

Empirische Sonderpädagogik, 2020, Nr. 2, S. 112-131
ISSN 1869-4845 (Print) · ISSN 1869-4934 (Internet)

Effekte eines Vorschultrainings zum induktiven Denken auf die Rechenleistung von Förderschulkindern vor und nach Durchführung eines Mathematiktrainings: Ergebnisse einer Mehrebenenanalyse von Einzelfalldaten

Eva Keller¹, Debora Meyer¹, Cornelia Glaser² & Joachim C. Brunstein¹

¹Justus-Liebig-Universität Gießen, ²Pädagogische Hochschule Heidelberg

Zusammenfassung

In einer Stichprobe von Kindern mit dem Förderschwerpunkt Lernen kombinierten wir sequenziell ein Training des induktiven Denkens (TID) mit einem Mathematiktraining. Auf diese Weise prüften wir, ob TID (a) einen unmittelbaren Transfereffekt auf die Leistung bei Mathematikaufgaben erzielt und/oder (b) einen mittelbaren Effekt, indem das Lernen in Mathematik verbessert und die Wirksamkeit des zugehörigen Trainings dadurch gesteigert wird. In der ersten Trainingsphase (8 Sitzungen) wurden 28 Dritt- und Viertklässler per Zufall entweder dem Denktraining (experimentelle Gruppe) oder einem Training sozial-emotionaler Fertigkeiten (Vergleichsgruppe) zugewiesen. Zu je 20 Messzeitpunkten wurden die Denk- und Matheleistungen vor, während und nach den beiden Trainings aufgezeichnet. Die Daten wurden mit Zwei-Ebenen-Modellen analysiert, die der Einzelfallforschung entlehnt wurden. Anders als Kinder der Vergleichsgruppe verbesserten sich Kinder der experimentellen Gruppe während des TID kontinuierlich in ihrer Denkleistung und konnten diesbezüglich substantielle Verbesserungen von der Grundrate zur Post-Trainingsphase erzielen. Jedoch verbesserten sich Kinder, die das TID durchlaufen hatten, weder in ihrer unmittelbar danach erfassten Mathematikleistung noch profitierten sie mehr als Kinder der Vergleichsgruppe von dem anschließend durchgeführten Mathematiktraining.

Schlüsselwörter: Induktives Denktraining, Rechentraining, unmittelbarer Transfer, mittelbarer Transfer, Einzelfallanalyse

Effects of prior training in inductive reasoning on the math performance of special education children before and after participation in a training of math skills: Results of a multilevel analysis of single-case data

Abstract

In a sample of children from special education schools with a focus on learning we combined in a sequential manner a training in inductive reasoning (TIR) with a math training. In this way, we examined if TIR would be able to produce (a) an immediate transfer effect on students' math

performance and/or (b) an indirect transfer effect through its beneficial impact on students' math learning and thereby on the efficaciousness of the respective training program. In the first training phase (8 sessions) 28 third and fourth graders were randomly assigned either to TIR (experimental group) or to a training of social-emotional abilities (comparison group). In the second training phase (8 sessions) all children participated in the math training. At 20 measurement occasions students' reasoning skills and math achievements were assessed before, throughout, and after the two training phases. Data were analyzed with two-level models adapted from research with single case designs. In contrast to students of the comparison group, TIR students significantly improved during the training phase in their reasoning task performance and thus showed substantial performance gains from the baseline to the post-training phase. Yet, TIR students neither improved in their math performance immediately after TIR nor did they benefit from the math training to a greater extent than students in the comparison group.

Keywords: training in inductive reasoning, math training, immediate transfer, indirect transfer, single case research

Einleitung

Schülerinnen und Schüler mit einem sonderpädagogischen Förderbedarf im Bereich Lernen zeigen Defizite in kognitiven Lernvoraussetzungen (Lauth, Brunstein & Grünke, 2014). Eine der wichtigsten kognitiven Voraussetzungen für erfolgreiches Lernen ist die fluide Intelligenz (Roth et al., 2015; Soares, Lemos, Primi & Almeida, 2015). Tests zur Messung der fluiden Intelligenz beinhalten häufig auch Aufgaben zum logisch schlussfolgernden bzw. induktiven Denken (Analogien, Klassifikationen, Folgen oder Matrizen) (Petermann & Petermann, 2010; Weiß & Osterland, 2012). Aufgrund der Bedeutsamkeit der fluiden Intelligenz für das schulische Lernen entwickelte Karl-Josef Klauer (Klauer, 1989) ein Training, das diese Form des logisch-schlussfolgernden Denkens fördern und damit die Lernvoraussetzungen bei lernschwachen Schülerinnen und Schülern auf breiter Front verbessern soll (Klauer, 2001, 2014a).

Konkret versteht Klauer unter induktivem Denken die *Entdeckung von Regelmäßigkeiten* bei gegebenen Objekten (z.B. verbal, numerisch, figural) durch die *Strategie des Vergleichens* (Klauer, 2001). Es geht um die Feststellung von Gleichheit oder Verschiedenheit oder der Gleichheit und Verschiedenheit von Merkmalen oder Relationen (Klauer, 2014a). Aus dieser Definition

ergeben sich nach Klauer (2014a) sechs Aufgabentypen des induktiven Denkens (für eine genaue Darstellung siehe Klauer (2014a)). Das induktive Denkvermögen wird in den Klauerschen Denktrainings durch die „Strategie des Vergleichens“ geschult. Die Schülerinnen und Schüler lernen, Objekte hinsichtlich ihrer Merkmale oder Relationen (Vergleichsaspekt) zu beschreiben. Daraufhin wird die Aufgabe durch systematische Vergleiche der Objekte hinsichtlich der Gleichheit, Verschiedenheit oder der Gleichheit und Verschiedenheit (Vergleichsprozess) der Merkmale oder Relationen gelöst. Im Verlauf des Trainings lernen die Schülerinnen und Schüler den Aufgabenklassen die jeweiligen Beschreibungen und Vergleiche zuzuweisen.

Bisherige Studien konnten bedeutsame Effekte des induktiven Denktrainings auf Tests zur Messung der fluiden Intelligenz nachweisen (Barkl, Porter & Ginns, 2012; Klauer, 1996; Klauer, Willmes & Phye, 2002; Klauer & Phye, 2008; Lenhard & Lenhard, 2011; Marx, 2005; Marx & Keller, 2010; Sonntag, 2004, 2010; Tomic, 1995; Tomic & Klauer, 1996). Die mittlere Stärke der Effekte, welche die Klauerschen Denktrainings auf Tests zur fluiden Intelligenz, wie z.B. den CFT 20-R von Weiß (2006) oder HAWIK-Subtests zum wahrnehmungsgebundenen logischen Denken nach Petermann und Petermann (2010) erzielen, liegt

nach Klauer und Phye (2008) bei $d = 0.60$ (zu einer erweiterten und aktualisierten Bestätigung dieser Wirksamkeit, s. Klauer 2014b). Die Effekte bleiben zeitlich stabil oder verstärken sich bei einer zeitverzögerten Testung noch zusätzlich, sodass von dauerhaften Verbesserungen im induktiven Denken auszugehen ist (Klauer & Phye, 2008; Klauer, 2014a). Zum anderen produziert das Klauer-Training bedeutsame Effekte auf schulische bzw. schulnahe Leistungen (Klauer, 1996; Klauer & Phye, 2008; Klauer, 2014b; Marx, 2006; Marx & Keller, 2010; Sonntag, 2004; Tomic & Klauer, 1996). Begründet wird dies damit, dass die meisten Lernstoffe die Strategie des induktiven Denkens erfordern (Klauer, 2014a). Die Effektstärken bezüglich der Verbesserung schulischer Leistungen liegen laut Klauer und Phye (2008) bei $d = 0.67$.

Im pfadanalytischen Mediationsmodell kann ein Teil des Effekts, den das Klauer-Training auf schulnahe Leistungen erzielt, über die verbesserte Intelligenztestleistung statistisch erklärt werden (Klauer, 2001). Da jedoch in diesem Mediationsmodell der direkte Pfad des Klauer-Trainings auf die Schulleistung meist signifikant bleibt, muss es noch andere unerklärte Einflüsse geben, über die das Klauer-Training seine Wirkung auf das schulische Lernen entfaltet. In Frage kommen die im Rahmen des Trainings erworbenen metakognitiven Fähigkeiten, die eine vertiefte Reflexion und gezielte Steuerung von Lernprozessen in unterschiedlichsten Aufgabengebieten möglich machen (Hasselhorn, 1998). Es wird zudem vermutet, dass die im Denktraining trainierte Fähigkeit, Regelmäßigkeiten zu erkennen, eine strukturierte Speicherung und Abrufleistung von Lerninhalten in nachfolgend durchgeführten Lerneinheiten erleichtert. Zum anderen könnten die schulischen Leistungseffekte auch darauf zurückgehen, dass die im Rahmen des Klauer-Trainings erworbene „Strategie des Vergleichens“ auf neue Lernsituationen und Aufgabenbereiche direkt zu übertragen ist. Dies ist beispielsweise dann der Fall, wenn

die sechs Aufgabenklassen des induktiven Denkens auf numerisches Material angewendet werden (Christou & Papageorgiou, 2007; Tomic & Klauer, 1996). So kann das Feststellen von Gleichheit oder Verschiedenheit das Kategorisieren von Zahlen und Mengen oder das Kategorisieren numerischer Informationen in Textaufgaben erleichtern. Das Erkennen von Relationen zwischen Zahlen unterstützt das Fortsetzen von Zahlenreihen und die Durchführung von Ergänzungsaufgaben bzw. Gleichungen und Ungleichungen.

Effekte induktiver Denktrainings auf schulische Leistungen wurden bisher im Bereich des Schriftspracherwerbs (Klauer, 1996; Marx, 2006; Marx & Keller, 2010), des Erwerbs mathematischer Fähigkeiten (Barkl et al., 2012; Sonntag, 2004; Tomic, 1995; Tomic & Klauer, 1996) und des Erwerbs allgemeiner Gedächtnisstrategien (Sonntag, 2010) untersucht. Potenzielle Effekte, die das Klauer'sche Denktraining auf rechnerische Fähigkeiten ausübt, wurden in zwei unterschiedlichen Situationen untersucht. Einige Studien überprüften im Anschluss an das Denktraining dessen Effekte auf Veränderungen in der Leistung bei ausgewählten Rechenaufgaben. Dadurch wurde überprüft, ob das Denktraining einen unmittelbaren, durch keine weitere Lerngelegenheit vermittelten Transfereffekt auf die Leistung bei Mathematikaufgaben erzielt (Barkl et al., 2012; Tomic, 1995; Tomic & Klauer, 1996). Andere Studien implementierten gleichfalls im Anschluss an das Denktraining eine mathematikrelevante Lehr-/Lernsituation und prüften, ob durch das Denktraining der Erwerb neuen Wissens erleichtert und verbessert wird (Sonntag, 2004). Der so untersuchte Transfereffekt kann in dem Sinne als mittelbar bezeichnet werden, als dass er erst durch den Wissenserwerb in einer neuen Lernsituation festzustellen ist. Kaum untersucht wurde eine spezielle Variante dieses mittelbaren Effekts, nämlich der Einfluss, den das Denktraining auf die Wirksamkeit eines zweiten, nachgeschalteten Trainings schulischer Fähigkeiten

ausübt (für eine Ausnahme im Bereich der Leseförderung s. Klauer, 1996).

Sonntag (2004) untersuchte den hier so bezeichneten mittelbaren Transfereffekt im Kontext des Lösen von Textaufgaben. Förderschüler der Mittel- und Oberstufe profitierten von einer 90-minütigen Unterrichtseinheit zum Thema „Textaufgaben“ stärker und erreichten nachfolgend bessere Rechenleistungen, wenn sie vor der Unterrichtseinheit das Denktraining nach Klauer absolviert hatten. Die Möglichkeit eines unmittelbaren Transfereffekts überprüfte Sonntag nicht. Studien, in denen die mittelbaren Effekte des Denktrainings auf ein anschließend implementiertes Mathe-Training überprüft worden wären, kennen wir nicht. Die Daten zum unmittelbaren Transfer sind wiederum uneinheitlich: In einigen Studien, die mit Gruppen von Regelschulkindern durchgeführt wurden, ließ sich ein unmittelbarer Transfereffekt auf die Mathematikleistung zuverlässig nachweisen (Tomic & Klauer, 1996); andere Studien verliefen diesbezüglich weniger erfolgreich (Barkl et al., 2012; Tomic, 1995). Generell fällt auf, dass die jeweils untersuchten Transfereffekte indirekt aus dem Vergleich von Vor- und Nachtestungen erschlossen werden, während Lernfortschritte, die im Verlauf der jeweils eingesetzten Intervention eintreten mögen, eigentlich gar nicht registriert werden. Sofern durch das Denktraining das Erlernen anderer bzw. neuer Sachverhalte tatsächlich erleichtert wird, wären gerade Daten zum Lernverlauf (und nicht nur zur Abschlussleistung) hierfür informativ. Methodische Ansätze aus dem Bereich der quantitativ experimentellen Einzelfallforschung, in der vor und nach, vor allem aber während der Intervention sehr häufig gemessen wird, könnten diesbezüglich aufschlussreich sein (Kazdin, 2011).

Zusammenfassend sind uns keine Studien bekannt, in denen (a) sowohl unmittelbare als auch mittelbare Transfereffekte des Denktrainings (nach Klauer) gemeinsam innerhalb der gleichen Studie überprüft wurden; (b) dem Training des induktiven Den-

kens ein regelrechtes Training rechnerischer Fähigkeiten nachgeschaltet wurde; (c) Verlaufsdaten zu den Lernfortschritten während der jeweiligen Trainingsphase erhoben wurden; und (d) etwaige unspezifische Transfereffekte des Denktrainings untersucht und ausgeschlossen wurden.

Fragestellung und Hypothesen

Mit der vorliegenden Studie sollte überprüft werden, ob durch das Training des induktiven Denkens die Wirksamkeit eines sich daran anschließenden Trainings rechnerischer Fähigkeiten erhöht werden kann. Das Training des induktiven Denkens sollte die Klassifikation von Aufgabentypen und Lösungsprozeduren und deren strukturierte Speicherung und Abrufleistung im Rechen-Training unterstützen. Dies wiederum sollte sich positiv auf das Lösen von Rechenaufgaben auswirken.

Die vorgenannten Punkte wurden dabei wie folgt berücksichtigt: (a) Unmittelbare und mittelbare Transfereffekte wurden gemeinsam in ein und demselben Kontext überprüft. (b) Im Unterschied zu Studien, in denen eng umrissene Lehrstunden als Transfersituationen herangezogen wurden, erfolgte die Untersuchung von (mittelbaren) Transfereffekten in dieser Studie anhand eines zeitlich ausgedehnten Rechen-Trainings. (c) Um beurteilen zu können, ob durch das Denktraining das Lernen mit Mathematikaufgaben verbessert wird, wurden nach dem Vorbild von Einzelfallstudien in allen Phasen der Untersuchung (vor, während und nach dem jeweiligen Training) Verlaufsdaten zum Denken und Rechnen erhoben; diese Daten wurden anschließend mit einem dafür geeigneten Modell der Mehrebenenanalyse analysiert. (d) Ob das Denktraining teils auch unspezifische Transfereffekte produziert, wurde durch Aufnahme eines alternativen Vortrainings zur Einübung sozial-emotionaler Kompetenzen in den Untersuchungsplan geprüft.

Mit diesen Zielsetzungen führten wir eine Einzelfallstudie mit 28 Dritt- und Viert-

klässlern des Förderschwerpunkts Lernen durch. Kinder der Experimentalgruppe durchliefen zunächst ein vierwöchiges Denktraining und daran anschließend ein gleichfalls vierwöchiges Rechentraining. Kinder der Vergleichsgruppe erhielten zuerst ein Training sozial-emotionaler Kompetenzen, gefolgt von dem gleichen Rechentraining. Alle Trainings wurden durch geschulte Trainerinnen angeleitet, wobei das Denktraining (Lenhard, Lenhard & Klauer, 2012) und das Rechentraining (Lenhard & Lenhard, 2009) PC-basiert durchgeführt und in einen gemeinsamen Makrokontext eingebettet wurden („Elfe und Mathis“). In dem Rechentraining wurden Rechenfertigkeiten in folgenden Bereichen geschult: Additions- und Subtraktionsaufgaben, Ergänzungsaufgaben sowie Gleichungen und Ungleichungen. Die Daten zur Denk- und zur Rechenleistung wurden in fünf Untersuchungsphasen mehrfach erhoben. Dies erfolgte in der Form eigens dafür entwickelter Kurztests, die standardisierten Verfahren entlehnt wurden. Eingesetzt wurden die Tests vor, während und nach dem Denktraining sowie vor, während und nach dem Rechentraining, wobei die Phase nach dem Denktraining der Phase vor dem Rechentraining entsprach. Alle Trainingseffekte wurden sowohl auf der Ebene der innerhalb der jeweiligen Trainingsphase ermittelten Verlaufstrends (Leistungsveränderungen über die Trainingssitzungen) als auch auf der Ebene mittlerer Niveauunterschiede im Prä-/Post-Vergleich erhoben (d.h. durch den Vergleich einer Phase vor mit einer Phase nach dem jeweiligen Training).

Auf dieser Grundlage formulierten wir fünf Hypothesen, die sich auf drei Arten von Fragestellungen bezogen und im Einzelnen wie folgt lauteten:

Trend- und Niveaueffekte des Denktrainings auf Tests zum induktiven Denken

H1: Während des Denktrainings werden Schülerinnen und Schüler einen signifikanten Leistungszuwachs (bzw. positiven Trend) in Tests zum induktiven

Denken zeigen. Bei Schülerinnen und Schülern der Vergleichsgruppe wird im Zeitraum des alternativen sozial-emotionalen Trainings kein vergleichbarer Zuwachs erwartet.

H2: Schülerinnen und Schüler, die am Denktraining teilgenommen haben, werden beim Vergleich der Vortestphase mit (a) der Phase der unmittelbaren Nachtestung und (b) der Phase der verzögerten Nachtestung jeweils ein höheres Niveau in ihrer Denkleistung zeigen. Bei Schülerinnen und Schülern der Vergleichsgruppe werden keine vergleichbaren Niveau-Unterschiede erwartet.

Unmittelbarer Transfereffekt des Denktrainings auf die Leistung in den Rechentests

H3: Schülerinnen und Schüler, die am Denktraining teilgenommen haben, werden ihr Leistungsniveau beim Lösen von Rechenaufgaben von der Vor- zur Nachtrainingsphase steigern. Bei Schülerinnen und Schülern der Vergleichsgruppe wird keine vergleichbare Steigerung erwartet.

Mittelbare Transfereffekte des Denktrainings auf die Wirksamkeit des Rechentrainings

H4: Schülerinnen und Schüler, die das Denktraining absolviert haben, werden während des Rechentrainings größere Lernfortschritte zeigen und daher einen stärkeren positiven Trend in ihrer Rechenleistung aufweisen als Schülerinnen und Schüler, die das alternative Vortraining absolviert haben.

H5: Schülerinnen und Schüler, die das Denktraining durchlaufen haben, werden von dem anschließend durchgeführten Rechentraining stärker profitieren als Schülerinnen und Schüler, die das sozial emotionale Vergleichstraining absolviert haben. In der erstgenannten Gruppe sollten die durch das Rechentraining bewirkten Niveauunterschiede im Prä-/Post-Vergleich da-

her stärker als in der zweitgenannten Gruppe ausgeprägt sein.

Um die Zufriedenheit und Akzeptanz des Trainings bei den teilnehmenden Kindern zu untersuchen, wurden Aspekte der sozialen Validität geprüft.

Methode

Stichprobe und Design

An der Untersuchung nahmen 12 Schülerinnen und 16 Schüler der 3. und 4. Klasse mit einem sonderpädagogischen Förderbedarf im Bereich Lernen teil. Das mittlere Alter lag bei $M = 9.75$ Jahren ($SD = 0.70$). Im Rahmen der Rekrutierung wurden drei Förderschulen kontaktiert, denen das Trainingskonzept vorgestellt wurde. Nach Zustimmung zur Teilnahme wurden die Klassenlehrkräfte gebeten, Kinder für die Förderung auszuwählen, die aufgrund ihrer Voraussetzungen im Lern- und Sozialverhalten (z.B. ausreichende Motivation und Ausdauer) am ehesten von den vorgestellten Trainings profitieren könnten. Vor der Teilnahme wurde das Einverständnis der Erziehungsberechtigten eingeholt. Im Rahmen einer Vortestung, die vier Wochen vor Beginn der interventionsbezogenen Erhebungen in einem Einzel-Setting stattfand, bearbeiteten die teilnehmenden Kinder die HAWIK-Untertests zum „wahrnehmungsgebundenen logischen Denken“ (WLD) nach Petermann und Petermann (2010). Aufgrund ihrer WLD-Testleistung wurden die Kinder in eine aufsteigende Rangfolge gebracht. Je zwei Kinder, die in der Rangfolge unmittelbar aufeinanderfolgten, wurden zu Paaren zusammengefasst. Innerhalb eines Paares wurde per Zufall bestimmt, welcher von zwei Bedingungen je ein Kind zugewiesen wurde. Kinder der Experimentalgruppe (EG) durchliefen als Vorschalttraining ein vierwöchiges Training zum induktiven Denken (zwei Sitzungen pro Woche), an das sich ein vierwöchiges Rechentraining (gleich-

falls zweimal wöchentlich) anschloss. Die Vergleichsgruppe (VG) durchlief ein alternatives (Vor-)Training zur Förderung sozial-emotionaler Kompetenzen, auf das dann das Rechentraining folgte. Aufgrund des oben beschriebenen Matching-Verfahrens bestand im WLD-Test zwischen der Experimentalgruppe (Wertpunktsumme: $M = 18.21$, $SD = 6.55$, mittlerer Rang = 11.04, Range = 0.10 bis 39.50) und der Vergleichsgruppe (Wertpunktsumme: $M = 17.57$, $SD = 7.59$, mittlerer Rang = 10.04, Range = 0.10 bis 50.00) praktisch kein Unterschied. Um die beiden Bedingungen hinsichtlich der anfänglichen Rechenleistung miteinander vergleichen zu können, bearbeiteten die Kinder zudem den „Heidelberger Rechentest“ (HRT) von Haffner, Baro, Marx und Schneider (2005). Hinsichtlich des Gesamtwerts (T-Skala) ergaben sich auch hier keine Unterschiede zwischen Experimentalgruppe ($M = 25.71$, $SD = 6.35$, Range = 17.00 bis 35.00) und Vergleichsgruppe ($M = 27.32$, $SD = 4.82$, Range = 17.00 bis 36.00).

Um die Effekte der Förderung zu untersuchen, wurde ein Einzelfalldesign gewählt (siehe Tabelle 1), bei dem relativ wenige Teilnehmer zu möglichst vielen Messzeitpunkten untersucht werden (Kazdin, 2011). Vorliegend wurden zwei A-B-Designs mit einer unmittelbaren und verzögerten Nachtestung kombiniert. Zwei Wochen vor Beginn des Vorschalttrainings wurden in einer vier Messzeitpunkte (zwei pro Woche) umfassenden Grundrate (Phase 1) die induktive Denkleistung und die Rechenleistung erfasst. Während des vierwöchigen Vorschalttrainings (Phase 2: zweimal pro Woche) wurde zu acht Messzeitpunkten die induktive Denkleistung erfasst. Daran anschließend folgte in einer zweiwöchigen Testungsphase (Phase 3) wiederum zu vier Messzeitpunkten die unmittelbare Nachtestung für das Denktraining und die unmittelbare Vortestung für das Rechentraining. In der daran anschließenden vierwöchigen Phase des Rechentrainings (Phase 4) wurde zu acht Messzeitpunkten die Rechenlei-

tung der Schülerinnen und Schüler erfasst. Nach dem Rechentraining erfolgte in einer zweiwöchigen Testungsphase (Phase 5) die unmittelbare Nachtestung für das Rechentraining sowie eine verzögerte Nachtestung für das Denktraining. Zur Sicherung der internen Validität wurden alle Testungen sowohl mit den Schülerinnen und Schülern der Experimentalgruppe als auch mit denen der Vergleichsgruppe durchgeführt. In jeder der drei Schulen wurden beide Untersuchungsbedingungen implementiert und jede Trainerin trainierte Kinder in beiden Untersuchungsbedingungen.

Ablauf der Test- und Trainings-sitzungen

Die 2 x 8 45-minütigen Trainingssitzungen in den Phasen 2 und 4 fanden in Schülerpaaren statt und folgten einem einheitlichen Ablauf. Zunächst gab es eine zehnminütige Einführungsphase. In der ersten Sitzung wurde den Kindern in dieser Phase das Token-System erläutert. Für das Einhalten bestimmter Verhaltensregeln („Ich lasse den anderen ausreden“, „Ich lasse den anderen in Ruhe arbeiten“ und „Ich bearbeite die Aufgaben ordentlich“) konnten die Schülerinnen und Schüler in jeder Sitzung insgesamt sechs Punkte sammeln. Die Punkte wurden am Ende des Denktrainings und des

Rechentrainings gegen eine Belohnung in Form von Schreibmaterialien eingetauscht. In den darauffolgenden Sitzungen wurden den Kindern in der Einführungsphase anhand einer Lernkurve Rückmeldungen zu ihrem Abschneiden in den Testungen zum induktiven Denken bzw. zum Rechnen erteilt. Nach der Einführungsphase folgte für 20 Minuten das eigentliche Training (Denktraining oder sozial-emotionales Training sowie Rechentraining). Danach wurde wiederum eine zehnminütige Testung (je nach Phase: induktives Denken oder Rechnen) durchgeführt. Anhand eines Protokollbogens erhielten die Schülerinnen und Schüler abschließend eine Rückmeldung zu ihrem Verhalten während des Trainings.

Auf drei fünfstufigen Skalen (1 = trifft überhaupt nicht zu bis 5 = trifft voll und ganz zu) beurteilten die Schülerinnen und Schüler jede Trainingssitzung nach folgenden Aspekten der sozialen Validität: Spaß, Lernzuwachs und Schwierigkeit der Lerninhalte. Über alle Sitzungen hinweg ergaben sich folgende Durchschnittswerte: Spaß (Denktraining: $M = 4.91$, $SD = 0.13$; sozial-emotionales Training: $M = 4.57$, $SD = 0.42$; Rechentraining: $M = 4.64$, $SD = 0.53$), Lernzuwachs (Denktraining: $M = 4.81$, $SD = 0.40$; sozial-emotionales Training: $M = 4.86$, $SD = 0.23$; Rechentraining: $M = 4.64$, $SD = 0.64$) und Schwierigkeit

Tabelle 1: Design der Studie

Untersuchungsphase	Messzeitpunkt	Leistungsmessung
Phase 1: Grundrate	1 - 4	Induktives Denken Rechenaufgaben
Phase 2: Vorschalttraining (Training zum induktiven Denken vs. sozial-emotionales Training)	5 - 12	Induktives Denken Rechenaufgaben
Phase 3: Unmittelbarer Nachtest für Denktraining/ Unmittelbarer Vortest für Rechentraining	13 - 16	Induktives Denken Rechenaufgaben
Phase 4: Rechentraining	17 - 24	Induktives Denken Rechenaufgaben
Phase 5: Unmittelbarer Nachtest für Rechentraining/ Verzögerter Nachtest für Denktraining	25 - 28	Induktives Denken Rechenaufgaben

(Denktraining: $M = 4.71$, $SD = 0.71$; sozial-emotionales Training: $M = 4.52$, $SD = 0.57$; Rechentraining: $M = 4.52$, $SD = 0.63$).

Die reinen Testsitzungen in den Phasen 1, 3 und 5 folgten dem gleichen Ablauf: Nach einer kurzen Einführung wurden zwei zehnmünütige Kurztestungen zur Erfassung der Rechenleistung und der induktiven Denkleistung durchgeführt. Um Reihenfolgeeffekte zu vermeiden, wurde die Abfolge der beiden Testarten innerhalb einer Sitzung zwischen den Messzeitpunkten randomisiert. Zudem erhielt jedes Kind über alle Messzeitpunkte hinweg die Tests in einer unterschiedlichen Reihenfolge. Zwischen den Tests gab es eine fünfminütige Pause. Jede Sitzung endete mit einer kurzen Spielphase.

Trainingsverfahren

Als Vorschalttraining wurden in der experimentellen Bedingung die „Denkspiele mit Elfe und Mathis“ für das Vor- und Grundschulalter (Lenhard et al., 2012) eingesetzt. Die Vergleichsgruppe erhielt im gleichen Zeitraum ein alternatives Training sozial-emotionaler Kompetenzen (Aebi, Perriard, Stiffler Scherer & Wettach, 2012). Anschließend wurden mit den Kindern beider Bedingungen die „Rechenspiele mit Elfe und Mathis I“ für Kinder der ersten bis dritten Jahrgangsstufe (Lenhard & Lenhard, 2009) durchgeführt.

Training des induktiven Denkens

Die Denkspiele mit Elfe und Mathis I (Lenhard et al., 2012) sind ein computergestütztes Programm zur Förderung der sechs Kernaufgaben des induktiven Denkens. Für die vorliegende Förderstudie wurde die Version für das Vor- und Grundschulalter ausgewählt, da es sich um eine Stichprobe mit Lernrückständen handelte und das Aufgabenmaterial nicht zu schwierig sein sollte. Das Programm basiert auf dem bewährten Training von Karl Josef Klauer (Klauer, 1989). Im Training wurde eingeübt, die Stra-

tegie des systematischen Vergleichens auf vielfältiges Aufgabenmaterial anzuwenden.

Das Training war derart gestaltet, dass es von den Schülerinnen und Schülern weitestgehend selbstständig absolviert werden konnte. Die Hinführung und Erarbeitung neuer Aufgabentypen bzw. Denkopoperationen erfolgte in Schülerpaaren. Die Trainerinnen klärten zentrale Begriffe (z.B. „Was ist ein Merkmal?“). Danach modellierten sie das Vorgehen beim Lösen der Aufgabenklasse, indem sie die spezifischen Charakteristika der dafür geeigneten Denkopoperation und das Lösungsvorgehen laut verbalisierten (1. Erkennen der Aufgabenstellung; 2. Analyse der Ausgangslage und Problemidentifikation; 3. Hypothesenbildung hinsichtlich des vermuteten Ergebnisses, der vermuteten Aufgabenklasse oder der Lösungsprozedur; 4. Hypothesenprüfung durch systematisches und zielgerichtetes Vergleichen; 5. Selbstkontrolle). Anschließend leiteten die Trainerinnen die Kinder in der selbstständigen Anwendung der Lösungsstrategie an. Dabei gingen die Trainerinnen nach der Methode des „gelenkten Entdeckenlassens“ vor, d.h. sie erteilten den Kindern während der Aufgabenbearbeitung Denk- und Strukturierungshilfen und leiteten sie zur Selbstinstruktion an, ohne die Aufgabenlösung selbst vorwegzunehmen.

Lenhard und Lenhard (2011) ermittelten mittlere bis große Effekte des Denktrainings auf einen Test zur Messung der fluiden Intelligenz (Subtests zum wahrnehmungsgebundenen logischen Denken aus dem HAWIK von Petermann und Petermann (2010)) bei Förderschülern der dritten Klasse.

Training sozial-emotionaler Kompetenzen

Die Vergleichsgruppe erhielt in der Vorschaltphase ein alternatives Training, mit dem etwaige unspezifische Trainingseffekte und Effekte des wiederholten Testens kontrolliert werden sollten. Eingesetzt wurden Übungen aus dem „Baghira-Training“ für Kinder mit oppositionellem und aggressivem Verhalten von Aebi et al. (2012) sowie dem Verhaltenstraining für Grundschulkin-

der von Petermann, Koglin, Natzke und Mares (2013). Beide Trainings zielen auf den Aufbau sozial kompetenten Verhaltens ab. Eingesetzt wurden Aufgaben zur Selbst- und Fremdwahrnehmung von Gefühlen, zum Umgang mit Wut, zu Strategien der Impuls- und Ärgerkontrolle sowie zu Strategien der Konfliktlösung.

Rechenttraining

Als Rechentraining wurden die „Rechenspiele mit Elfe und Mathis I“ für Kinder der ersten bis dritten Jahrgangsstufe von Lenhard und Lenhard (2009) verwendet. Da die teilnehmenden Kinder aus dem Förderschwerpunkt Lernen inhaltliche Lernrückstände von zwei bis drei Schuljahren aufwiesen, wurden die Rechenspiele für jüngere Schülerinnen und Schüler eingesetzt. Das Training umfasste 17 Aufgabenblöcke aus den Bereichen Mengen (Mengen erfassen, vergleichen, aufteilen und Zahlen zuordnen), Zahlen (Zahlendarstellungen und Zahlbeziehungen/Relationen), Sachaufgaben (Anwendung mathematischer Fertigkeiten auf quasi-realistische Szenarien), Bilder (räumliches Vorstellungsvermögen) und Rechnen (Addition, Subtraktion). Alle Inhaltsbereiche wurden in drei Schwierigkeitsstufen durchgeführt und an das Fähigkeitsniveau des einzelnen Kindes angepasst. Auch dieses Training wurde von den Schülerinnen und Schülern weitestgehend selbstständig absolviert, wobei die Kinder die Aufgaben aus allen Bereichen auf verschiedenen Niveaustufen durchliefen.

Wie im Manual empfohlen, wurde das Vorgehen für jeden Aufgabentyp zunächst in Schülerpaaren modelliert, indem die Kinder Selbstinstruktionen für das Lösen der Aufgaben erlernten. Daraufhin arbeiteten die Kinder individuell am Computer. Bei auftretenden Schwierigkeiten gaben die Trainerinnen Hilfestellungen, wobei wiederum auf die Methode des „gelenkten Entdeckenlassens“ zurückgegriffen wurde.

Abhängige Variablen

Induktives Denken

Für die Erfassung des induktiven Denkens wurden - in Entsprechung zur Gesamtzahl der Messungen - 20 Tests erstellt, die aus Aufgabenmaterial des CFT-20-R (Weiß, 2006) bestanden. Jeder Test beinhaltete vier Subtests (Reihen fortsetzen, Unpassendes streichen, Topologien, Matrizen) mit jeweils vier Aufgaben, sodass jeder Test 16 Aufgaben umfasste. Alle Aufgaben wurden mit 0 oder 1 (korrekt gelöst) bewertet, sodass in jedem Test maximal 16 Punkte erreicht werden konnten.

Die zu den Subtests existierenden Aufgaben aus dem CFT-20-R wurden randomisiert den 20 Tests zugeordnet. Dazu wurden die Aufgaben pro Kategorie nummeriert und den Kurztests zufällig zweimal zugewiesen. Um Reihenfolgeeffekte zu vermeiden und die interne Validität zu stärken, wurde die Reihenfolge der vier Subtests in jedem Test variiert. Des Weiteren wurde die Zuordnung der 20 Tests zu den 20 Messzeitpunkten zwischen den Schülerinnen und Schülern derart variiert, dass die Reihenfolge der Tests innerhalb einer Bedingung unterschiedlich ausfiel, zwischen den Bedingungen aber gleich blieb. Dazu wurde jeweils für ein anhand des Tests zum wahrnehmungsgebundenen logischen Denken (Petermann & Petermann, 2010) gebildetes Schülerpaar, das aus einem Kind der Experimentalgruppe und einem Kind der Vergleichsgruppe bestand, die Abfolge der 20 Tests per Zufall festgelegt. So wurde sichergestellt, dass die Schülerinnen und Schüler innerhalb einer Bedingung zu einem Messzeitpunkt unterschiedliche Tests absolvierten, aber die Testungen zu einem Messzeitpunkt zwischen den Bedingungen miteinander vergleichbar waren. Alle Testungen folgten einer standardisierten Instruktion. Vor jeder Testung wurde für jeden Subtest eine Beispielaufgabe mit den Schülerinnen und Schülern besprochen. Die Inhaltsvalidität der erstellten Tests wurde durch die oben beschriebenen Konstruktionsprinzipi-

en der Kurztestungen sichergestellt. Um die Konstruktvalidität der erstellten Tests zu prüfen, wurde die Korrelation zwischen dem mittleren Wert der vier Grundratentests mit dem Rohwert des wahrnehmungsgebundenen logischen Denkens (Petermann & Petermann, 2010) berechnet. Die Korrelation betrug $r = 0.45$. Die interne Konsistenz der vier Grundratentests lag bei $\alpha = 0.66$.

Rechenleistung

Für die Erfassung der Rechenleistung wurden ebenfalls 20 Tests erstellt. Jeder Test bestand aus vier Subtests (Rechenaufgaben mit Addition und Subtraktion, einfache Ergänzungsaufgaben mit Rechenoperation auf einer Gleichungsseite, komplexe Ergänzungsaufgaben mit Rechenoperationen auf beiden Gleichungsseiten sowie Gleichungen bzw. Ungleichungen) mit jeweils acht Aufgaben, sodass insgesamt 32 Punkte erreicht werden konnten. In Anlehnung an den DEMAT I (Krajewski, Küspert & Schneider, 2002) und den DEMAT II (Krajewski, Liehm & Schneider, 2004) wurde ein Aufgabenpool erstellt, aus dem die Aufgaben randomisiert den Testungen zugeordnet wurden. Um Reihenfolgeeffekte zu vermeiden, wurde bei allen 20 Tests die Abfolge der vier Subtests variiert. Zudem wurde die Reihenfolge der Testungen bei allen Schülerinnen und Schülern variiert, indem die ausgelosten Abfolgen der 20 Tests den für die beiden Bedingungen jeweils gebildeten Schülerpaaren zugewiesen wurden. Die Inhaltsvalidität der erstellten Kurztestungen wurde durch die Konstruktionsprinzipien gewährleistet. Zur Überprüfung der Konstruktvalidität wurde die Korrelation zwischen dem mittleren Testwert der vier Grundratentests mit dem Rohwert des Heidelberger Rechen-tests (Haffner et al., 2005) berechnet. Die Korrelation betrug $r = 0.80$. Die interne Konsistenz der vier Grundratentests lag bei $\alpha = 0.93$. Die Korrelation zwischen den Durchschnittsleistungen der vier Grundratentests im Denken und Rechnen betrug $r = 0.67$.

Treatment-Integrität

Die Treatment-Integrität wurde durch drei Verfahren sichergestellt:

Alle Trainerinnen durchliefen eine Schulung, in der sie den Ablauf und die Umsetzung der einzelnen Trainingskomponenten einübten. Den Trainerinnen lag ein Manual vor, in dem der Ablauf jeder Sitzung sowie das gewünschte Instruktionsverhalten detailliert beschrieben wurden. Nach jeder Trainingssitzung fand eine Rücksprache mit der Trainingsleiterin (Erstautorin) statt, in der Fragen und aufgekommene Schwierigkeiten besprochen wurden. Zudem machten alle Trainerinnen für jede Trainingssitzung Angaben zur Vollständigkeit der durchgeführten Trainingsmaßnahmen. Dazu verwendeten sie Checklisten, auf denen sie für jede Sitzung angaben, ob alle Phasen (Einführung, Training, Testung, Verhaltensfeedback und soziale Validität) umgesetzt wurden. Gemäß dieser Selbstbeurteilung zeigte sich, dass alle Trainerinnen alle Trainingsphasen umgesetzt hatten. Für das Denktraining wurde die Qualität der Durchführung der Trainingssitzungen von drei unabhängigen Beobachterinnen (Psychologiestudierende) erfasst. Das Denktraining wurde aufgrund der in den Schülerpaaren gemeinsam stattfindenden Erarbeitung neuer Aufgabentypen und der Modellierungs-Phasen als instruktionsintensiver eingeschätzt. Jede Trainerin wurde im Trainingsverlauf dreimal während der Trainingsdurchführung beobachtet. Die Beobachterinnen schätzten das Instruktionsverhalten hinsichtlich der Klarheit der Vermittlung, der Verhaltensbeobachtung, der Individualisierung und der Motivierung auf einer fünfstufigen Skala (1 = Umsetzung gar nicht gelungen bis 5 = Umsetzung hervorragend gelungen) ein (Helmke, 2006). Über alle Trainerinnen hinweg ergaben sich folgende Werte: Klarheit: $M = 2.00$, $SD = 1.13$; Monitoring: $M = 1.50$, $SD = 0.55$; Individualisierung: $M = 1.77$, $SD = 0.67$; Motivierung: $M = 1.60$, $SD = 0.54$.

Statistische Analysen

Die Auswertung der Einzelfall- bzw. Messwiederholungsdaten erfolgte mit Hilfe des Statistikprogramms SAS University Edition. Fehlende Werte wurden mittels multipler Imputation anhand des EM-Algorithmus (Anzahl an Imputationen = 25) fallweise geschätzt (Peng & Chen, 2018). Die Analysen wurden zunächst für jede Imputation einzeln durchgeführt (BY IMPUTATION) und daran anschließend durch die Prozedur PROC MIANALYZE zusammengeführt. Auf diese Weise ergab sich für jede Analyse ein über alle 25 Imputationen gepooltes Ergebnis. Für die Darstellung der deskriptiven Statistiken wurden ebenfalls die über 25 Imputationen gemittelten Kennwerte verwendet. Der Anteil fehlender Daten, der durch krankheitsbedingte Fehlzeiten der Kinder entstand, betrug 8.92% bei den Denktests und 11.96% bei den Rechentests. Kinder, die über alle Phasen hinweg mehr als 50% fehlende Werte aufwiesen, wurden von der Analyse ausgeschlossen. Beim induktiven Denken gingen die Daten aller 28 Schülerinnen und Schüler in die Analyse ein. Für die Analyse der Daten zur Rechenleistung wurde ein Schüler der Experimentalgruppe aufgrund eines zu hohen Anteils fehlender Werte ausgeschlossen.

Die Auswertung wurde an das von Moeyaert, Ugille, Ferron, Beretvas und van den Noortgate (2014) beschriebene Vorgehen zur Anwendung von Mehrebenen-Regressions-Ansätzen auf Einzelfalldaten angelehnt. Dazu wurden die Daten in einem Zwei-Ebenen-Modell (Ebene 1: Messzeitpunkte; Ebene 2: Personen) analysiert (Moeyaert, Ferron, Beretvas & van den Noortgate, 2014; Moeyaert, Ugille et al., 2014). Um Aussagen über (a) mittlere Niveauunterschiede zwischen zwei Phasen sowie (b) über Trends bzw. Steigungen innerhalb einer Phase treffen zu können, wurden für die Daten zum induktiven Denken und zur Rechenleistung mehrere Regressionsmodelle aufgestellt (PROC MIXED). Jedes Modell bildete die vier untersuchten Phasen des

Trainings ab. Dementsprechend wurden drei Dummy-Variablen kodiert, die jeweils den mittleren Unterschied zwischen den Werten zweier Phasen anzeigten. Zudem wurden für jedes Modell drei Zeit-Variablen kodiert, die den Trend der Daten innerhalb einer Phase abbildeten. Autokorrelationen in den Daten wurden für jede Untersuchungsphase modelliert. Die Auswertung von Einzelfalldaten mit 2-Ebenen-Modellen zeigt laut Simulationsstudien bereits ab sechs Personen mit mindestens sechs Messzeitpunkten eine ausreichende Power (> 0.80), um starke Treatmenteffekte nachweisen zu können (Lobo, Moeyaert, Baraldi Cunha & Babik, 2017).

Daten zum induktiven Denken. Dazu wurden zwei Modelle aufgestellt. Die Phasen waren die Grundrate, das Vorschalttraining sowie die unmittelbare und die verzögerte Nachtstung für das Denktraining. Für die Modellierung der vier Phasen wurden drei Dummy-Variablen gebildet, die jeweils den Wechsel von einer Phase zur nächsten Phase anzeigten ($d_1 = 0-1-0-0$; $d_2 = 0-0-1-0$; $d_3 = 0-0-0-1$). Zur Schätzung der Trends in den einzelnen Phasen wurden vier Zeit-Variablen kodiert: T0 wurde auf den Start der Grundrate zentriert, stieg in eben dieser an und wurde ab der zweiten Phase (Vorschalttraining) konstant gehalten. T1 wurde auf den Start der Vorschalttrainingsphase zentriert und stieg in dieser an, bevor sie in den nachfolgenden Phasen konstant gehalten wurde. T2 wurde auf den Beginn der unmittelbaren Nachtstung zum Denktraining zentriert, stieg in dieser an und wurde in der letzten Phase (verzögerte Nachtstung für das Denktraining) konstant gehalten. T3 wurde auf den Beginn der verzögerten Nachtstung für das Denktraining zentriert und stieg in dieser Phase an. Das erste Modell (M1), in das nur die Dummy-Variablen eingefügt wurden, lieferte Aussagen über die mittleren Unterschiede zwischen einer ausgewählten Phase und der Grundrate. Das zweite Modell (M2), in das die Zeitvariablen zusätzlich aufgenommen wurden,

zeigte durch die Interaktionsterme ($T \times d$) den Trend in den einzelnen Phasen an.

$$\text{Modell 1: } Y = \beta_0 + \beta_1 * d_1 + \beta_2 * d_2 + \beta_3 * d_3$$

$$\begin{aligned} \text{Modell 2: } Y = & \beta_0 + \beta_1 * T_0 + \beta_2 * d_1 + \\ & \beta_3 * T_1 * d_1 + \beta_4 * d_2 + \\ & \beta_5 * T_2 * d_2 + \beta_6 * d_3 + \\ & \beta_7 * T_3 * d_3 \end{aligned}$$

Daten zur Rechenleistung. Aufgestellt wurden drei Modelle. Die untersuchten Phasen waren die Grundrate, die unmittelbare Vortestung für das Rechentraining, das Rechentraining sowie die unmittelbare Nachtstung für das Rechentraining. Um die unmittelbare Nachtstung für das Rechentraining mit den vorangehenden Phasen vergleichen zu können, wurde zusätzlich zu den bereits existierenden Dummy-Variablen (d_1 bis d_3) eine vierte Variable d_4 erstellt ($d_4 = 1-0-0-0$). Die Zeitvariablen wurden aus der Analyse zum induktiven Denken übernommen. Das erste Modell (M1) erlaubte Aussagen zu den mittleren Unterschieden einer Phase im Vergleich zur Grundrate. Das zweite Modell (M2) lieferte durch die Interaktionsterme Aussagen zu den Trends innerhalb der einzelnen Phasen. Das dritte Modell (M3) ermöglichte Aussagen über mittlere Phasen-Unterschiede im Vergleich zur unmittelbaren Nachtstung für das Rechentraining. Aufschluss über einen unmittelbaren Leistungseffekt des Denktrainings auf die Rechenleistung würde demzufolge der Vergleich zwischen Grundrate und unmittelbarer Nachtstung für das Denktraining liefern (s. Modell 1). Die Überprüfung eines Lerneffektes erfolgte zum einen durch Modell 3, in dem die Phase der unmittelbaren Nachtstung mit der Phase der unmittelbaren Vortestung des Rechentrainings und der Grundratenerhebung kontrastiert wird. Zum anderen wurde der Trend innerhalb des Rechentrainings (s. Modell 2) betrachtet.

$$\text{Modell 1: } Y = \beta_0 + \beta_1 * d_1 + \beta_2 * d_2 + \beta_3 * d_3$$

$$\begin{aligned} \text{Modell 2: } Y = & \beta_0 + \beta_1 * T_0 + \beta_2 * d_1 + \\ & \beta_3 * T_1 * d_1 + \beta_4 * d_2 + \\ & \beta_5 * T_2 * d_2 + \beta_6 * d_3 + \\ & \beta_7 * T_3 * d_3 \end{aligned}$$

$$\text{Modell 3: } Y = \beta_0 + \beta_1 * d_1 + \beta_2 * d_2 + \beta_3 * d_4$$

Mittelwertvergleiche zwischen den Bedingungen erfolgten anhand einer einfaktoriellem Kovarianzanalyse (PROC GLM), wobei die jeweilige Vortestleistung kontrolliert wurde. Für alle Intergruppenvergleiche wurde die für Vortestunterschiede kontrollierte Effektstärke (d_{kor}) nach Klauer (2001) berechnet. Zur Berechnung der Effektstärken wurde auf einen einschlägigen Online-Rechner zurückgegriffen.

Ergebnisse

Deskriptive Statistiken

Getrennt nach Experimentalgruppe (Denktraining + Rechentraining) und nach Vergleichsgruppe (Alternativtraining + Rechentraining) zeigt Abbildung 1 den Verlauf der Leistungstests zum induktiven Denken und zum Rechnen in der jeweiligen Untersuchungsphase. Die zugehörigen Phasen-Mittelwerte werden in Tabelle 2 berichtet.

Tests zum induktiven Denken

Die Berechnungen für Modell 1, bei dem es innerhalb der beiden Bedingungen um die Bestimmung der mittleren Phasen-Unterschiede relativ zur Grundrate ging, ergaben Folgendes (s. Tabelle 3): In der Experimentalgruppe wurden, relativ zur Grundrate, in der Phase des Denktrainings ($b_1 = 0.82$, $p = .009$, $d = 0.91$) sowie in der Phase der unmittelbaren ($b_2 = 1.46$, $p < .001$, $d = 1.14$) und der verzögerten Nachtstung ($b_3 = 1.62$, $p < .001$, $d = 0.98$) höhere Leistungen in den Denktests erzielt. In der Vergleichsgruppe ergab sich ein gleichgerichteter Phasenunterschied zwischen Grundrate und verzögerter Nachtstung ($b_3 = 1.07$, $p = .030$, $d = 1.15$). Alle übrigen Phasenunterschiede waren nicht bedeutsam. Unter Kontrolle der Daten der Grundrate ergaben sich in der Nachtstungsphase (Phase 3) zwischen den beiden Bedingungen keine signifikanten Unterschiede in der über die vier Zeitpunkte gemittelten

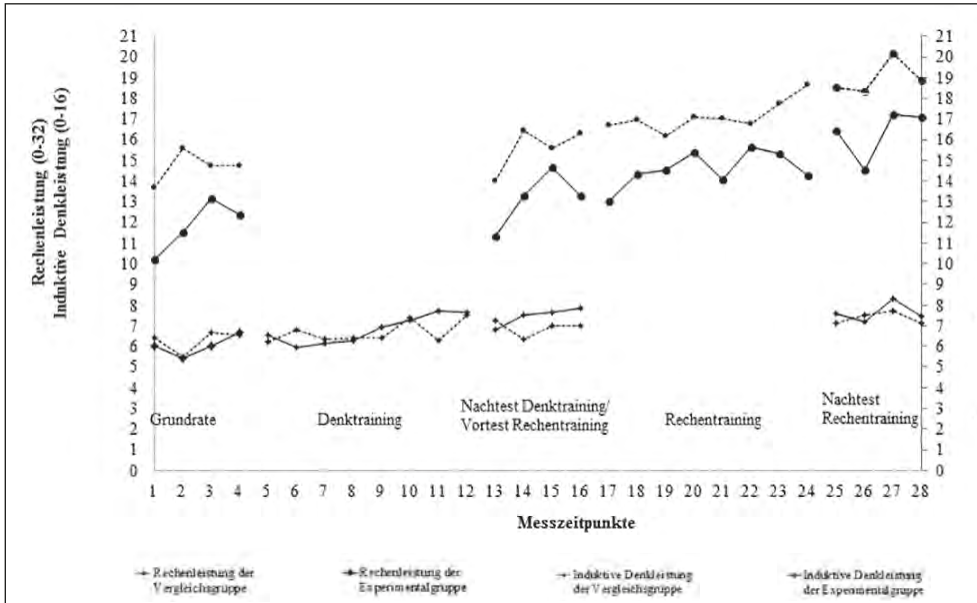


Abbildung 1. Verlauf der Denkleistung und der Rechenleistung für die Experimentalgruppe (Denktraining + Rechentraining) und für die Vergleichsgruppe (Alternativtraining + Rechentraining).

Tabelle 2: Deskriptive Statistiken für die Denkleistung und die Rechenleistung in den zugehörigen Erhebungsphasen

	Experimentalgruppe		Vergleichsgruppe	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Induktives Denken				
Grundratenerhebung	6.03	2.30	6.29	1.07
Denktraining	6.81	2.01	6.67	1.50
Unmittelbarer Nachttest für das Denktraining	7.46	2.58	6.92	1.85
Verzögerter Nachttest für das Denktraining	7.63	2.19	7.38	2.25
Rechenleistung				
Grundratenerhebung	11.82	6.89	14.68	7.04
Unmittelbarer Vortest für das Rechentraining	13.13	7.30	15.58	7.60
Rechentraining	14.56	8.24	17.13	6.92
Unmittelbarer Nachttest für das Rechentraining	16.32	9.14	18.97	7.62

Anmerkungen. Wegen eines zu hohen Anteils fehlender Werte wurde bei der Bestimmung der Rechenleistung 1 Schüler der Experimentalgruppe von der Auswertung ausgeschlossen. Ansonsten gilt: $n = 14$ in beiden Untersuchungsgruppen.

Denkleistung ($b = 0.79$, $SE = 0.58$, $p = .180$, $d_{korr} = 0.39$).

Die Testung von Modell 2 ergab, dass während der Phase des Denktrainings in der Experimentalgruppe ein positiver Trend über die wiederholten Denktests hinweg zu verzeichnen war ($b_3 = 0.24$, $p = .003$). In

der Vergleichsgruppe zeigte sich während der Vorschaltphase keine vergleichbare Entwicklung ($b_3 = 0.12$, $p = .229$). In keiner der beiden Bedingungen wurden signifikante Trends während der Grundrate oder in einer der übrigen Phasen festgestellt ($p > .05$).

Tabelle 3: Schätzung der festen Effekte bei der Analyse der induktiven Denkleistung getrennt für die Experimental- und die Kontrollgruppe

	Experimentalgruppe		Vergleichsgruppe	
	<i>b</i> (<i>SE</i>)	<i>p</i>	<i>b</i> (<i>SE</i>)	<i>p</i>
Modell 1				
b_0 : durchschnittliche Leistung: Grundrate	6.00 (0.60)	< .001	6.27 (0.38)	< .001
b_1*d_1 : durchschnittlicher Leistungsunterschied: Grundrate vs. Denktraining	0.82 (0.32)	.009	0.41 (0.32)	.201
b_2*d_2 : durchschnittlicher Leistungsunterschied: Grundrate vs. unmittelbare Nachtstung zum Denktraining	1.46 (0.37)	< .001	0.67 (0.41)	.098
b_3*d_3 : durchschnittlicher Leistungsunterschied: Grundrate vs. verzögerte Nachtstung zum Denktraining	1.62 (0.45)	< .001	1.07 (0.49)	.030
Modell 2				
b_0 : Intercept	5.59 (0.68)	< .001	5.98 (0.44)	< .001
b_1*T_0 : Trend in der Grundrate	0.27 (0.22)	.228	0.19 (0.22)	.381
$b_3*d_1*T_1$: Trend im Denktraining	0.24 (0.08)	.003	0.12 (0.10)	.229
$b_5*d_2*T_2$: Trend in der unmittelbaren Nachtstung zum Denktraining	0.35 (0.25)	.171	-0.03 (0.27)	.922
$b_7*d_3*T_3$: Trend in der verzögerten Nachtstung zum Denktraining	0.06 (0.27)	.808	0.01 (0.28)	.972

Anmerkungen. Aus Gründen der Übersichtlichkeit werden nur die festen Effekte berichtet.

Rechentests

Bezüglich der Frage, ob das Denktraining einen unmittelbaren Effekt auf das Ergebnis der Rechentests erzielte, ergab die Prüfung von Modell 1 Folgendes (s. Tabelle 4): In der experimentellen Gruppe ließ sich kein signifikanter Unterschied zwischen Grundrate und unmittelbarer Nachtstung des Denktrainings feststellen ($b_1 = 1.36$, $p = .072$, $d = 0.55$). Ebenso wenig zeigten sich in der Vergleichsgruppe überzufällige Unterschiede zwischen den Testungen vor (Grundrate) und nach (unmittelbarer Nachtstung) der Durchführung des alternativen Vorschalttrainings ($b_1 = 1.10$, $p = .242$, $d = 0.34$). Unter Kontrolle der Grundrate waren nach der Vorschaltphase (bzw. in Phase 3) keine bedeutsamen Unterschiede in der über die Zeitpunkte gemittelten Rechenleistung zwischen den beiden Bedingungen festzustellen ($b = 0.43$, $SE = 1.04$, $p = .679$, $d_{\text{kor}} = 0.08$).

Bezüglich eines mittelbaren Effektes lieferte die Testung von Modell 3 in der Experimentalgruppe folgende Befunde: Die Leistung bei den Rechentests fiel unmittelbar nach dem Rechentraining höher aus als unmittelbar davor ($b_1 = -3.41$, $p < .001$, $d = 1.09$). Ein gleichgerichteter Effekt ergab sich beim Vergleich der Grundrate mit der Nachtstungphase für das Rechentraining ($b_3 = -4.77$, $p < .001$, $d = 1.46$). Allerdings wurden in der Vergleichsgruppe analoge Effekte festgestellt: Für den Vergleich zwischen unmittelbarer Vor- und Nachtstung zum Rechentraining ($b_1 = -3.26$, $p < .001$, $d = 0.98$) und für den Vergleich zwischen Grundrate und unmittelbarer Nachtstung zum Rechentraining ($b_3 = -4.24$, $p < .001$, $d = 1.30$). Unter Kontrolle der Rechenleistung in Phase 3 (vor dem Rechentraining) unterschieden sich die beiden Bedingungen unmittelbar nach dem Rechentraining nicht signifikant in der über die Zeitpunkte gemittelten Rechenleistung ($b = -0.18$, $SE = 1.51$, $p = .907$, $d_{\text{kor}} = 0.01$).

Die Prüfung von Modell 2 zeigte, dass innerhalb der Phase des Rechentrainings weder für die Experimentalgruppe ($b_5 = 0.19, p = .133$) noch für die Vergleichsgruppe ($b_5 = 0.24, p = .131$) ein signifikanter Trend über die Messzeitpunkte hinweg festzustellen war.

Ergänzend wurden etwaige Transfereffekte des Denktrainings auf einzelne Aufgabentypen der Rechentests überprüft. Indizien für einen unmittelbaren Effekt des Denktrainings auf die Rechenleistung (Vergleich der Phasen 1 und 3) ergaben sich in der Experimentalgruppe bei den einfachen

($b_1 = 0.57, p < .001$) und bei den komplexen Ergänzungsaufgaben ($b_1 = 0.82, p = .030$). Additions- und Subtraktionsaufgaben ($b_1 = -1.45, p = .467$) sowie Gleichungen und Ungleichungen ($b_1 = -0.36, p = .311$) lieferten hingegen keine Hinweise auf einen unmittelbaren Transfereffekt. Für Kinder mit alternativem Vorschultraining fiel keiner der vier aufgabenspezifischen Vergleiche bedeutsam aus. Hinweise auf einen durch das Denktraining mittelbar bewirkten Transfereffekt (Vergleich der Phasen 3 und 5) lieferten in der Experimentalgruppe die komplexen Ergänzungsaufgaben

Tabelle 4: Schätzung der festen Effekte in der Analyse der Rechenleistung getrennt für die Experimentalgruppe und die Vergleichsgruppe

	Experimentalgruppe		Vergleichsgruppe	
	$b(SE)$	p	$b(SE)$	p
Modell 1				
b_0 : durchschnittliche Leistung: Grundrate	11.67 (2.22)	< .001	14.67 (1.94)	< .001
b_1*d_1 : durchschnittlicher Leistungsunterschied: Grundrate vs. unmittelbarer Nachtest zum Denktraining (Leistungseffekt)	1.36 (0.76)	.072	1.10 (0.85)	.242
b_2*d_2 : durchschnittlicher Leistungsunterschied: Grundrate vs. Rechentraining	2.86 (0.72)	< .001	2.50 (0.71)	< .001
b_3*d_3 : durchschnittlicher Leistungsunterschied: Grundrate vs. unmittelbarer Nachtest zum Rechentraining	4.64 (1.10)	< .001	4.31 (1.00)	< .001
Modell 2				
b_0 : Intercept	10.62 (1.87)	< .001	14.40 (1.93)	< .001
b_1*T_0 : Trend in der Grundrate	0.80 (0.39)	.039	0.20 (0.51)	.698
$b_3*d_1*T_1$: Trend unmittelbarer Vortest zum Rechentraining	0.73 (0.36)	.039	0.47 (0.32)	.142
$b_5*d_2*T_2$: Trend Rechentraining	0.19 (0.12)	.133	0.24 (0.16)	.131
$b_7*d_3*T_3$: Trend unmittelbarer Nachtest zum Rechentraining	0.67 (0.42)	.111	0.30 (0.42)	.470
Modell 3				
b_0 : durchschnittliche Leistung: unmittelbarer Nachtest zum Rechentraining	16.46 (2.28)	< .001	18.93 (1.98)	< .001
b_1*d_1 : durchschnittlicher Leistungsunterschied: unmittelbarer Vor- und Nachtest zum Rechentraining (Lerneffekt)	-3.41 (0.89)	< .001	-3.26 (0.94)	< .001
b_2*d_2 : durchschnittlicher Leistungsunterschied: unmittelbarer Nachtest zum Rechentraining	-1.91 (0.86)	.025	-1.77 (0.79)	.025
b_3*d_3 : durchschnittlicher Leistungsunterschied: Grundrate vs. unmittelbarer Nachtest zum Rechentraining (Lerneffekt)	-4.77 (1.00)	< .001	-4.24 (0.97)	< .001

Anmerkungen. Aus Gründen der Übersichtlichkeit werden nur die festen Effekte berichtet.

($b_1 = -1.08$, $p = .008$) sowie die Gleichungen und Ungleichungen ($b_1 = -0.97$, $p = .023$). Da sich aber auch die Vergleichsgruppe in den einfachen ($b_1 = -0.75$, $p = 0.027$) und in den komplexen Ergänzungsaufgaben ($b_1 = -1.54$, $p < .001$) verbesserte, ließen sich auch aus dieser Analyse keine dezidierten Belege für spezifische Transfereffekte des Klauer-Trainings ableiten.

Diskussion

In dieser mit lernschwachen Kindern durchgeführten Trainingsstudie ergaben sich bezüglich der eingangs formulierten Hypothesen folgende Ergebnisse:

(a) Wie erwartet führte das Denktraining zu besseren Leistungen bei Aufgaben, die einem Test zur fluiden Intelligenz entstammten. Die Verbesserung zeigte sich anhand von Leistungszuwächsen, die während der Trainingsphase zu verzeichnen waren (H1), ebenso wie anhand von Unterschieden im Durchschnittsniveau zwischen Grundrate einerseits sowie unmittelbarer und verzögerter Nachtstung andererseits (H2). Kinder mit sozial-emotionalem Vortraining zeigten weder während des Trainings noch unmittelbar danach einen Zuwachs in der Denkleistung. Allerdings übertrafen sie in der Phase der verzögerten Nachtstung ihre während der Grundrate erzielte Leistung. Für diesen unerwarteten Effekt könnten generelle Gewinne in der Denkentwicklung der Kinder verantwortlich gewesen sein; oder auch Retest-Effekte bzw. Effekte wachsender Testerfahrung, wogegen allerdings spricht, dass in keiner der drei reinen Testungsphasen (1, 3 und 5) ein signifikanter Trend bei der Lösung der Denkaufgaben festzustellen gewesen ist. Wurden die Analysen innerhalb der Individuen um den Vergleich zwischen den Bedingungen ergänzt, so ergab sich für die Überlegenheit der experimentellen Gruppe im unmittelbaren Nachtstung ein schwacher

Effekt in Höhe von $d = 0.39$ (kontrolliert für die durchschnittliche Leistung in der Grundrate). Bei der Einordnung dieses Effekts ist zu beachten, dass die Vergleichsgruppe in Phase 2 gleichfalls eine Trainingsmaßnahme durchlief, sodass trainingsunspezifische Effekte, wie z.B. Zuwendungseffekte, als Erklärung des Gruppenunterschieds wohl auszuschließen sind.

(b) Keine Unterstützung fand die Idee, dass das Denktraining unmittelbar bzw. aus sich selbst heraus Transfereffekte auf das Lösen von Rechenaufgaben erzielen kann (H3). Zwar verfehlte der Vergleich der Rechenleistungen in den Phasen 1 und 3 in der experimentellen Gruppe nur knapp die statistische Signifikanz ($p = .072$, zweiseitig). Zum einen traten in dieser Gruppe aber signifikante Trends innerhalb der beiden Testungsphasen auf (s. Tabelle 6), was auf die Möglichkeit von Retest-Effekten hindeuten mag. Zum anderen wies auch die Vergleichsgruppe von Phase 1 zu Phase 3 leichte Verbesserungen in der Matheleistung auf. Im Ergebnis ließ sich zwischen den Bedingungen keinerlei Unterschied ($d = 0.08$, korrigiert für Unterschiede in der Grundratenleistung) in der Wirksamkeit der beiden Vorschalttrainings bei der Lösung curricular valider Rechenaufgaben feststellen.

(c) Anders als vorhergesagt (H4 und H5) konnte das Denktraining im Vergleich zum alternativen Vorschalttraining die Wirksamkeit des Rechentrainings nicht erhöhen. Zwischen den Phasen 3 und 5 fielen die durchschnittlichen Niveauunterschiede in beiden Bedingungen signifikant aus. Umso erstaunlicher war, dass während des Trainings in keiner der beiden Gruppen ein signifikanter Zuwachs in der Rechenleistung festzustellen war. Denkbar ist, dass die Testung am Ende einer Trainingssitzung Ermüdungseffekten unterlag, sodass die Wirkung des Trainings erst im Anschluss an die Trainingsphase zu sehen war. In zukünftigen Studien wäre zu prüfen, ob sich Testungen zu Beginn statt am Ende einer Trainingssitzung besser dafür eignen, um potenzielle

Lernzuwächse deutlich zu machen. Jedenfalls weist die Verbesserung der Rechenleistung von Phase 3 zu Phase 5 darauf hin, dass das Rechenttraining durchaus, wenn gleich bedingungsunabhängig wirksam gewesen ist. Die interne Validität dieses Ergebnisses wird dadurch eingeschränkt, dass aufgrund der Fragestellung der Studie keine Kontrollgruppe ohne Rechenttraining in den Versuchsplan integriert wurde.

Unsere Studie unterscheidet sich von früheren Untersuchungen, in denen Transfereffekte von Trainings des induktiven Denkens besser nachzuweisen waren, reiht sich aber in andere Studien ein, die diesbezüglich gleichfalls weniger erfolgreich waren (siehe Einführung). Zu den Stärken und Besonderheiten der Studie zählt, dass in ein und demselben Kontext unmittelbare und mittelbare Transfereffekte gemeinsam analysiert wurden; dass nach dem Vorbild von Einzelfallstudien Verlaufsinformationen während der Trainings aufgezeichnet wurden; dass die Datenanalyse mit einem der Einzelfallforschung entlehnten Mehrebenen-Ansatz erfolgte; und dass unspezifische (Transfer-)Effekte des Denktrainings durch ein alternatives Vorschalttraining kontrolliert wurden. Zudem gehört die Studie zu den wenigen Untersuchungen (Klauer, 1996), in denen Transfereffekte des Klauerschen Denktrainings nicht anhand einer eng umrissenen Lehrstunde (Sonntag, 2004), sondern im Rahmen eines intensiven Trainings – hier: des rechnerischen Denkens – überprüft wurden. Enttäuschend mag wirken, dass bei diesem Vorgehen nicht nur der unmittelbare Transfereffekt ausblieb; keine Unterstützung fand auch die Vorstellung, dass die Verbesserung einer kognitiven Lernvoraussetzung, wie hier des induktiven Denkens, die Wirksamkeit des Trainings einer schulischen Fähigkeit erhöht.

Limitationen

Kritisch betrachtet bieten sich mehrere Erklärungen an, warum in dieser Studie weder unmittelbare noch mittelbare Transfer-effekte des Denktrainings nachzuweisen waren:

Erstens lässt sich einwenden, dass der Effekt des acht Sitzungen umfassenden Denktrainings zu schwach war, um Transfereffekte sichtbar werden zu lassen. Aus dem während der Trainingsphase (Phase 2) ermittelten Trend lässt sich sogar hochrechnen, wie viele zusätzliche Sitzungen notwendig wären, um substanziellere Förder-effekte erzielen zu können. Skeptisch anzumerken ist, dass weder der unmittelbare Transfereffekt, schon gar nicht aber der mittelbare, irgendwelche Tendenzen erkennen ließ, die bei einer Verlängerung des Denktrainings mutmaßlich signifikant werden müssten.

Zweitens verbesserte sich die Vergleichsgruppe in den Tests zum induktiven Denken von der Grundrate zur verzögerten Nachtestung. Dies mag den Nachweis eines Lernvorteils der Experimentalgruppe während und in Folge des Rechentrainings erschwert haben. Um diesem (unerwarteten) Effekt auf den Grund zu gehen, wäre es aufschlussreich, dem Design eine Gruppe hinzuzufügen, die in der Vorschaltphase gar kein Treatment erhält. So ließe sich ermitteln, ob der betreffende Effekt als Ergebnis von Reifungsprozessen, von Zuwendungseffekten oder von anderen trainingsunspezifischen Einflüssen auftritt.

Drittens lässt sich der Einspruch erheben, dass die hier verwendeten Rechenaufgaben nicht hinreichend sensitiv für Transfereffekte des Denktrainings waren. Denkbar wäre, dass die trainierten und getesteten Rechenleistungen zu elementar waren, als dass die Schülerinnen und Schüler bei ihrer Lösung von der Strategie des induktiven Denkens profitieren konnten. Gestützt wird dieses Argument durch Nachanalysen zum unmittelbaren Transfereffekt: Ergänzungsaufgaben waren für Effekte des Denktrainings

nings empfänglicher als Additions- oder Subtraktionsaufgaben. Andererseits ist zu bedenken, dass die hier verwendeten, an den DEMAT angelehnten Verlaufstests, trotz der homogenen Stichprobe, mit einem komplexeren Test mathematischer Fähigkeiten (HRT) substantiell korreliert waren. Außerdem waren die Grundraten-Ergebnisse der Denktests mit denen der Rechentests so hoch korreliert ($r = .67$), dass die Bedeutung des induktiven Denkens für das Lösen der hier verwendeten Rechenaufgaben nicht einfach von der Hand zu weisen ist. Dennoch wäre es lohnenswert genauer zu analysieren, welche Rechenaufgaben von der Strategie des induktiven Denkens besonders profitieren und für welche Aufgaben diese Strategie weniger relevant ist.

Viertens wurde in der vorliegenden Studie wenig getan, um den Transfer vom Denk- zum Rechentraining gezielt anzubahnen. Lediglich die Einbettung der Denk- und der Rechenispiele in einen gemeinsamen Makrokontext („Elfe und Mathis“) hätte potentiell transferfördernd wirken können. Daher lässt sich nicht ausschließen, dass sich die teilnehmenden Schülerinnen und Schüler erst gar nicht bemühten, die Strategie des induktiven Denkens auf die Lösung der Mathematikaufgaben anzuwenden. Dieses Argument läuft auf die Vorstellung hinaus, dass ein solcher Transfer zumindest bei lernschwachen Kindern nicht oder nur in Ausnahmefällen selbst initiiert auftritt. Explizite Hinweise und instruktionale Unterstützungen, Demonstrationen am Beispiel oder auch die Bildung von Vorsätzen (oder „Wenn-Dann“-Plänen) zur Anwendung der erlernten Strategie könnten sich als nützlich erweisen, um den Transfer von Denkstrategien auf die Lösung mathematischer Aufgaben wahrscheinlicher zu machen.

Fünftens lässt sich kritisieren, dass in dieser Studie die Größe der Stichprobe und damit die Testpower zu gering waren, um die erwarteten Effekte statistisch absichern zu können. Bezüglich der Analysen, die innerhalb der Individuen durchgeführt wur-

den, greift dieses Argument kaum. Tatsächlich überschreiten Stichproben in Einzelfallstudien (z.B. mit multiplen Grundraten oder mit alternierenden Treatments) nur selten den einstelligen Bereich. Schwerer wiegt das Argument, wenn es um Vergleiche zwischen den Bedingungen geht. Relativ zur Gruppe mit alternativem Training ließ sich z.B. der Effekt der „Denkspiele“ auf die Nachtest-Leistung zum induktiven Denken nicht gegen den Zufall absichern. Im Hinblick auf die Transfereffekte dürfte die bloße Erhöhung der Stichprobengröße aber kein probates Mittel sein, um statistische Signifikanz zu erreichen. Tendenzen in die vorhergesagte Richtung zeichneten sich beim unmittelbaren Transfer allenfalls spurenweise und beim mittelbaren Transfer eigentlich gar nicht ab. Immerhin: Rein methodisch gesehen, würden es größere Stichproben ermöglichen, die im Einzelfalldesign gewonnenen Informationen über Phasenunterschiede und Verlaufstrends auf den Einfluss moderierender Variablen, wie z.B. der Art des Vorschalttrainings, hin zu analysieren (so wie dies z.B. in Metaanalysen über Einzelfallstudien gemacht wird).

Zusammenfassend ergaben sich in dieser Studie keine belastbaren Belege, dass sich ein Training des induktiven Denkens unmittelbar oder mittelbar bzw. vermittelt über eine weitere Trainingsmaßnahme vorteilhaft auf den Nachweis und das Erlernen rechnerischer Fähigkeiten auswirkt. Obgleich sich die hier untersuchten lernschwachen Kinder im intraindividuellen Vergleich sowohl im Denk- als auch im Rechentraining verbesserten, blieben die vorhergesagten Transfereffekte aus. Weitere Studien sollten analysieren, welche Arten von Rechenaufgaben vom induktiven Denken besonders profitieren und welche instruktionalen Maßnahmen erforderlich sind, um sicherzustellen, dass die Strategie des induktiven Denkens auf die Lösung von Rechenaufgaben auch tatsächlich angewendet wird.

Literatur

- Aebi, M., Perriard, R., Stiffler Scherer, B. & Wettach, R. (2012). *Kinder mit oppositionellem und aggressivem Verhalten: Das Baghira Training*. Göttingen: Hogrefe.
- Barkl, S., Porter, A. & Ginns, P. (2012). Cognitive training for children: Effects on inductive reasoning, deductive reasoning, and the mathematics achievement in an Australian school setting. *Psychology in the Schools*, 49(9), 828–842.
- Christou, C. & Papageorgiou, E. (2007). A framework of mathematics inductive reasoning. *Learning and Instruction*, 17, 55–66.
- Haffner, J., Baro, K., Marx, H. & Schneider, W. (2005). *Heidelberger Rechentest: Erfassung mathematischer Basiskompetenzen im Grundschulalter*. Göttingen: Hogrefe.
- Hasselhorn, M. (1998). Metakognition. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 348–351). Weinheim: Psychologie Verlags Union.
- Helmke, A. (2006): Unterrichtsqualität. In Rost, D. H. (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (3. Aufl., S. 812–820). Weinheim: Beltz.
- Kazdin, A. E. (2011). *Single-case research designs. Methods for clinical and applied settings* (2. Aufl.). New York: Oxford University Press.
- Klauer, K. J. (1989). *Denktraining für Kinder I: Ein Programm zur intellektuellen Förderung*. Göttingen: Hogrefe.
- Klauer, K. J. (1996). Denktraining oder Lesetraining? Über die Auswirkungen eines Trainings zum induktiven Denken sowie eines Lesetrainings auf das Leseverständnis und induktives Denken. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, (1), 67–89.
- Klauer, K. J. (2001). Training des induktiven Denkens. In K. J. Klauer (Hrsg.), *Handbuch Kognitives Training* (S. 165–209). Göttingen: Hogrefe.
- Klauer, K. J. (2014a). Förderung des induktiven Denkens und Lernens. In G. Lauth, M. Grünke & J. C. Brunstein (Hrsg.), *Intervention bei Lernstörungen. Förderung, Training und Therapie in der Praxis* (2., überarbeitete und erweiterte Auflage, S. 331–340). Göttingen: Hogrefe.
- Klauer, K. J. (2014b). Training des induktiven Denkens – Fortschreibung der Metaanalyse von 2008. *Zeitschrift für Pädagogische Psychologie*, 28(1-2), 5–19.
- Klauer, K. J. & Phye, G. D. (2008). Inductive Reasoning: A training approach. *Review of Educational Research*, 78(1), 85–123.
- Klauer, K. J., Willmes, K. & Phye, G. D. (2002). Inducing inductive reasoning: Does it transfer to fluid intelligence? *Contemporary Educational Psychology*, 27, 1–25.
- Krajewski, K., Küspert, P. & Schneider, W. (2002). DEMAT 1+: Deutscher Mathematiktest. Göttingen: Hogrefe.
- Krajewski, K., Liehm, S. & Schneider, W. (2004). *DEMAT 2+: Ein Mathematiktest für zweite Klassen*. Göttingen: Hogrefe.
- Lauth, G., Brunstein, J. C. & Grünke, M. (2014). Lernstörungen im Überblick: Arten, Klassifikation, Verbreitung und Erklärungsperspektiven. In G. Lauth, M. Grünke & J. C. Brunstein (Hrsg.), *Intervention bei Lernstörungen. Förderung, Training und Therapie in der Praxis* (2., überarbeitete und erweiterte Auflage, S. 17–31). Göttingen: Hogrefe.
- Lenhard, A. & Lenhard, W. (2011). Computerbasierte Intelligenzförderung mit den „Denkspielen mit Elfe und Mathis“ – Vorstellung und Evaluation eines Computerprogramms für Vor- und Grundschüler. *Empirische Sonderpädagogik*, (2), 105–120.
- Lenhard, A., Lenhard, W. & Klauer, K. J. (2012). *Denkspiele mit Elfe und Mathis: Förderung des logischen Denkvermögens für das Vor- und Grundschulalter*. Göttingen: Hogrefe.
- Lenhard, W. & Lenhard, A. (2009). *Rechen-spiele mit Elfe und Mathis I: Ein Mathematiktraining für Kinder der ersten bis dritten Jahrgangsstufe*. Göttingen: Hogrefe.
- Lobo, M. A., Moeyaert, M., Baraldi Cunha, A. & Babik, I. (2017). Single-Case Design,

- Analysis, and Quality Assessment for Intervention Research. *Journal of Neurologic Physical Therapy*, 41(3), 187–197.
- Marx, E. (2005). Bewirkt ein kognitives Training das, was es bewirken soll? Programme zur Denk- und Aufmerksamkeitsförderung im Vergleich. *Zeitschrift für Pädagogische Psychologie*, 19(4), 237–247.
- Marx, E. (2006). Profitiert das kindliche Sprachsystem von anderen kognitiven Entwicklungsbereichen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 38(3), 139–145.
- Marx, E. & Keller, K. (2010). Effekte eines induktiven Denktrainings auf die Denk- und Sprachentwicklung bei Vorschulkindern und Erstklässlern in benachteiligten Stadtteilen. *Zeitschrift für Pädagogische Psychologie*, 24(2), 139–146.
- Moeyaert, M., Ferron, J. M., Beretvas, N. & van den Noortgate, W. (2014). From a single-level analysis to a multilevel analysis of single-case experimental designs. *Journal of School Psychology*, 52, 191–211.
- Moeyaert, M., Ugille, M., Ferron, J. M., Beretvas, N. & van den Noortgate, W. (2014). The influence of the design matrix on treatment effect estimates in the quantitative analyses of single-subject experimental design research. *Behavior Modification*, 38(5), 665–704.
- Peng, J. & Chen, L. (2018). Handling missing data in single-case studies. *Journal of Modern Applied Statistical Methods*, 7 (1), 1–33.
- Petermann, F., Koglin, U., Natzke, H. & Mares, N. von. (2013). *Verhaltenstraining in der Grundschule: Ein Programm zur Förderung emotionaler und sozialer Kompetenzen*. Göttingen: Hogrefe.
- Petermann, F. & Petermann, U. (2010). *Hamburg-Wechsler-Intelligenztest für Kinder – IV (HAWIK-IV)*. Göttingen: Hogrefe.
- Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F. & Spinath, F. M. (2015). Intelligence and school grades. A meta-analysis. *Intelligence*, 53, 118–137. <https://doi.org/10.1016/j.intell.2015.09.002>
- Soares, D. L., Lemos, G. C., Primi, R. & Almeida, L. S. (2015). The relationship between intelligence and academic achievement throughout middle school. The role of students' prior academic performance. *Learning and Individual Differences*, 41, 73–78. <https://doi.org/10.1016/j.lindif.2015.02.005>
- Sonntag, W. (2004). Der Einfluss des Klauerschen Denktrainings auf mathematisches Denken und Lernen von lernbehinderten Sonderschülern. *Zeitschrift für Pädagogische Psychologie*, 18(2), 101–111.
- Sonntag, W. (2010). Fördert induktives Denken die Gedächtnisstrategie des Kategorisierens bei lernbehinderten Sonderschülern? *Empirische Sonderpädagogik*, (1), 5–21.
- Tomic, W. (1995). Training in inductive reasoning and problem solving. *Contemporary Educational Psychology*, 20, 483–490.
- Tomic, W. & Klauer, K. J. (1996). On the effects of training inductive reasoning: How far does it transfer and how long do the effects persist? *European Journal of Psychology of Education*, (3), 283–299.
- Weiß, R. H. (2006). *CFT 20-R: Grundintelligenztest Skala 2 – Revision mit Wortschatztest und Zahlenfolgentest*. Göttingen: Hogrefe.
- Weiß, R. H. & Osterland, J. (2012). *CFT 1-R: Grundintelligenztest Skala 1*. Göttingen: Hogrefe.

Dr. Eva Keller

*Justus-Liebig-Universität Gießen
Abteilung für Pädagogische Psychologie
Otto-Behaghel-Str. 10 F
35394 Gießen*

Erstmalig eingereicht: 15.11.2019

Überarbeitung eingereicht: 24.03.2020

Angenommen: 26.03.2020