lernende dar, die häufig daran scheitern, das in institutionalisierten Bildungsprozessen erworbene Wissen auf unbekannte Situationen zu übertragen. Nach wie vor fehlt es an empirisch gesicherten Erkenntnissen über didaktische Settings, die den Erwerb transferierbaren Wissens unterstützen.

2 Stand der Forschung

2.1 Transferfähigkeit als Ziel sachunterrichtlichen Lernens

Mit dem Begriff *Transfer* wird die erfolgreiche Anwendung angeeigneten Wissens bzw. erworbener Fertigkeiten in neuen, beim Erlernen noch nicht vorgekommenen Situationen bezeichnet (Klauer 2011; Mähler und Stern 2018). In Abhängigkeit von der Nähe bzw. Distanz, die zwischen der Lern- und der Anwendungssituation besteht, kann Transfer unterschiedlich herausfordernd für Lernende sein. Das Rahmenmodell von Barnett und Ceci (2002) bietet eine theoretische Grundlage, die

Schwierigkeit von Transfer vorherzusagen. Die erste Dimension, die *Inhaltskomponente*, fokussiert auf das Wissen oder die Fertigkeiten, die von einer Lern- in eine Anwendungssituation übertragen werden sollen: Unterschieden wird dabei, ob spezifische, eng umschriebene Kenntnisse oder ein allgemeines Prinzip, eine generelle Problemlösestrategie auf neue Beispiele übertragen werden sollen. Anhand der zweiten Dimension, der *Kontextkomponente*, kann entlang von sechs Kriterien – Wissensdomäne, Modalität der Informationsverarbeitung sowie situativer, zeitlicher, funktionaler und sozialer Kontext – die Nähe bzw. die Distanz zwischen der Lern- und der Anwendungssituation beschrieben werden. Die Spezifikation im Hinblick auf diese Kriterien, die jeweils auf einem Kontinuum verortet sind, ermöglicht im Einzelfall eine Bestimmung zwischen nahem und weitem Transfer und somit eine Aussage über den jeweiligen Anspruch bzgl. der Transferierbarkeit des Lernens.

Befunde aus TIMSS (Trends in International Mathematics and Science Study) 2015 zeigen, dass es vielen Schülerinnen und Schülern im Bereich des naturwissenschaftlichen Lernens nicht gelingt, ihr (vorhandenes) Wissen anzuwenden. Kompetenzstufe IV, definiert als die Fähigkeit, Wissen „in alltäglichen und abstrakteren Kontexten an[z]uwenden“ (Steffensky et al. 2016, S. 155), wird zwar in Deutschland immerhin von 39,6% der Schülerinnen und Schülern erreicht (ebd., S. 169). Umgekehrt bedeutet dies jedoch auch, dass deutlich über die Hälfte der hier untersuchten Kinder nicht in der Lage ist, ihr Wissen anzuwenden oder dieses Wissen nicht hat. Den Sprung von Kompetenzstufe III („Die Schülerinnen und Schüler verfügen über ein Basiswissen“; ebd., S. 155) auf die Kompetenzstufe IV kann man in der Logik der verwendeten Kompetenzstufenmodelle als Fähigkeit zum (nahen) Transfer verstehen, was bedeutet, dass die 38,8% der Schülerinnen und Schüler, die sich auf Stufe III befinden, zwar das erforderliche Wissen haben, nicht jedoch in der Lage sind, dieses angemessen anzuwenden.

Die Untersuchung von Transfer ist vor diesem Hintergrund sowohl wissenschaftlich als auch unterrichtspraktisch relevant. Studien im Sachunterricht der Grundschule, deren zentraler Gegenstand die Überprüfung der Transferfähigkeit von schulisch erworbenem Wissen ist, liegen nach unserem Kenntnisstand nicht vor. In einzelnen Interventionsstudien wurde jedoch erhoben, inwieweit Schülerinnen und Schüler das in diesen Interventionen erworbene Wissen in verschiedenen Kontexten anwenden können. Untersucht wurden dabei v. a. perspektivenübergreifende Wissensselemente, wie z. B. das Wissenschaftsverständnis (Grygier et al. 2008) oder das Verständnis der Variablenkontrollstrategie (Bohmann und Möller 2017).

Mit Blick auf gelingenden Transfer fokussiert der Sachunterricht insbesondere auf den Aufbau eines „grundlegenden konzeptuellen Wissens und Verständnisses, das naturwissenschaftlichen Leitideen zugeordnet werden kann“ (BLK und Prenzel et al. 2004, S. 10). Ein solches konzeptuelles Wissen wird aufgebaut, indem durch die Auseinandersetzung mit Konkretionen eine allgemeine Erkenntnis gewonnen wird, die dann auf andere Konkretionen mit der gleichen Tiefenstruktur bezogen werden kann. Konzeptuelles Wissen ist somit per Definition auf Transfer angelegt. Verschiedene Forschungsbefunde zeigen, dass abstrakte Repräsentationen Transfer besser unterstützen als spezifisches Wissen (Chi et al. 1989; Chi und VanLehn 2012; Fong et al. 1986; Fong und Nisbett 1991); bei Letzterem erkennen Lernende die

Ähnlichkeit zwischen dem vorhandenen Wissen und der neuen Situation vielfach nicht.

Schülerinnen und Schüler können im Sachunterricht häufig auf vor- oder außerschulisch erworbenes Wissen zurückgreifen, das oft prozeduraler Natur ist. Spreckelsen (1995, 2004) geht davon aus, dass auch dieses Wissen transferiert werden kann – er spricht dabei von Transduktionen. Ohne auf ein allgemeines (naturwissenschaftliches) Gesetz Bezug zu nehmen, kann von einer Konkretion auf eine andere, oberflächlich ähnliche Konkretion gefolgert werden.

2.2 Förderung von Transferfähigkeit

Verschiedene Disziplinen beschäftigen sich mit der Frage, wie Lernsituationen gestaltet sein müssen, damit der Erwerb anwendungsfähigen Wissens gefördert wird. Dabei stellen sich auch differenzielle Fragestellungen, da Lernsettings für Lernende mit bestimmten Lernvoraussetzungen wirksamer sein können als für andere (vgl. dazu die Forschungstradition der Aptitude-Treatment-Interaction; zsf. Münzer und Brünken 2018).

Sowohl aus Sicht der Allgemeinen Didaktik (*exemplarische Bedeutung*; Klafki 1985), der Fachdidaktik (*Phänomenkreise*; Spreckelsen 2004), der Grundschuldidaktik (Lohrmann 2015) als auch der Lehr-Lernforschung (zsf. Lipowsky 2015) wird es als sinnvoll angesehen, Lernumgebungen basierend auf Konkretionen zu gestalten, um ausgehend davon allgemeine Gesetzmäßigkeiten zu verdeutlichen (zsf. Lohrmann et al. 2013). Dabei müssen Konkretionen für den zu lernenden Inhalt exemplarisch sein (Klafki 1962), also einen allgemeinen Bildungsgehalt in sich bergen. Letztlich decken sich diese Ansätze aus verschiedenen Disziplinen mit Überlegungen und Befunden aus der Wissenspsychologie (zsf. Lohrmann 2014): Durch Situierungen, also durch die Auseinandersetzung mit konkreten Ausgangssituationen, soll die Wissensstruktur um Einzelheiten elaboriert werden. Abstrahierungen vom jeweiligen Kontext wiederum sollen es ermöglichen, die Wissensstruktur auf wesentliche Anteile zu reduzieren und relationales Wissen sowie konzeptuelles Verständnis auf der Symbolebene aufzubauen (Gick und Holyoak 1983; Hummel und Holyoak 1997, 2003). Der stetige Wechsel zwischen konkreten und abstrahierten Wissensbestandteilen soll den Aufbau einer elaborierten, organisierten, vertikal und horizontal vernetzten sowie flexibilisierten Wissensstruktur fördern. Eine solche Wissensstruktur gilt als Voraussetzung für Wissenstransfer und Problemlösen.

Zur Förderung von Transfer werden meist drei Grundideen genannt (Mähler und Stern 2018): Der *Umgang mit mentalen Werkzeugen* zielt darauf, allgemeine Voraussetzungen für den Wissenstransfer zu fördern. Die *Förderung metakognitiver Kontrolle* fokussiert auf die Vermittlung von Wissen zur Steuerung und Regulation bewussten Lernens und des gezielten Einsatzes von Arbeitstechniken. Für die vorliegende Untersuchung sind die Überlegungen von besonderer Bedeutung, die *bewusste Vergleiche und Analogiebildungen* in den Fokus setzen. So konnten Alfieri et al. (2013) in einem Review von Labor- und Unterrichtsstudien zeigen, dass sich der *Vergleich multipler Konkretionen* zu einem Inhaltsbereich förderlich auf den Wissenserwerb und die Transferfähigkeit auswirkt. Die Vergleiche ermöglichen zum einen zusätzliche Elaborationen zur besseren Verankerung des neuen Wissens

und eröffnen zum anderen zusätzliche Möglichkeiten logischer Ableitungen und Schlussfolgerungen. Der Abstrahierungsprozess kann so mehrfach vollzogen werden, das Allgemeine wird an verschiedenen Beispielen deutlich. Damit trägt die Beschäftigung mit vielfältigen Konkretionen dazu bei, die Anwendbarkeit des Wissens zu erkennen und Transfer einzuüben (Chen und Daehler 1989; Gentner et al. 2003; Gerjets et al. 2008; Gick und Holyoak 1980, 1983; Loewenstein et al. 2003; Richland und McDonough 2010).

Analoges Enkodieren stellt eine Möglichkeit dar, das Lernen mit multiplen Konkretionen didaktisch umzusetzen (Gentner et al. 2003; Lohrmann et al. 2014; Schwelle 2016): Die Schülerinnen und Schüler werden dabei gleichzeitig mit mehreren, noch nicht durchgearbeiteten Konkretionen zu einem Funktionsprinzip konfrontiert. Dadurch sollen wechselseitige Vergleichsprozesse zwischen den Konkretionen angeregt werden, was den Blick von Oberflächenmerkmalen lösen und auf die gemeinsame Tiefenstruktur lenken soll. Voraussetzung ist das Erkennen struktureller Gemeinsamkeiten auf der Ebene von Elementen, Relationen und Systemen (Gentner 1989; Gentner und Kurtz 2006; Gentner und Markman 1994). Wenn Schülerinnen und Schüler bereits in der Phase des Wissenserwerbs zum Vergleichen und Schlussfolgern angeregt werden und erkennen, dass das erlernte Wissen für verschiedene Kontexte relevant und nicht an eine spezifische Situation gebunden ist, können das Abrufen von strukturell ähnlichen Beispielen aus dem Gedächtnis sowie ein Übertragen des erworbenen Wissens auf neue Situationen frühzeitig eingeübt und nachfolgender Transfer gefördert werden (Alfieri et al. 2013).

Das Vergleichen multipler Konkretionen erfordert eine Entscheidung darüber, wie diese Konkretionen beschaffen sein sollen, um den Lernprozess möglichst positiv anzuregen. Anzunehmen ist, dass der Aufbau einer flexiblen Wissensstruktur durch Phänomene, die auf der Oberflächenstruktur unähnlich sind, besser gefördert werden kann, da diese eine verstärkte Auseinandersetzung mit der Tiefenstruktur erforderlich machen. Voraussetzung ist jedoch, dass die Schülerinnen und Schüler beim Wissenserwerb die Gemeinsamkeiten zwischen den jeweiligen Konkretionen erkennen. Ob dieses Mapping zwischen zwei Konkretionen gelingt, wird – neben der jeweiligen (Un-)Ähnlichkeit der zu vergleichenden Phänomene – v. a. durch Intelligenz und inhaltspezifisches (Vor-)Wissen beeinflusst (Blanchette und Dunbar 2001; Chi et al. 1981; Kotovsky und Gentner 1996; Novick 1988; Rittle-Johnson und Star 2009; Vosniadou 1989): Personen mit einer ausgeprägten Intelligenz und reichhaltigem Vorwissen sind aufgrund ihrer kognitiven Möglichkeiten und Sachkenntnis zu höherwertigem Mapping fähig.

Inwiefern die Variation der Oberflächenstruktur (ähnlich, unähnlich) von naturwissenschaftlichen Phänomenen die Fähigkeit zum *Erwerb von Wissen* im Sachunterricht der Grundschule beeinflusst, wurde im Rahmen der hier berichteten Studie bereits untersucht (z. B. Lohrmann et al. 2014; Schwelle 2016). Am Beispiel des Hebelgesetzes zeigte sich, dass der auf der Grundlage des Analoges Enkodierens gestaltete Unterricht sowohl in der ähnlichen als auch in der unähnlichen Lernbedingung einen deutlichen Lernfortschritt bewirkte, wenngleich keine signifikanten Unterschiede zwischen den Versuchsgruppen aufgezeigt werden konnten. Anders als bei Kindern der unähnlichen Lernbedingung ging dieses Wissen jedoch bei Kindern der ähnlichen Bedingung drei Monate nach der Intervention etwas zurück (wenn

auch nicht signifikant). Mit Blick auf Schülervoraussetzungen konnte gezeigt werden, dass das Wissen unmittelbar (Post-Messzeitpunkt) und drei Monate nach der Intervention (Follow-Up-Messzeitpunkt) erheblich stärker durch das Vorwissen als durch die Intelligenz der Kinder begünstigt wird. Die erwarteten differenziellen Unterschiede zwischen den beiden Lernbedingungen konnten nicht festgestellt werden.

Offen ist allerdings, inwieweit der Vergleich ähnlicher oder unähnlicher Phänomene beim Lernen Bedeutung für die Fähigkeit zum *Transfer des erworbenen Wissens* hat. Die Beantwortung dieser Frage erweitert den eben berichteten Befund um die Untersuchung der Qualität des Wissens.

3 Forschungsfragen

Folgende Fragestellungen stehen damit im Fokus dieses Beitrags:

1. Wie wirkt sich die Lernbedingung (Lernen an oberflächlich ähnlichen vs. unähnlichen Phänomenen) auf die Fähigkeit zum Transfer aus?
2. Welche Zusammenhänge bestehen zwischen verschiedenen Formen des Wissens (prozedural, konzeptuell) zum Follow-Up-Messzeitpunkt und der Fähigkeit zum Transfer?
3. Gibt es mit Blick auf die ähnliche vs. unähnliche Lernbedingung differenzielle Effekte bezüglich
 - a) der Intelligenz der Schülerinnen und Schüler und der Fähigkeit zum Transfer?
 - b) der im Unterricht erworbenen verschiedenen Formen des Wissens (prozedural, konzeptuell) und der Fähigkeit zum Transfer?

4 Methodisches Vorgehen

4.1 Untersuchungsdesign

In einer quasi-experimentellen Unterrichtsstudie wurden Daten von 367 Schülerinnen und Schülern aus 22 dritten Klassen aus 20 Schulen des Großraums Freiburg erfasst. Es handelte sich um einen nicht-randomisierten Drei-Gruppen-Plan mit drei Messzeitpunkten (Prä: vor der Intervention; Post: unmittelbar nach der Intervention; Follow-Up: ca. drei Monate nach der Intervention). Für die Beantwortung der oben genannten Fragen sind die Befunde des Prätests sowie des Follow-Up-Tests von Bedeutung.

Die Intervention umfasste sieben Unterrichtsstunden und fand zwischen dem Prä- und Post-Messzeitpunkt statt (vgl. hierzu im Einzelnen Schwelle 2016). Als Unterrichtsgegenstand der Intervention wurde das Hebelgesetz gewählt, fokussiert auf die Prinzipien Gleichgewicht (drei Unterrichtsstunden) und Kraftverstärkung (vier Unterrichtsstunden). Durch die Variation der Ähnlichkeitsbeziehung der Beispiele (ähnlich bzw. unähnlich) ergaben sich die beiden Versuchsgruppen. Die Kinder der Kontrollgruppe nahmen an allen Tests, nicht jedoch an der Intervention teil.

4.2 Eingesetzte Testinstrumente: Transfertest, CFT 20-R, Wissenstest

Im Mittelpunkt dieses Beitrags steht die Fähigkeit zum Transfer des Wissens, das zum Follow-Up-Messzeitpunkt mittels eines Transfertests (12 Items zu Gleichgewicht, 15 Items zu Kraftverstärkung) untersucht wurde. Die Bestimmung der Transferdistanz erfolgte anhand der Kriterien von Barnett und Ceci (2002) mit Blick auf die Differenz von Lernsituation (Unterricht) und Anwendungssituation (Transfertest) (vgl. Abschn. 2). Im Fokus stand die Inhaltskomponente: Überprüft wurde die Fähigkeit zur Anwendung des Hebelgesetzes (Gleichgewicht, Kraftverstärkung) auf neue, im Unterricht nicht thematisierte Phänomene. So hatten z. B. die Kinder beider Versuchsgruppen relevante Merkmale des Gleichgewichtsprinzips (Drehpunkt, zwei Hebelarme, wirkende Kräfte) im Unterricht anhand der Wippe kennengelernt, was

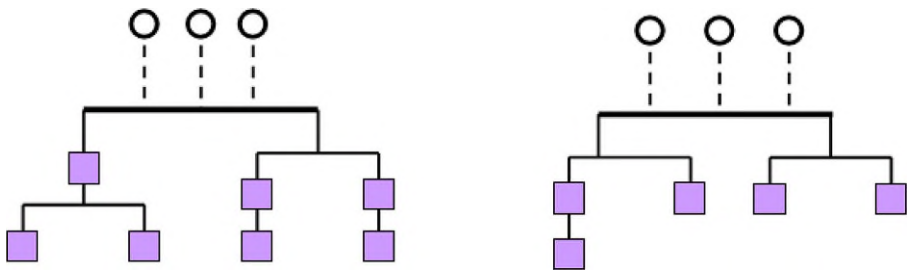


Abb. 1 Beispielitem Gleichgewicht (Transfertest).

Aufgabenstellung: An welcher Stelle muss Anton den Faden befestigen, damit die Mobiles gerade hängen? Kreuze an



Abb. 2 Beispielitem Kraftverstärkung (Transfertest).

Aufgabenstellung: Ein starker Sturm hat einen Baum entwurzelt, ohne den Stamm abzubringen. Kannst du das mit dem Hebelgesetz erklären? (Antworten der Kinder ohne Bezug zum Hebelgesetz wurden als falsch gewertet, wohingegen Antworten als richtig gewertet wurden, in denen relevante Aspekte des Hebelgesetzes benannt wurden. Die Interraterreliabilität wurde auf der Grundlage von 100 Antworten überprüft. Cohens κ (Wirtz und Caspar 2002) beträgt 0,93 und liegt damit im sehr guten Bereich).

Bildquelle: <http://www.christoph-kranich.de/hamburg.html>

auch im Wissenstest abgefragt wurde. Im Transfertest sollten sie das im Unterricht erworbene Wissen dann z. B. auf Mobiles (Phänomen: Gleichgewicht, vgl. Abb. 1) oder das Entwurzeln von Bäumen (Phänomen: Kraftverstärkung, vgl. Abb. 2) übertragen, was im Sinne von Alfieri et al. (2013, S. 101) als naher Transfer klassifiziert werden kann.

Um überprüfen zu können, inwieweit sich die Intelligenz der Kinder auf die Fähigkeit zum Transfer auswirkt, wurde zum Prä-Messzeitpunkt der Culture-Fair-Test (CFT 20-R, erster Testteil; Weiß 2006) eingesetzt.

Zur Erfassung des Wissens wurde an den drei Messzeitpunkten ein und derselbe Wissenstest mit 17 Aufgaben (und 42 Items) zum Hebelgesetz eingesetzt. Dieser erfasste prozedurales und konzeptuelles Wissen zu Gleichgewicht und Kraftverstärkung, bezogen auf jene Phänomene, die Gegenstand des Unterrichts waren. Die Reliabilität der Skalen war akzeptabel bis gut (Schwelle 2016).

4.3 Stichprobe

Grundlage für die vorgestellten Ergebnisse sind die Daten von 367 Schülerinnen und Schüler, die zu allen drei Messzeitpunkten an den Erhebungen teilnahmen. Die Zuweisung zu einer der beiden Versuchsgruppen (ähnlich vs. unähnlich) erfolgte mittels der Matched-Samples-Methode (Kupper et al. 1981).

Zwischen den drei Gruppen zeigen sich vor der Intervention keine Unterschiede bezüglich Geschlecht, Alter und Intelligenz (vgl. Tab. 1). Auch gibt es in keiner Gruppe überzufällige Differenzen bei den CFT-Werten im Vergleich zu den Werten, die im Testmanual für die Normierungsstichprobe der Acht- bis Neunjährigen berichtet werden (Weiß 2006). Die drei Gruppen unterscheiden sich außerdem nicht hinsichtlich des prozeduralen und konzeptuellen Vorwissens zum Hebelgesetz, das zum Prä-Messzeitpunkt mit einem Wissenstest erfasst wurde (vgl. Schwelle 2016). Die Schülerinnen und Schüler zeigen zwischen Prä- und Follow-Up-Messzeitpunkt

Tab. 1 Deskriptive Statistiken

	Ähnlich (VG1)	Unähnlich (VG2)	Kontrollgruppe (KG)
<i>n</i>	128	128	111
Weiblich	45,3%	50,0%	47,7%
Alter: <i>M</i> (<i>SD</i>)	8,4 (0,6)	8,4 (0,6)	8,4 (0,6)
CFT 20-R: <i>M</i> (<i>SD</i>)	103,8 (14,7)	101,9 (15,0)	105,5 (15,2)
<i>Wissen zum Hebelgesetz (Prä)</i>			
Prozedural: <i>M</i> (<i>SD</i>) ^a	5,0 (1,8)	4,4 (1,8)	5,1 (1,8)
Konzeptuell: <i>M</i> (<i>SD</i>) ^b	13,5 (4,9)	13,1 (5,3)	14,5 (5,3)
<i>Wissenszuwachs (FU minus Prä)</i>			
Prozedural: <i>M</i> (<i>SD</i>)	-0,4 (1,8)	0,4 (1,8)	0,0 (2,0)
Konzeptuell: <i>M</i> (<i>SD</i>)	4,8 (6,0)	5,7 (6,8)	2,1 (5,2)

VG1 Versuchsgruppe ähnlich, VG2 Versuchsgruppe unähnlich, *n* Anzahl der Schülerinnen und Schüler, *M* Mittelwert, *SD* Standardabweichung, *Prä* Prä-Messzeitpunkt, *FU* Follow-Up-Messzeitpunkt

^aMaximale Punktzahl: 8 Punkte bei 6 Aufgaben

^bMaximale Punktzahl: 34 Punkte bei 11 Aufgaben

einen durchschnittlichen Punktergebnis von $M=0,02$ ($SD=1,92$) im prozeduralen Wissen und von $M=4,17$ ($SD=6,12$) im konzeptuellen Wissen. Hier ergeben sich signifikante Unterschiede zwischen den drei Gruppen (prozedural: $F_{(0,05; 2)}=3,80$; partielles $\eta^2=0,02$; konzeptuell: $F_{(0,05; 2)}=10,80$; partielles $\eta^2=0,06$). Auffällig ist der Wissenszuwachs mittlerer Effektstärke beim konzeptuellen Wissen, insbesondere bei den beiden Versuchsgruppen.

4.4 Analyseverfahren

In einem ersten Schritt wurden sowohl die Testitems des *Transfertests* als auch des *CFT* auf Basis der Item-Response-Theorie (R-Paket TAM; Kiefer et al. 2016) ausführlich analysiert. Aus dem *Transfertest* mit ursprünglich 27 Items wurde ein Item ausgeschlossen, das von keiner Person gelöst wurde. Weitere sieben Items wurden wegen schlechtem Item-Fit ($0,80 > WMNSQ > 1,20$; Adams 2002) bzw. zu niedriger Trennschärfe ($r < 0,20$) aus den Analysen ausgeschlossen, woraus ein Pool aus 19 Items (Kraftverstärkung: 13 Items; Gleichgewicht: 6 Items) resultierte. Für die *CFT*-Items ergab sich durchgehend ein guter Item-Fit, sodass in den weiteren, nachfolgend vorgestellten Modellen die in dem Testmanual (Weiß 2006) vorgegebenen Altersnormwerte (8,5–9,0 Jahre) berücksichtigt wurden (vgl. Tab. 1). Die psychometrische Prüfung des Wissenstests ist ausführlich dargestellt bei Schwelle (2016).

Die psychometrische Struktur des *Transfertests* wurde durch konkurrierende Modellvergleiche für den unbereinigten und den bereinigten Itemsatz analysiert (vgl. Tab. 2): Geprüft wurden ein eindimensionales Modell (Hebelgesetz) und ein zweidimensionales Modell (Gleichgewicht, Kraftverstärkung), jeweils spezifiziert als einparametrisches (1PL) und als zweiparametrisches Modell (2PL). Die zweidimensionalen Modelle wurden so spezifiziert, dass Items nur auf einer Dimension (Gleichgewicht oder Kraftverstärkung) laden (Between-Item-Mehrdimensionalität; Adams et al. 1997). Für den Vergleich der Modelle wurden verschiedene informationstheoretische Maße (Fit-Indizes) herangezogen: Dabei misst der AIC die Passung des Modells mit den Daten und berücksichtigt die Modellkomplexität, der BIC berück-

Tab. 2 Fit-Statistiken zum Vergleich der konkurrierenden Modelle vor und nach Ausschluss von Items

Skalierungslauf	Modell	Final Deviance	n_p	AIC	BIC	CAIC
Unbereinigt (26 Items)	Modell 1a: eindimensional (1PL)	10.009,66	27	10.064	10.169	10.196
	Modell 1b: eindimensional (2PL)	9784,79	52	9889	10.092	10.144
	Modell 2a: zweidimensional (1PL)	9773,60	29	9832	9945	9974
	Modell 2b: zweidimensional (2PL)	9387,51	53	<i>9494</i>	<i>9700</i>	<i>9753</i>
Bereinigt (19 Items)	Modell 1a: eindimensional (1PL)	8019,99	20	8060	8138	8158
	Modell 1b: eindimensional (2PL)	7867,42	38	7943	8092	8130
	Modell 2a: zweidimensional (1PL)	7774,36	22	7818	7904	7926
	Modell 2b: zweidimensional (2PL)	7601,95	39	<i>7680</i>	<i>7832</i>	<i>7871</i>

Stichprobengröße: $N=384$

Die Kennwerte der besten Lösung sind kursiv hervorgehoben

n_p Anzahl der geschätzten Parameter, *AIC* Akaike Information Criterion, *BIC* Bayesian Information Criterion, *CAIC* Consistent Akaike Information Criterion

sichtigt die Stichprobengröße, und der CAIC wird ergänzend als robuster Schätzer empfohlen (vgl. Tab. 2). Als Entscheidungsregel gilt, dass das Modell mit den niedrigsten Werten am besten zu den Daten passt (z. B. Schermelleh-Engel et al. 2003).

Die Fit-Statistiken der vier konkurrierenden Modelle zeigen, dass das zweidimensionale, zweiparametrische Modell (2b) die niedrigsten Werte in den informationstheoretischen Maßen sowohl im unbereinigten als auch im bereinigten Item-Pool hat. Durch den Transfertest scheinen demnach zwei Dimensionen bzw. Kompetenzfacetten, *Gleichgewicht* und *Kraftverstärkung*, erfasst zu werden. Dass das zweiparametrische (2b) dem einparametrischen (2a) Modell überlegen ist, verdeutlicht, dass die Items unterschiedlich stark zwischen Schülerinnen und Schülern mit hohen bzw. geringen Transferleistungen trennen. Dieses Modell bildete die Grundlage für die weiteren Analysen (vgl. Abschn. 5).

Gegenüber konventionellen dimensionsanalytischen Ansätzen zeichnen sich die hier angewendeten probabilistischen Testmodelle dadurch aus, dass Personenfähigkeiten und Aufgabenschwierigkeiten unabhängig voneinander geschätzt und auf einer gemeinsamen Skala angeordnet sowie interpretiert werden können. Die mittlere Itemschwierigkeit liegt für die Dimension Gleichgewicht (GG) bei $\delta_{GG}=0,81$ ($SD=0,78$; $\min=-0,77$; $\max=1,36$) und für die Dimension Kraftverstärkung (KV) bei $\delta_{KV}=0,44$ ($SD=1,17$; $\min=-1,05$; $\max=3,37$). Damit ist der Test mit einer mittleren Lösungshäufigkeit von 39% (GG) bzw. 45% (KV) als anspruchsvoll zu werten.

Die Reliabilität (Expected a posteriori/Plausible Value-Reliabilität, vergleichbar mit Cronbachs α , vgl. Rost 2004) – kann für die Dimension Gleichgewicht mit 0,71 als akzeptabel bezeichnet werden und fällt für die Dimension Kraftverstärkung mit 0,81 gut aus. Die Leistungen im Transfertest liegen für beide Dimensionen und über alle Gruppen hinweg bei rund null Logits (GG (EAP): $SD=0,84$; KV (EAP): $SD=0,90$). Dabei bewegen sich die 367 Kinder in einem Wertebereich von \min (EAP) = -1,36 und \max (EAP) = 1,84 Logits auf der Gleichgewichtsskala und in einem etwas größeren Wertebereich von \min (EAP) = -1,82 und \max (EAP) = 2,21 Logits auf der Kraftverstärkungsskala. Mit anderen Worten bedeutet dies, dass es im Transfertest sowohl sehr geringe als auch sehr hohe Transferleistungen von Schülerinnen und Schülern gibt.

Eine inhaltliche Analyse der verbleibenden Items in beiden Skalen zeigt, dass es eine gewisse Konfundierung gibt: Items, die in die Skala Gleichgewicht eingingen, erfordern überwiegend ein Wissen, das als prozedural einzuschätzen ist, wohingegen die Items zu Kraftverstärkung einen Schwerpunkt im konzeptuellen Wissen haben.

Um den Einfluss von Merkmalen auf die Fähigkeit zum Transfer zu untersuchen, wurden anschließend latente Regressionsmodelle analysiert (vgl. Tab. 3), in welche die Gruppenzugehörigkeit zur ähnlichen oder unähnlichen Versuchsgruppe (dummycodiert, Modell 1), der Leistungsstand zum Follow-Up-Messzeitpunkt (Modell 2), die Wechselwirkung zwischen Gruppenzugehörigkeit und Intelligenz (Modell 3) sowie zwischen Gruppenzugehörigkeit und Lerngewinn (erfasst durch den Differenzwert [Follow-Up minus Prä] im Wissenstest; Modell 3) aufgenommen wurden. Um die Regressionsgewichte der Interaktionseffekte gruppenspezifisch interpretieren zu können, wurden die Werte um den jeweiligen Gruppenmittelwert zentriert.

Tab. 3 Ergebnisse für die latenten Regressionsanalysen (zweidimensionales, zweiparametrisches Modell, standardisierte Lösung)

Modell	β (VG1) ^b	β (VG2) ^b	β (IQ) ^a	β proz ^a	β konz ^a	β (Δ prozedural) ^a	β (Δ konzeptuell) ^a	R ² (in %)
1	-0,08	0,15	0,24*	-	-	-	-	6,3
	0,96*	0,96*	0,44*	-	-	-	-	35,7
2	-0,07	0,11	0,12*	0,25*	0,12*	-	-	15,5
	0,87*	0,82*	0,35*	-0,03	0,29*	-	-	42,0
3	-0,11	0,08	0,19*/0,07/0,12*	-	-	0,07/-0,04/0,09	0,08/0,11/-0,05	10,2
	0,90*	0,84*	0,28*/0,32*/0,11*	-	-	-0,02/0,01/-0,04	0,10*/0,10*/-0,10	40,4

N = 367 Schülerinnen und Schüler

Signifikante Regressionskoeffizienten ($p < 0,05$; zweiseitige Testung) sind durch * hervorgehoben

Zur Beurteilung standardisierter Lösungen können die Angaben zur Interpretation der Effektstärke von Regressionskoeffizienten herangezogen werden, wonach ein $\beta \geq 0,10$ einem kleinen Effekt, $\beta \geq 0,30$ einem mittleren und $\beta \geq 0,50$ einem großen Effekt entspricht (Kline 2010)

β Regressionskoeffizient, auf die zweite Nachkommastelle gerundet, GG Gleichgewicht, KV Kraftverstärkung, VG1 Versuchsgruppe ähnlich, VG2 Versuchsgruppe unähnlich, KG Kontrollgruppe, IQ Ergebnisse aus dem CFT 20-R, proz prozedurales Wissen zum Follow-Up-Messzeitpunkt, konz konzeptuelles Wissen zum Follow-Up-Messzeitpunkt, Δ prozedural Differenz aus Ergebnis des prozeduralen Wissens im Wissenstest zum Follow-Up-Messzeitpunkt minus Ergebnis des prozeduralen Wissens im Wissenstest zum Prä-Messzeitpunkt, Δ konzeptuell Differenz aus Ergebnis des konzeptuellen Wissens im Wissenstest zum Follow-Up-Messzeitpunkt minus Ergebnis des konzeptuellen Wissens im Wissenstest zum Prä-Messzeitpunkt, R² prozentuale Varianzaufklärung durch das Modell

^aFür metrische Prädiktoren sind vollständig standardisierte Regressionskoeffizienten angegeben

^bFür dichotome Prädiktoren sind Regressionskoeffizienten angegeben, die nur für die Logit-Skala des Transfertestes standardisiert wurden

5 Ergebnisse

Für den Vergleich der beiden Versuchsgruppen wurde zunächst in latenten Regressionsanalysen die Zugehörigkeit zu einer der beiden Gruppen als Prädiktor aufgenommen (vgl. Forschungsfrage 1), kontrolliert durch die CFT-Leistung. Die Ergebnisse in Modell 1 (vgl. Tab. 3) zeigen, dass sich die Zugehörigkeit zu den Versuchsgruppen ausschließlich in Bezug auf den Transfer von Wissen im Bereich Kraftverstärkung positiv auswirkt (großer Effekt). Anders als angenommen unterscheiden sich die beiden Versuchsgruppen jedoch nicht signifikant voneinander. Die Intelligenz begünstigt die Transferleistungen in beiden Dimensionen signifikant mit einem kleinen Effekt (Gleichgewicht) bzw. mittleren Effekt (Kraftverstärkung). Die Modellaufklärung durch die in das Modell eingefügten Variablen beträgt 6,3 % (GG) und 35,7 % (KV).

Im Modell 2 geht das Wissen der Kinder zum Follow-Up-Messzeitpunkt als zusätzlicher Prädiktor ein (vgl. Forschungsfrage 2). Die Ergebnisse zeigen, dass die Befunde für die zwei verschiedenen Wissensformen unterschiedlich sind: Das konzeptuelle Wissen begünstigt den Transfer sowohl bei Kraftverstärkung (mit einem mittleren Effekt) als auch bei Gleichgewicht (mit einem kleinen Effekt). Das prozedurale Wissen bewährt sich nur als Prädiktor für die Transferleistung bei den Gleichgewichtsaufgaben. Die Modellaufklärung beträgt nun 15,5 % (GG) bzw. 42,0 % (KV).

Zur Erfassung der differenziellen Effekte (vgl. Forschungsfrage 3) wurden in Modell 3 Interaktionseffekte der Gruppenzugehörigkeit mit der Intelligenz bzw. dem Lernzugewinn zwischen dem Prä- und dem Follow-Up-Messzeitpunkt in das Messmodell aufgenommen. Dadurch wird zusätzlich deutlich (vgl. Forschungsfrage 3a), dass sich eine hohe Intelligenz in allen Gruppen positiv auf die Fähigkeit zum Transfer auswirkt; Interaktionseffekte sind v. a. in den beiden Versuchsgruppen im Bereich Kraftverstärkung wirksam, jedoch unterscheiden sich die beiden Versuchsgruppen nicht signifikant. Bei der Analyse der Leistungsveränderungen zwischen dem Prä- und dem Follow-Up-Messzeitpunkt in Interaktion mit der Gruppenzugehörigkeit (Forschungsfrage 3b) zeigt sich, dass das hinzugewonnene konzeptuelle Wissen nur in den Versuchsgruppen und nur in Bezug auf Kraftverstärkung eine signifikant bessere Transferleistung vorhersagt (kleiner Effekt); dabei unterscheiden sich die Versuchsgruppen wiederum nicht signifikant. Ein Wissenszuwachs im prozeduralen Wissen hingegen scheint das Lösen von Transferaufgaben in keiner der drei Gruppen vorherzusagen.

6 Diskussion

Die Regressionsanalysen zeigen über alle Modelle hinweg, dass sich die untersuchten Merkmale unterschiedlich auf die Fähigkeit zum Transfer auswirken, je nachdem, ob es sich um Fragen zu Gleichgewicht oder zu Kraftverstärkung handelt. Bei der Dimension Kraftverstärkung ist der Einfluss des Unterrichts sowie des konzeptuellen Wissens zum Follow-Up-Messzeitpunkt deutlich höher. Dies könnte durch die unterschiedlichen Anforderungen erklärt werden, die den beiden Dimensionen

zugrunde liegen: Die Kraftverstärkungsaufgaben sind stärker auf konzeptuelles Wissen ausgerichtet, die Gleichgewichtsaufgaben überwiegend auf prozedurales Wissen (vgl. Abschn. 4). Hinzu kommt, dass Kraftverstärkung das physikalisch schwierigere Konzept ist, bedingt durch die Unterscheidung der Hebelarme in einen Kraft- und einen Lastarm, die damit verbundene Unterscheidung der wirkenden Kräfte in Kraft und Last sowie die Notwendigkeit, diese Elemente miteinander in Beziehung zu setzen. Die Transferaufgaben zu Kraftverstärkung erfordern daher ein echtes konzeptuelles Wissen – bereits vorhandene strukturelle Wissens Elemente müssen auf ein neues Phänomen angewandt werden. Diese Fokussierung auf Gemeinsamkeiten beim bewussten Vergleich von Phänomenen konnten die Schülerinnen und Schüler der Versuchsgruppen bereits im Unterricht beim Analogen Enkodieren (Mapping) üben. Beim Lösen von Gleichgewichtsaufgaben gibt es hingegen keinen Vorteil der Versuchsgruppen gegenüber der Kontrollgruppe. Eine Erklärung für diesen Befund könnte in der oberflächlichen Ähnlichkeit des Lerninhalts (Wippe) und der Transferaufgabe (Mobile) liegen. Nach Barnett und Ceci (2002) könnte es hier gegebenenfalls ausreichen, auf der Inhaltskomponente die spezifischen Kenntnisse zum Wippen zu übertragen, ohne auf das allgemeine Prinzip rekurrieren zu müssen. Transfer ist bei diesen Aufgaben (bedingt durch die oberflächliche Ähnlichkeit von Wippe und Mobile) offenbar durch Transduktionen (vgl. Abschn. 2) möglich.

Insgesamt zeigen sich keine überzufälligen Unterschiede zwischen den beiden Versuchsgruppen. Für die Fähigkeit, Wissen zu transferieren, scheint es ohne Bedeutung zu sein, ob das Wissen anhand oberflächlich ähnlicher oder unähnlicher Phänomene erworben wurde. Dieses Ergebnis ist analog zu den Ergebnissen zum Wissenserwerb aus dieser Studie (Lohrmann et al. 2014; Schwelle 2016) (vgl. Abschn. 2). Wir vermuten, dass deutlichere Unterschiede zwischen den beiden Versuchsgruppen erneut durch eine gute Unterrichtsqualität nivelliert wurden: Den Kindern wurde eine strukturierte und kognitiv hoch aktivierende Lernumgebung geboten, sodass es in beiden Gruppen gleichermaßen gelungen zu sein scheint, die den Phänomenen zugrundeliegende Tiefenstruktur herauszuarbeiten. Für diese Annahme spricht auch, dass der Einfluss der Intelligenz in den beiden Versuchsgruppen bei der Dimension Kraftverstärkung gleich groß ist.

Pädagogisch zu beachten sind die differenziellen Effekte der Intelligenz, die zeigen, dass die Intelligenz in den Versuchsgruppen einen höheren Effekt auf die Fähigkeit zum Transfer hat als in den Kontrollgruppen. Hier deuten weitere (deskriptive) Analysen (u. a. explorativ anhand eines Mediansplits der Intelligenz) darauf hin, dass dieser signifikante Effekt darauf zurückzuführen ist, dass Kinder mit höherer Intelligenz in den beiden Versuchsgruppen deutlich bessere Leistungen erbringen als jene aus der Kontrollgruppe. Auch bei geringen CFT-Werten sind die Schülerinnen und Schüler der Versuchsgruppen in ihren Transferleistungen der Kontrollgruppe überlegen – der Unterschied zur Kontrollgruppe ist jedoch geringer. Dies deutet unserer Einschätzung nach darauf hin, dass Kinder mit günstigeren Grundfähigkeiten durch die anregende Lernumgebung und die Herausforderungen des Analogen Enkodierens in besonderem Maße gefördert werden, dass es aber keine Benachteiligung der weniger intelligenten Schülerinnen und Schüler in den Versuchsgruppen gibt, da auch diese bessere Transferleistungen erzielen. Analoges Enkodieren hat demnach das Potenzial, alle Kinder zu Transferleistungen anzuregen.

Betrachtet wurde auch der differenzielle Effekt des Lerngewinns auf die Transferleistung, also jene Lernentwicklung, die die Kinder zwischen Prä- und Follow-Up-Messzeitpunkt zeigen. Hier ist bei den beiden Versuchsgruppen nur in Bezug auf die Dimension Kraftverstärkung ein Effekt nachweisbar: Das im Rahmen des Analogenen Enkodierens erworbene Wissen unterstützt die Transferfähigkeit, wohingegen der gleiche Lerngewinn in der Kontrollgruppe (z. B. aufgrund des dreimaligen Bearbeitens der Wissenstests oder anderer, nicht kontrollierbarer Einflüsse) sich nicht auf die Transferfähigkeit auswirkt. Diese Analysen relativieren somit den in der Kontrollgruppe durchaus beachtlichen Lerngewinn. Pointiert lässt sich formulieren: Die Kinder der Versuchsgruppen haben intelligentes, anwendbares und transferierbares Wissen erworben, die Kinder der Kontrollgruppe lediglich träges Wissen. Dieser Befund zeigt die Bedeutsamkeit von lehr-lerntheoretisch fundierten und didaktisch entsprechend gestalteten Lernsituationen im Unterricht: Durch exemplarisches Lernen kann konzeptuelles, intelligentes Wissen aufgebaut werden. Dieses unterscheidet sich in seiner Qualität und seiner Transferierbarkeit auf neue Kontexte von wenig angeleitetem Lernen.

Ausgangspunkt der dargestellten Studie war das Ziel, für Schülerinnen und Schüler interessante und – mit Blick auf intelligentes, anwendbares Wissen – anregende und empirisch überprüfte Lernumgebungen zur Verfügung zu stellen. Die Befunde zeigen, dass sich die – vorrangig in Laborsettings untersuchten – positiven Effekte des Analogenen Enkodierens erfolgreich auf Unterrichtssituationen im Sachunterricht der Grundschule übertragen lassen. Die vorliegende Studie ergänzt die bereits publizierten Befunde zum Wissenserwerb (Lohrmann et al. 2014; Schwelle 2016) durch Erkenntnisse zum nahen Transfer des erworbenen Wissens. Die Übertragung lernpsychologischer Erkenntnisse (wie z. B. zum Analogenen Enkodieren) in die Gestaltung von Unterricht und damit verbunden die Übertragung von Laborbefunden in die ökologische Validität quasi-experimenteller Unterrichtsstudien ist aus unserer Sicht eine wichtige Herausforderung für die Didaktik des Sachunterrichts.

Die Einschränkung auf das naturwissenschaftliche Lernen bzw. das Thema Hebelgesetz in der vorliegenden Studie bedeutet zugleich eine Einschränkung bezüglich der allgemeinen Gültigkeit der Aussagen dieser Studie. Für die Generalisierbarkeit der Befunde wäre die Übertragung auf andere Perspektiven des Sachunterrichts (und über den Sachunterricht hinaus) erforderlich. Diesbezüglich ist auch die fehlende Repräsentativität der vorliegenden Untersuchung zu betonen. Eine wichtige Limitation der vorliegenden Studie ist auch, dass aufgrund des vorliegenden Designs keine Aussagen über die Lernprozesse der Schülerinnen und Schüler zu treffen sind, mithilfe derer gegebenenfalls die Unterschiede bezüglich des Einflusses von konzeptuellem bzw. prozeduralem Wissen auf die Transferleistungen oder auch unsere Erklärungen bzgl. der fehlenden Unterschiede zwischen den beiden Versuchsgruppen überprüft werden könnten. Dieser vertiefende Blick auf das Mapping beim Analogenen Enkodieren könnte Gegenstand zukünftiger Forschungsarbeiten sein.

Danksagung Die Autorinnen und der Autor danken Dr. Veronika Kainz (geb. Schwelle) für die engagierte und professionelle Mitarbeit im Projekt sowie Dr. Alexander Robitzsch vom Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik in Kiel für die methodische Beratung im Rahmen dieser Publikation. Der Dank gilt ferner Christoph Kranich für die Genehmigung zum Abdruck des Fotos (Abb. 2).

Förderung Der Beitrag beruht auf Erkenntnissen eines von der Deutschen Forschungsgemeinschaft mit dem Aktenkennzeichen LO 1706/1-1 geförderten Projekts.

Literatur

- Adams, R. J. (2002). Scaling PISA cognitive data. In R. J. Adams & M. L. Wu (Hrsg.), *PISA 2000 technical report* (S. 99–108). Paris: OECD.
- Adams, R. J., Wilson, M., & Wang, W.-C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21(1), 1–23.
- Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, 48(2), 87–113.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128, 612–637.
- Blanchette, I., & Dunbar, K. (2001). Analogy use in naturalistic settings: the influence of audience, emotion and goals. *Memory and Cognition*, 29(5), 730–735.
- BLK (Bund-Länder-Kommission), & Prenzel, M., et al. (2004). *Sinus – Transfer GS*. Weiterentwicklung des mathematischen und naturwissenschaftlichen Unterrichts an Grundschulen. Heft 112. Kiel: IPN.
- Bohrmann, M., & Möller, K. (2017). Welcher Magnet ist stärker? In H. Giest, A. Hartinger & S. Tänzler (Hrsg.), *Vielperspektivität im Sachunterricht* (S. 91–99). Bad Heilbrunn: Klinkhardt.
- Chen, Z., & Daehler, M. W. (1989). Positive and negative transfer in analogical problem solving by 6-year-old children. *Cognitive Development*, 4, 327–344.
- Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, 47, 177–188.
- Chi, M., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121–152.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanation: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145–182.
- Fong, G. T., & Nisbett, R. E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology: General*, 120(1), 34–45.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, 18, 253–292.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Hrsg.), *Similarity and analogical reasoning* (S. 199–241). New York: Cambridge University Press.
- Gentner, D., & Kurtz, K. J. (2006). Relations, objects, and the composition of analogies. *Cognitive Science*, 30, 609–624.
- Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: No difference without similarity. *American Psychological Society*, 5(3), 152–158.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95(2), 393–408.
- Gerjets, P., Scheiter, K., & Schuh, J. (2008). Information comparisons in example-based hypermedia environments: supporting learners with processing prompts and an interactive comparison tool. *Educational Technology Research and Development*, 56, 73–92.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306–355.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1–38.
- Grygier, P., Jonen, A., Kircher, E., Sodian, B., & Thoermer, C. (2008). „Wissenschaftsverständnis“ und Erwerb von naturwissenschaftlichem Wissen und Experimentierfähigkeit in der Grundschule. In H. Giest & J. Wiesemann (Hrsg.), *Kind und Wissenschaft* (S. 69–81). Bad Heilbrunn: Klinkhardt.
- Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, 104, 427–466.
- Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, 110, 220–263.
- Kiefer, T., Robitzsch, A., & Wu, M. L. (2016). Test Analysis Modules (TAM) (Version 1.995-0 (2016-05-31)). <http://www.edmeasurementsurveys.com/TAM/Tutorials/>. Zugegriffen: 20. Februar 2018.
- Klafki, W. (1962). Didaktische Analyse als Kern der Unterrichtsvorbereitung. In H. Roth & A. Blumenthal (Hrsg.), *Grundlegende Aufsätze aus der Zeitschrift Die Deutsche Schule* (S. 5–34). Hannover: Schroedel.
- Klafki, W. (1985). *Neue Studien zur Bildungstheorie und Didaktik*. Weinheim: Beltz.

- Klauer, K.J. (2011). *Transfer des Lernens. Warum wir oft mehr lernen als gelernt wird*. Stuttgart: Kohlhammer.
- Kline, R. B. (2010). *Principles and practice of structural equation modeling*. New York: Guilford.
- Köhnlein, W. (2015). Aufgaben und Ziele des Sachunterrichts. In J. Kahlert, M. Fölling-Albers, M. Götz, A. Hartinger, S. Miller & S. Wittkowske (Hrsg.), *Handbuch Didaktik des Sachunterrichts* (S. 88–97). Bad Heilbrunn: Klinkhardt.
- Kotovskiy, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, 67, 2797–2822.
- Kupper, L. L., Karon, J. M., Kleinbaum, D. G., Morgenstern, H., & Lewis, D. K. (1981). Matching in epidemiologic studies: Validity and efficiency considerations. *Biometrics*, 37(2), 271–291.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 69–105). Heidelberg: Springer.
- Loewenstein, J., Thompson, L., & Gentner, D. (2003). Analogical learning in negotiation teams: Comparing cases promotes learning and transfer. *Academy of Management Learning and Education*, 2, 119–127.
- Lohrmann, K. (2014). Kontextualisierung und Dekontextualisierung im Unterricht der Grundschule. In W. Einsiedler, M. Götz, A. Hartinger, F. Heinzel, J. Kahlert & U. Sandfuchs (Hrsg.), *Handbuch Grundschulpädagogik und Grundschuldidaktik* (S. 414–418). Bad Heilbrunn: Klinkhardt.
- Lohrmann, K. (2015). Instruktionsorientierter Sachunterricht. In J. Kahlert, M. Fölling-Albers, M. Götz, A. Hartinger, S. Miller & S. Wittkowske (Hrsg.), *Handbuch Didaktik des Sachunterrichts* (S. 408–413). Bad Heilbrunn: Klinkhardt.
- Lohrmann, K., Hartinger, A., & Schwelle, V. (2013). Exemplarisches Lehren und Lernen durch das Arbeiten mit Beispielen – theoretische Bezüge zwischen Allgemeiner Didaktik, Fachdidaktik und Lehr-Lernforschung. *Zeitschrift für Grundschulforschung*, 1(6), 158–171.
- Lohrmann, K., Hartinger, A., Schwelle, V., & Hartig, J. (2014). Die Bedeutung der (Un-)Ähnlichkeit von Beispielen für den Aufbau von konzeptuellem Wissen. *Zeitschrift für Grundschulforschung*, 7(2), 60–73.
- Mähler, C., & Stern, E. (2018). Transfer. In D.-H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 842–852). Weinheim: Beltz.
- Münzer, S., & Brünken, R. (2018). Aptitude-Treatment-Interaktion. In D.-H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 14–20). Weinheim: Beltz.
- Novick, L. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(3), 510–520.
- Prenzel, M., Geiser, H., Langeheine, R., & Lobemeier, K. (2003). Das naturwissenschaftliche Verständnis am Ende der Grundschule. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valentin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 143–187). Münster: Waxmann.
- Richland, L. E., & McDonough, I. M. (2010). Learning by analogy: discriminating between potential analogs. *Contemporary Educational Psychology*, 35, 28–43.
- Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology*, 101(3), 529–544.
- Rost, J. (2004). *Lehrbuch Testtheorie/Testkonstruktion*. Bern: Huber.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research – Online*, 8(2), 23–74. Verfügbar unter: http://www.dgps.de/fachgruppen/methoden/mpr-online/issue20/art2/mpr130_13.pdf. Zugegriffen: 20. Februar 2018.
- Schwelle, V. (2016). *Lernen mit (un-)ähnlichen Beispielen. Zur Bedeutung der Oberflächenstruktur von Beispielen im naturwissenschaftlichen Sachunterricht*. Münster: Waxmann.
- Spreckelsen, K. (1995). Analogieübungen als erste Schritte von Schülern auf dem Wege der Physik. *Naturwissenschaften im Unterricht Physik*, 6(27), 30–32.
- Spreckelsen, K. (2004). Phänomenkreise als Entwicklungskerne für das Verstehen-Lernen. In W. Köhnlein & R. Lauterbach (Hrsg.), *Verstehen und begründetes Handeln* (S. 133–144). Bad Heilbrunn: Klinkhardt.
- Steffensky, M., Kleickmann, T., Kasper, D., & Köller, O. (2016). Naturwissenschaftliche Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In H. Wendt, W. Bos, C. Selzer, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015 Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 133–184). Münster: Waxmann.

- Vosniadou, S. (1989). Analogical reasoning as a mechanism in knowledge acquisition: A developmental perspective. In S. Vosniadou & A. Ortony (Hrsg.), *Similarity and analogical reasoning* (S. 413–437). Cambridge: University Press.
- Weinert, F.E. (1998). Neue Unterrichtskonzepte zwischen gesellschaftlichen Notwendigkeiten, pädagogischen Visionen und psychologischen Möglichkeiten. In Bayerisches Staatsministerium für Unterricht, Kultus, Wissenschaft & Kunst (Hrsg.), *Wissen und Werte für die Welt von morgen* (S. 101–125). Donauwörth: Auer.
- Weiß, R. H. (2006). *CFT 20-R. Grundintelligenztest. Skala 2. Revision*. Göttingen: Hogrefe.
- Wirtz, M., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.