Empirische Pädagogik
2015, 29. Jahrgang, Heft 2, S. 161-188
© Verlag Empirische Pädagogik

Originalarbeiten

Ein Quasi-Experiment zur Gruppenrallye im Mathematikunterricht: Hält die Methode, was sie verspricht?

Christian Wandeler, Alois Niggli, Caroline Villiger,
Marcel Aebischer & Philippe Leopold

Die vorliegende quasi-experimentelle Interventionsstudie untersuchte Effekte einer Gruppenrallye im Fach Mathematik. Insgesamt nahmen 497 Fünftklässler aus dem deutschsprachigen Teil des Schweizer Kantons Freiburg teil. Je 15 Schulklassen wurden nach dem Zufall der Interventions- bzw. der Kontrollgruppe zugewiesen. Mehrebenenanalysen erbrachten die folgenden Hauptergebnisse: Unter der Kontrolle des Geschlechts, der sozialen Herkunft, der Intelligenz sowie des Leseverständnisses auf der Individualebene erreichten die Schüler, die in einer Rallye-Klasse unterrichtet worden waren, etwas höhere Leistungen als die Kontrollgruppe. Von den Kontrollvariablen hatte das Geschlecht den größten Einfluss. Mädchen erreichten geringere Leistungen. Interaktionseffekte zwischen den Kontrollvariablen und der Intervention kamen keine zustande. Insofern schienen alle Schüler von der Intervention in gleichem Masse profitiert zu haben. Die Effektstärken waren eher moderat, so dass sich weitere Untersuchungen rechtfertigen würden. Diese könnten dazu beitragen, Empfehlungen in der didaktischen Ratgeberliteratur zur Gruppenrallye für das Fach Mathematik der Grundstufe differenzierter zu beurteilen.

Schlagwörter: Gruppenrallye - Kooperatives Lernen - Mathematikleistung

1 Einleitung

Untersuchungen zum kooperativen Lernen gelten als eine der bemerkenswertesten Erfolgsgeschichten der Unterrichtsforschung (Borsch, 2010; Johnson & Johnson, 1999; Slavin, Hurley & Chamberlain, 2003). Hinsichtlich kognitiver Zielvariablen scheinen kooperative Lernformen individualisiertem und kompetitivem Lernen insgesamt überlegen zu sein (Hattie, 2009; Johnson & Johnson, 1999; Slavin, 1995; Springer, Stanne, & Donovan, 1999). Als theoretische Erklärungen kommen kognitiv- bzw. sozio-konstruktivistische Ansätze in Frage (Borsch, 2010; Lipowsky, 2015), die annehmen, dass die kognitive Entwicklung durch Aushandlungs- und Entwicklungsprozesse gefördert wird. Zu den Effekten auf der motivational-affektiven Ebene (z. B. Einstellungen zum Lernen, Selbstwertgefühl, Motivation) liegen deutlich weniger Studien vor als für die Leistungsentwicklung (Lipowsky, 2015). Zugängliche Metaanalysen indizieren jedoch auch in diesen Fällen ein positives Bild (Ginsburg-Block, Rohrbeck & Franzutto, 2006; Lou et al., 1996; Springer, Stanne, & Donovan, 1999). Vor allem das Erleben sozialer Eingebundenheit scheint die Lernmotivation der Gruppenmitglieder zu fördern.

Aus der Tradition behavioraler Lerntheorien können ferner auch gruppenbezogene Belohnungen, die auf der Basis individueller Verantwortlichkeiten erteilt werden, einen entscheidenden Wirkmechanismus darstellen (Lipowsky, 2015, S. 87).

Die Wirksamkeit kooperativer Methoden setzt voraus, dass bestimmte Basiselemente realisiert werden (vgl. Hasselhorn & Gold, 2013, S. 310f.). Darin unterscheiden sie sich von gängigen Gruppenarbeiten, bei denen Lernende bei der Bewältigung einer Aufgabe nicht zwingend aufeinander angewiesen sind. Kennzeichnend für Kooperatives Lernen ist dagegen das Vorhandensein wechselseitiger Abhängigkeit (positive Interdependenz) der Lernenden, ferner individueller Verantwortlichkeit, förderlicher Interaktionen, kooperativer Arbeitstechniken und gemeinsamer Reflexion über abgelaufene Lernprozesse. Oftmals sind diese Voraussetzungen jedoch nicht gegeben (Huber, 1997; Kostsopoulos, 2010; Veenman, Kenter & Post, 2000), so dass vorteilhafte Ergebnisse kooperativen Lernens ausbleiben. Auch Drittvariablen wie das Geschlecht, Zielsetzungen und Freiheitsgrade der Aufgaben scheinen eine bedeutende Rolle zu spielen (Rohrbeck, Ginsburg-Block, Fantuzzo & Miller, 2003). Eine Metaanalyse von Kyndt et al. (2013), die Untersuchungen aus jüngerer Zeit analysiert haben, hat zudem ergeben, dass kooperatives Lernen kultur- und domänenspezifisch variiert. Auch das Alter der Lernenden kann die Ergebnisse beeinflussen. Auf der Primar- und Tertiärstufe wurden höhere Effektstärken gefunden als auf der Sekundarstufe.

2 Die Gruppenrallye (STAD) als prototypische Form kooperativen Lernens

Die vorliegende Studie untersucht die Wirksamkeit des Gruppenrallyes (STAD: Student teams achievement divisions) nach Slavin (1994) im Mathematikunterricht der Grundschule. Mit der Gruppenrallye wird dem Anspruch nachgekommen, eine Methode gemäß den Basiselementen kooperativen Lernens konzipiert zu haben. Es handelt sich dabei um eine Kombination aus einführendem Unterricht im Klassenverband, anschließender Gruppenarbeit, Leistungsüberprüfung und Gruppenbelohnung (vgl. Lipowsky, 2015, S. 86). Jede Gruppe bekommt die gleichen Aufgaben und Lernmaterialien. Im Test arbeitet jeder Schüler allein. Die erzielten Leistungen werden als Verbesserungswerte mit einer zuvor erfassten Baseline verglichen und pro Gruppe aufsummiert. Die Gruppe mit den meisten Punkten gewinnt. Eine Gruppe – so die Erwartung – ist nur dann erfolgreich, wenn sich die Lernenden in ihrem Verstehensprozess gegenseitig unterstützen. Von anderen kooperativen Lernformen unterscheidet sich die Gruppenrallye durch ihren Wettbewerbscharakter. Große Bedeutung wird im Weiteren der individuellen Leistungsüberprüfung eingeräumt. Im Vergleich zur Bewertung eines von der Gruppe gemeinsam er-

stellten Produkts konnte Slavin (1996) deutliche Vorzüge dieses Verfahrens nachweisen.

Empirisch ist die Wirksamkeit der Gruppenrallye vor allem für kognitive Lernziele gut belegt. Slavin (1995) berichtet in seiner Metaanalyse einen mittleren Effekt von d = 0.32 zugunsten der Gruppenrallye; Johnson, Johnson und Stanne (2000) Effekte von d = 0.51 gegenüber kompetitiven und d = 0.29 gegenüber individualistisch strukturierten Kontrollbedingungen. In praxisorientierten Publikationen zum kooperativen Lernen kann die Gruppenrallye den Lehrkräften deshalb meist vorbehaltlos empfohlen werden (Borsch 2010; Huber, 2008; Traub, 2004). Dafür sprechen neben ihrer Wirksamkeit auch pädagogische Gründe: Gruppen sind nur dann erfolgreich, wenn sich ihre Mitglieder die Konzepte gegenseitig erklären. Begleitendes Training kann diese Gruppenprozesse anregen (Wellenreuther, 2014, S. 459). Aufgrund der individuellen Bezugsnorm beim Berechnungsverfahren erhalten schwächere Schüler ferner durch unterstützendes Dazulernen dieselben Chancen, einen positiven Beitrag zum Gruppenergebnis zu leisten wie gute Schüler.

Gegenüber der durchwegs positiven Rezeption in der didaktischen Literatur lassen sich inzwischen dennoch Gründe anführen, die dafür sprechen, die Methode erneut empirisch zu prüfen. Im deutschsprachigen Raum existieren beispielsweise keine Studien über ihre Wirksamkeit. Ferner variieren die Effektstärken zwischen den Metaanalysen relativ stark. Dies spricht für differentielle Analysen, beispielsweise hinsichtlich unterschiedlicher Fächer oder Altersgruppen. Im vorliegenden Fall ist die Metaanalyse von Slavin und Lake (2008) über Grundschul-Lernprogramme im Fach Mathematik von Bedeutung. Zwar hat er die Gruppenrallye als eine der wenigen Methoden hervorheben können, die deutliche Effekte („strong evidence of effectiveness“) erbracht haben. Bei näherer Betrachtung stellen sich jedoch nach wie vor Anschlussfragen, die es rechtfertigen, den Erkenntnisstand zu erweitern. Tabelle 1 gibt einen Überblick über die vier Studien, die seiner Metaanalyse zur Wirksamkeit der Methode im Mathematikunterricht der Grundstufe zugrunde liegen. Eine zusätzliche Studie von Tarim & Akdeniz (2008) konnte von Slavin noch nicht berücksichtigt werden. Sie genügt jedoch den Auswahlkriterien, die seiner Metaanalyse zugrunde liegen.

Tabelle 1: Überblick über geeignete Studien zur Wirksamkeit des Gruppenrallyes im Mathematikunterricht der Grundschule

Studie	Dauer	N	Stufe	Kultureller Kontext	Effekt-stärke	Bemerkungen
Glassmann (1989)	6 Monate	441	3-5	USA	0.01	Mädchen profitierten mehr; Leistungsgruppen: nicht sig.
Mevarech* (1985)	15 Wochen	34	5	Israel	0.19	Math. Berechnungen: sig. Math. Verständnis: nicht sig. Vergleichsstichprobe mit Mastery Learning erreichte höhere Werte
Mevarech* (1991)	12 Wochen	32	3	Israel	0.60	Low achievers profitierten mehr
Suyanto (1998)	4 Monate	664	3-5	Indonesien	0.40	4.-Klässler: nicht sig. Leistungsgruppen: nicht sig.
Tarim & Akdeniz (2008)	14 Wochen	175	4	Türkei	0.40	TAI-Methode** war wirksamer als STAD; Einstellungen gegenüber Mathematik: nicht sig.

Anmerkungen: *Mevarech (1985; 1991) verglich jeweils 4 Schulklassen (3 Treatmentbedingungen und 1 Kontrollklasse). Die Angaben in Tabelle 1 beziehen sich ausschliesslich auf die STAD-Probanden. Neben der Kontrollgruppe wurde das STAD-Konzept in einer Schulklasse mit Mastery-Learning Elementen kombiniert. Eine weitere Schulklasse erfuhr ausschliesslich ein Mastery-Learning-Treatment.
** TAI: Team Assisted Individualization (Slavin, 1984); Das N dieser Studie bezieht sich nur auf die involvierten STAD-Probanden.

Die Effektstärken der Studien in Tabelle 1 schwanken zwischen 0.01 (Glassmann, 1989) und 0.60 (Mevarech, 1991). In den beiden Arbeiten aus Israel fällt die geringe Probandenzahl auf. Die israelische Studie mit der höchsten Effektstärke stützt sich auf eine einzelne Rallye-Klasse mit lediglich 32 Kindern. Die Studien aus Indonesien und aus der Türkei scheinen im Vergleich zur Studie aus den USA den Trend zu bestätigen, dass sich kooperatives Lernen in nicht-westlichen, eher kollektivistischen Kulturen als wirksamer erwiesen hat (Kyndt et al., 2013). Verwendet wurden in allen Studien quasi-experimentelle Designs mit Prä-Post-Messungen, die varianzanalytisch ausgewertet wurden. Teilweise wurden zwei Gruppierungsvariablen miteinander kombiniert (z. B. Niveaugruppen und Treatmentbedingungen). Neben der Leistung wurde von Tarim und Akdeniz (2008) auch die zusätzliche abhängige Variable „Einstellungen gegenüber dem Fach Mathematik" untersucht. Aktuellere mehrebenenanalytische Verfahren, die es erlauben, Effekte von Faktoren unterschiedlicher Aggregatniveaus in einer Analyse simultan zu berücksichtigen, wurden bis anhin nicht eingesetzt. Der Einfluss moderierender Faktoren konnte somit lediglich ansatzweise geprüft werden.

Die Beurteilung der Methode in der Metaanalyse von Slavin und Lake (2008) stützt sich infolgedessen auf zwei Studien mit sehr geringen Stichproben und auf zwei Dissertationen mit deutlich auseinanderliegenden Effektstärken, die Ende der 80er und 90er Jahre des letzten Jahrhunderts abgeschlossen worden sind. Eine jüngere Studie aus der Türkei bestätigte die Resultate aus Indonesien. Der relativ schmale Forschungsstand zur Wirksamkeit der Methode im Fach Mathematik rechtfertigt somit weitere Untersuchungen. Diese sollten vorzugsweise in westlichen Kulturen durchgeführt werden und den Einfluss zentraler moderierender Faktoren berücksichtigen.

3 Eckpunkte bei der Gestaltung einer Gruppenrallye

Eine Rallye wird einerseits durch ihre organisatorische Anlage (Klassenlernen – Kooperatives Lernen), den abschließenden Test und andererseits durch eine Anleitung zur Kooperation gesteuert, an der sich die Gruppen orientieren können (vgl. ausführlich Huber, 1985; Slavin 1988).

Organisatorische Anlage: Grundsätzlich eignet sich die Gruppenrallye für die Weiterführung und Festigung von Inhalten, die im Unterricht von der Lehrperson mit der Klasse eingeführt und erläutert worden sind. Die Anteile von Klassenunterricht und Gruppenlernen halten sich dabei in etwa die Waage. Eine thematische Einheit umfasst in der Regel ca. vier Lektionen. Das Gruppenlernen findet in heterogenen 4er oder 5er-Gruppen statt. In den Gruppen versuchen die Schüler, zuerst die Aufgaben in 2er (oder 3er)-Gruppen zu lösen. Bei Schwierigkeiten fragen sie die anderen Gruppenmitglieder. Wenn ein schwächerer Schüler eine Aufgabe nicht lösen kann, sind die anderen verantwortlich, ihm die Lösung zu erklären.

Vorgehensskript zur Steuerung der Gruppenprozesse: Lernwirksame Gruppenprozesse kommen durch rein organisatorische Maßnahmen kaum automatisch zustande. Den Lernenden sollten, wie jüngere Forschungsergebnisse gezeigt haben, geeignete Anleitungen zur Strukturierung der Kooperation zur Verfügung stehen (Ertl & Mandl, 2006; Kollar, Fischer & Hesse, 2006). Dies gilt folgerichtig auch für eine Rallye. Slavin (1988, S. 28) hat sich ursprünglich eher auf allgemeine Empfehlungen beschränkt. Scripts beschreiben demgegenüber eine Abfolge von Vorgehensschritten und dazugehöriges Rollenverhalten präziser (Dennen & Hoadley, 2013, S. 395). Es kann sich dabei um erforderliche Angaben zum gegenseitigen Erklären oder über Rollen- und Aufgabenverteilungen handeln. Strukturierendes und lenkendes „cooperative scripting" vermindert den „workload" beim Lernen. Die Schüler können sich auf die Bewältigung der Aufgaben konzentrieren und erreichen bessere Leistungen (Kollar, Ufer, Reichersdorfer, Vogel & Fischer, 2014; Prichard, Bizo & Stratford, 2011). Kooperative Skripts für eine Rallye haben Rotering-Steinberg, Kopensteiner & Tesar (1988) sowie Wahl (2004, S. 88) formuliert.

Tests und mögliche Belohnung: Eine Lerneinheit endet mit einem ca. 30-minütigen Leistungstest. Eine Gruppe kann nur dann erfolgreich abschließen, wenn möglichst alle den Stoff gut beherrschen und einander geholfen haben. Diese Voraussetzung wird mit einem speziellen Berechnungsverfahren sichergestellt. Bei der Berechnung des Gruppenresultates werden Verbesserungswerte bzw. individuelle Lernfortschritte verglichen zwischen einem Basistest, der zu Beginn der Rallye durchgeführt worden ist und einem Test nach einer durchgeführten Lerneinheit. Durch dieses Verfahren können schwächere Schüler den gleichen Beitrag zum Gruppenergebnis beisteuern wie gute Schüler. Die Ergebnisse werden den Gruppen öffentlich zurückgemeldet und besprochen. Gruppen mit hohen Punktzahlen können belohnt werden. Belohnungen sind ein Zusatzelement, das Lernende in einer Rallye motivieren kann, sich anzustrengen. Wenn Belohnungen gewährt werden, verbleibt als wesentlichste Unterstützungsmaßnahme jedoch nach wie vor die Einstellung der Lehrkräfte zur Kooperation und zum Lernerfolg in der Gruppe (Slavin, 1988, S. 30.). Neben den Gruppenleistungen werden im Übrigen auch Erfahrungen im Lernprozess, z. B. Probleme bei der Zusammenarbeit in den Gruppen, angesprochen.

4 Lernvoraussetzungen auf Seiten der Lernenden

Theoretische Überlegungen und empirische Befunde lassen vermuten, dass Merkmale der Lernenden den Lernerfolg im Fach Mathematik und damit auch die Wirksamkeit einer Rallye maßgeblich beeinflussen können. Diese individuellen Merkmale sollen in der Analyse kontrolliert werden. Neben den üblichen soziodemographischen Variablen sind es vor allem das Geschlecht, die Intelligenz und die sprachlichen Kompetenzen, die einen bedeutsamen Einfluss haben können.

SES: Die soziale Herkunft ist ein wichtiger schulischer Erfolgsfaktor. Kinder aus unteren Bildungsschichten haben weniger Gelegenheit, mathematikbezogene Erfahrungen zu sammeln, die förderlich für den schulischen Erfolg sind (Roberts & Bryant, 2011). Eine deutliche Koppelung zwischen dem sozialen Status und dem Bildungserfolg im Fach Mathematik berichten ferner Bonsen, Frey und Bos (2008), Pietsch und Krauthausen (2006) sowie Howie (2005). Hoffungsvoll stimmen hingegen die Befunde von Rohrbeck et al. (2003), nach denen insbesondere sozial benachteiligte Kinder von gegenseitigem Lernen in Gruppen (PAL: „peer-assisted learning“) profitieren.

Geschlecht: Mädchen erreichen im Fach Mathematik häufig tiefere Leistungen als Jungen. In der LAU-Studie berichten Lehmann, Peek und Gänsfuss (2011) eine Effektstärke in Mathematiktests von d = -.18 zu Ungunsten der Mädchen. Diese Ergebnisse werden auch bestätigt durch die Timss-Studie (Brehl, Wendt & Bos,

2011) sowie von der IGLU-Studie (Bos, Lankes, Prenzel, Schwippert & Walther 2004). Unterschiede bleiben auch im Längsschnitt bestehen (Wai, Cacchio, Putallaz & Makel, 2010).

Intelligenz: Fluide Intelligenz ist ein zentraler Prädiktor im Hinblick auf die Leistung im Fach Mathematik (Kyttäla & Letho, 2008; Primi, Ferraro & Almeida, 2010; Weber, Lu, Shi & Spinath, 2013). Allerdings wird dieser Zusammenhang durch zusätzliche Merkmale der Person wie die Aufmerksamkeit oder sprachliche Kompetenzen moderiert (Steinmayr, Ziegler & Trauble, 2010).

Sprachliche Kompetenzen: Eine mediierende Rolle beim Zusammenhang zwischen Intelligenz und Leistung wird insbesondere sprachlichen Kompetenzen zugeschrieben (Vista, 2013). Mathematische Probleme werden vielfach im Medium der Sprache präsentiert. Dies hat zur Folge, dass Informationen in Textform in mathematische Beziehungen zu transformieren sind (Tolar et al. 2012). Die Fähigkeit, solche Probleme zu lösen, ist abhängig von formalen Aspekten des dargebotenen Textmaterials sowie von der Erfahrungswelt und den sprachlichen Kompetenzen der Problemlöser (Blum et al., 2004; Kintsch & Greeno, 1985; Nathan, Kintsch & Young, 1992; Staub & Reusser, 1995). Sprachliche Fertigkeiten haben als Prädiktoren für mathematische Leistungen deshalb zunehmend Beachtung gefunden (Vukovich, 2012). Selbst Schüler mit verbalen Kompetenzen, die üblichen Anforderungen genügen, sind vielfach nicht in der Lage, spezifische Ansprüche zu erfüllen, die ein formales, mathematisches Problemverständnis verlangt (Lott Adams & McKoy Lowery, 2007). In besonderem Masse gilt dies für Lernende in multilingualen Klassen, die in kontext-reduzierten Settings wie dem Mathematikunterricht besonderen Schwierigkeiten ausgesetzt sind (Haag, Heppt, Stanat, Kuhl & Anand, 2013; Lehmann, Peek & Gänsfuss, 2011).

5 Fragestellungen

Eine Rallye ist so konzipiert, dass lernwirksame Grundprinzipien kooperativen Lernens erfüllt sein sollten (Slavin, 1988). Ein erster Grundsatz besagt, dass ein positives Ergebnis nur zustande kommt, wenn die individuelle Verantwortlichkeit jedes einzelnen Gruppenmitgliedes sichergestellt ist. Dies ist in einer Rallye der Fall, weil sich das Gruppenergebnis auf individuelle Verbesserungswerte stützt. Ferner hat beim kooperativen Lernen positive Interdependenz vorherrschend zu sein. Man erreicht das Ziel nicht auf Kosten anderer, sondern nur zusammen mit den übrigen Gruppenmitgliedern. Dies trifft in einer Rallye ebenfalls zu. Nur wenn Stärkere den Schwächeren helfen, ist die Gruppe erfolgreich. Im Weiteren müssen sich alle anstrengen, das gemeinsame Ziel zu erreichen. Dieser Kerngedanke wird in der Rallye erfüllt, indem alle Gruppenmitglieder den Stoff möglichst gut beherrschen und als Gruppe entsprechend viele Verbesserungswerte erreichen sollen. Es

wird somit erwartet, dass sich die Teilnahme an der Rallyeintervention im Vergleich zur Kontrollgruppe signifikant positiv auf den Lernerfolg auswirkt. Im Weiteren werden Merkmale identifiziert, die die Wirksamkeit der Intervention moderieren. Zu diesen Kontrollvariablen gehören Angaben zur sozialen Herkunft, zum Geschlecht, zur Intelligenz und im Hinblick auf kooperative Interaktionen in der Rallye auch zu sprachlichen Kompetenzen.

6 Methode

6.1 Design

Die Studie wurde im Rahmen des geltenden Stundenplans im Fach Mathematik durchgeführt. Daten der Vortests wurden im März 2011 erhoben. Die Intervention fand in den Monaten Mai/Juni 2011 statt und dauerte 3 ½ Wochen (ca. 15 Lektionen). Interventions- und Kontrollklassen behandelten zeitgleich denselben Stoff. Der Unterricht in den Interventionsklassen wurde als Gruppenrallye arrangiert. Die Kontrollklassen folgten dem üblichen Unterricht. Die Posttests schlossen unmittelbar an die Intervention an, bzw. in den Kontrollklassen zeitlich ebenfalls unmittelbar an den behandelten Lernstoff. Ein Follow-up Test fand fünf Monate später, im November 2011, statt. Die Tests wurden durch Testleiter administriert.

6.2 Stichprobe

Insgesamt nahmen 497 Fünftklässler aus dem deutschsprachigen Teil des Kantons Freiburg (Schweiz) am Projekt teil (48.6 % weiblich; Alter: M = 11.85 Jahre). Sie stammten aus 30 Klassen (22 Schulen). Die Aufteilung in jeweils 15 Interventions- (N = 262) und Kontrollklassen (N = 235) erfolgte nach dem Zufall. Kontrollklassen wurden nur aus Schulen rekrutiert, in denen keine Intervention stattfand. 95.2 % der Kinder, die an der Studie teilnahmen, waren in der Schweiz geboren. 83.3 % gaben an, mit ihren Eltern deutsch zu sprechen. Die größte Gruppe der Kinder aus Migrantenfamilien stammte aus dem ehemaligen Jugoslawien (5.5 % der Schüler).

6.3 Curriculare Festlegung und Intervention

Ein erstes Anliegen der Studie betraf curriculare Aspekte. Die STAD-Methode sollte für Lernaufgaben eingesetzt werden, die hauptsächlich in Textform und ikonischen Darstellungen vorlagen und in mathematische Modelle zu transformieren waren. Durch sozialen Austausch in der Rallye sollte ein gemeinsames Verständnis der zu lösenden Probleme herbeigeführt werden können. Ausgewählt wurden drei Lernumgebungen aus dem Lehrmittel „Schweizer Zahlenbuch 5“ von Affolter, Amstad, Doebeli und Wieland (2009).

Lernumgebung 1 „Staffellauf“: Dieser Kontext aus dem Sportbereich eignete sich, um die vier Grundoperationen praxisnah anzuwenden sowie Grafiken zu lesen und zu verstehen.

Lernumgebung 2 „Spitze“: Verglichen und berechnet wurden Geschwindigkeiten von Vögeln, Säugetieren und dem Menschen.

Lernumgebung 3 „Alter von Bäumen“: Aus Sachtexten waren Daten zur Lösung von Problemstellungen zum Pflanzenwachstum zu entnehmen.

Die Instruktion der Lehrpersonen erfolgte in drei Workshops. Diese fanden im September, im November und vor der Intervention im April des laufenden Schuljahres statt. Jede Veranstaltung dauerte drei Stunden. Zwischen den einzelnen Workshops konnten Erfahrungen mit dem Konzept gesammelt werden, über die man sich in der nachfolgenden Veranstaltung austauschen konnte. Theoretische und praxisbezogene Grundlagen der Workshops sind in (Niggli, 2013, S. 172ff.) detailliert dargestellt.

Im Workshop 1 (September 2010) wurden die Lehrkräfte mit dem STAD-Konzept vertraut gemacht. Im Zentrum stand die organisatorische Anlage (s. oben, erster Eckpunkt). Die Lehrkräfte erhielten eine ausgearbeitete Unterrichtseinheit, die sie im Unterricht einsetzen konnten. Sie hatten jedoch Spielraum bei der Umsetzung (vgl. „middle ground solutions“, Chavez-Lopes, 2003, S. 163). Das prototypische Beispiel war zusammen mit einem Expertenlehrer entwickelt worden, der zuvor Erfahrung mit der Rallyemethode gesammelt hatte und an allen drei Workshops mitwirkte. Ausgewählt wurden ebenfalls drei Lernumgebungen aus dem oben erwähnten Lehrmittel.

Im Workshop 2 (November 2010) wurden die Erfahrungen mit der durchgeführten Unterrichtseinheit aus dem Workshop 1 ausgewertet. Den thematischen Schwerpunkt in diesem Workshop bildete der Umgang mit einem speziell für diese Studie entwickelten Kooperationsskript (Niggli, 2013, S. 188, s. oben, zweiter Eckpunkt). Dieses lag als grafische Skizze vor, an der sich die Schüler in den Gruppen orientieren sollten. Im Zeitraum zwischen Workshop 2 und 3 wurden in den Schulklassen im Rahmen des üblichen Unterrichts Gelegenheiten geschaffen, Erfahrungen mit dem kooperativen Skript zu sammeln.

Im Workshop 3 (April 2011) wurde die Planung der drei Lernumgebungen an die Hand genommen, die für die Intervention und für die Datenerhebung maßgebend waren. Die Projektverantwortlichen moderierten den gesamten Planungsprozess. Dabei wurden primär die Lehrpersonen als Experten für die Gestaltung ihres Unterrichts angesehen, die selbst am besten entscheiden können, wie sie die Umsetzung unter den spezifischen Bedingungen in ihrer Klasse vornehmen können (Gräsel, 2011). Sie hatten die Lernumgebung organisatorisch zu planen und die

Leistungstests sowie die Veröffentlichungsform der Testresultate (s. oben, Eckpunkt 3) für die drei curricular festgelegten Lernumgebungen zu bestimmen. Den Lehrpersonen war freigestellt, ob sie die Gruppen mit den positivsten Resultaten belohnen wollten. Als Belohnung wurden ein oder zwei Bonuspunkte für die Schlussprüfung vereinbart. Diese Belohnungsstrategie fokussierte den Nutzen kooperativen Lernens zur Erreichung der Lernziele. Die Aufmerksamkeit wurde nicht auf Gegenstände verlagert, die keinen Bezug zum Lernen haben (vgl. Brophy, 2004, S. 176).

6.4 Treatmentcheck

Interventionsgruppe: Laut den Angaben der Lehrpersonen beanspruchte der Anteil selbständigen kooperativen Lernens gemäß dem Rallyekonzept durchschnittlich 54 % der Gesamtlernzeit (s. Tabelle 2). Dieser Wert entsprach den Empfehlungen der Projektverantwortlichen. Der vorangehende Unterricht mit der Klasse lag bei 43 % (3 % wurden für andere Formen aufgewendet). Über die Hälfte des zeitlichen Anteils, den die Organisationsform „Klassenunterricht" beanspruchte, wurde im Plenum unterrichtet. Die restliche Zeit mit der Klasse verteilte sich auf die drei gängigsten Sozialformen. Die vorgesehenen drei Lernumgebungen konnten in den 15 Klassen vollständig nach der Rallyemethode bearbeitet werden. In 10 Klassen wurden den Gruppen mit den positivsten Resultaten als Belohnung ein oder zwei Bonuspunkte gewährt, die für die individuelle Schlussprüfung angerechnet werden konnten. Bei dieser Belohnung handelt es sich wie oben erwähnt um ein Zusatzelement der Gruppenrallye.

Tabelle 2: Eingesetzte Organisations- und Sozialformen in der Interventions- und Kontrollgruppe (Prozentangaben)

	Interventionsgruppe				Kontrollgruppe			
Anteil Unterricht im Klassenverband	43.0				74.1			
	davon				davon			
	Plenum	Einzel	Partner	Gruppen	Plenum	Einzel	Partner	Gruppen
	54.4	15.1	21.8	8.6	48.0	16.3	24.3	10.0
Anteil „Rallyeko-operation" gem. Skript	54.0							
Sonstige offene Organisations-formen	Andere Nennungen				Wochenplan Stationenlernen		Andere Nennungen	
	3.0				21.6		4.3	

Kontrollgruppe: In der Kontrollgruppe wurden durchschnittlich 74.1 % der Lernzeit (gegenüber 43 % bei den Interventionsklassen) für den Unterricht mit der Klasse aufgewendet. Ca. die Hälfte dieser Organisationsform wurde im Plenum bestritten. Die restlichen im Klassenunterricht eingesetzten Sozialformen unterschieden sich unmerklich von den Werten der Interventionsgruppe. 21.6 % der Lernzeit waren durch Wochenplanunterricht oder durch Stationenlernen mit offenen Sozialformen bestimmt. Andere Nennungen („Parcours", „Staffellauf", „Hausaufgaben selbstständig Üben") beanspruchten 4.3 % der Unterrichtszeit.

Der Anteil feststellbaren kooperativen Lernens in Gruppen lag gemäß Tabelle 2 in den Interventionsklassen (54 % + 8.6 % Anteil im Klassenunterricht) infolgedessen deutlich höher als in der Kontrollgruppe (10 % Anteil im Klassenunterricht + nicht genau feststellbarer Anteil im Rahmen sonstiger offener Organisationsformen).

6.5 Messinstrumente

Mathematikleistungen: Zur Ermittlung relevanter Vorkenntnisse für die Bewältigung des Lernstoffs wurde ein standardisierter Vortest entwickelt. Ein weiterer curricularer Test wurde zum Lernstoff konstruiert, der in den drei Lernumgebungen zu behandeln war. Die Entwicklung der Tests wurde in einer Pilotstudie ein Jahr vor der Intervention an einer speziellen Stichprobe vorgenommen. Die Testskalierung wurde mit einem „3 parameter logistic item response theory model" mit der Software Bilog-MG (Zimowski, Muraki, Mislevy, & Bock, 2003) durchgeführt. Der zum Messzeitpunkt T1 eingesetzte Test zur Erfassung bedeutsamer Vorkenntnisse

enthielt 19 Aufgaben mit 23 dichotomen Items. Der curriculare Test, der zu T2 und T3 (follow-up) verwendet wurde, umfasste 18 Aufgaben mit 21 dichotomen Items. Die interne Konsistenz der beiden Tests (Kuder-Richardson-Formula 20) betrug 0.74 zu T1 und je 0.75 zu T2 und zu T3.

SES: Der Sozioökonomische Status (SES) wurde mit dem ISEI-Index (International Socio-Economic Index of Occupational Status; Ganzeboom, De Graaf, Treiman & De Leeuw, 1992) bestimmt. Die Angaben der Kinder zum Beruf ihrer Mutter bzw. ihres Vaters wurden von zwei Ratern gesondert eingeschätzt. Nicht-Übereinstimmungen wurden kommunikativ validiert. Der höhere Score des jeweiligen Elternteils ging in die Analysen ein.

Familiensprache: Die Kinder wurden ebenfalls zur gesprochenen Sprache zu Hause befragt (Welches ist deine Muttersprache?). Elf der am häufigsten gesprochenen Sprachen waren vorgegeben und konnten angekreuzt werden. Zudem konnte eine Kategorie „andere Sprachen" markiert werden. Die Variable wurde in eine Dummy-Variable (deutsch vs. andere Sprache) transformiert.

Intelligenz: Die kognitiven Fähigkeiten wurden mit dem sprachunabhängigen Intelligenztest CFT 20-R (Weiss, 1998) gemessen. Dieser Test enthält vier Subtests mit Aufgaben zu folgenden Bereichen: Reihenfortsetzen, Klassifikationen, Matrizen und topologische Schlussfolgerungen. Aus den vier Subtests wurde ein Gesamtscore gebildet (Cronbachs $\alpha = .59$).

Sprachliche Kompetenzen: Zur Erfassung des Leseverständnisses wurde der Lesetest ELFE 1-6 von Lenhard und Schneider (2006) verwendet. Eingesetzt wurde der Subtest, der das Leseverständnis auf Textebene misst (Auffinden von Informationen, satzübergreifendes Lesen, schlussfolgerndes Denken). Die interne Konsistenz (Kuder-Richardson-Formula 20) betrug 0.80 zu T1 und 0.79 zu T2.

6.6 Statistisches Vorgehen

Aufgrund der hierarchischen Datenstruktur wurden zur Überprüfung der Hypothesen Mehrebenenanalysen durchgeführt (Raudenbush & Bryk, 2002). Das verwendete Softwarepaket für Mehrebenenanalysen (HLM 6.04, Raudenbush, Bryk, Cheong & Congdon, 2004) bietet allein unstandardisierte Regressionskoeffizienten in der Ergebnisausgabe an, was die Interpretation der Regressionsgewichte aufgrund der oftmals arbiträren Metriken der Prädiktoren und Kriterien erschwert. Zur vereinfachten Interpretation von Effekten der Prädiktorvariablen wurden im vorliegenden Fall alle metrischen Variablen y-standardisiert ($M = 0$, $SD = 1$). Die Regressionskoeffizienten zeigen dann an, um welchen Anteil einer Standardabweichung sich die abhängigen Variablen bei der Zu- oder Abnahme einer Prädiktorvariable um eine Einheit verändern. Die deskriptiven Analysen wurden mit dem Statistikpaket SPSS 17 gerechnet.

Der Durchschnitt der fehlenden Werte lag bei 2.1 % pro Variable. Aufgrund der tiefen Anzahl Ausfälle und der Annahme, dass die Daten völlig zufällig fehlten (engl. missing completely at random, MCAR; vgl. Little & Rubin, 2002), wurden die MDM-files (HLM 6.04, Raudenbush et al., 2004) aufgrund der Rohdaten erstellt. Dieses Vorgehen resultiert für die Analysen in einer „listwise deletion".

Für die Berechnung der Effektstärke wurde die Formel für dichotome Variablen in Mehrebenenmodellen verwendet (vgl. Tymms, Merrell & Henderson, 1997):

$$\Delta = \frac{\beta_1}{\sigma_e}$$

In Worte gefasst bedeutet dies: Effektstärke = Regressionskoeffizient des dichotomen Prädiktors/gepoolte Standardabweichung (Wurzel der Varianz auf Schülerebene). Verwendet wurde jeweils die Varianz auf Schülerebene nach der Kontrolle der im Modell verwendeten Variablen. Diese Varianz ist kleiner als die Varianz auf der Schülerebene im Nullmodell und führt daher zu einer höheren Effektstärke. Im vorliegenden Fall ist die Verwendung der reduzierten Varianz zur Berechnung der Effektstärke angebracht, da das Interesse auf dem Effekt der Intervention unter Kontrolle von Vorleistung und weiterer Kontrollvariablen beruht.

7 Ergebnisse

7.1 Deskriptive Analysen und Interkorrelationen der Variablen

In Tabelle 3 werden Mittelwerte, Standardabweichungen und Interkorrelationen der Variablen berichtet. Zu beachten ist, dass sich der Mathematiktest stoffrelevanter Vorkenntnisse zu T1 inhaltlich vom curricularen Test zu T2 und T3 unterscheidet, mit dem die in den drei Lernumgebungen erreichten Leistungen gemessen worden sind. Zwischen den Messzeitpunkten T2 und T3 ist ein Anstieg der mathematischen Leistungen zu registrieren. Die Korrelationen zwischen den Mathematiktests und dem sprachunabhängigen Intelligenztest liegen höher als die Zusammenhänge mit dem Textverständnis. Hoch sind die Interkorrelationen der drei Mathematiktests ausgefallen. Die Korrelationen zwischen den Leistungstests und dem beruflichen Status der Eltern liegen höher als mit der gesprochenen Sprache zu Hause. Das Geschlecht (Jungen = 1) korreliert negativ mit dem Textverständnis, jedoch positiv mit den Mathematikleistungen.

Tabelle 3: Mittelwerte, Standardabweichungen und Korrelationen

	M	SD	1	2	3	4	5	6	7	8
Geschlecht (1=männl.)	a	-								
Deutsch (=1)	a	-	.05							
ISEI_ELTERN	49.92	16.93	-.02	.16**						
CFT	35.87	5.92	-.06	.00	.16**					
ELFE-Text-verständnis	16.11	3.11	-.17**	.11*	.19**	.26**				
Mathtest t1	10.59	4.15	.09*	.12**	.22**	.45**	.38**			
Mathtest t2	9.82	3.97	.16**	.16**	.22**	.48**	.39**	.69**		
Mathtest t3	10.59	3.88	.20**	.11*	.16**	.43**	.37**	.67**	.74**	
Intervention (=1)	a	-	.01	-.02	.00	-.02	.00	-.02	.08	.07

Anmerkungen: a dichotome Variablen
*p < .05
**p<.01

In Tabelle 4 sind die Mittelwerte und Standardabweichungen der Mathematiktests zu den drei Messzeitpunkten für beide Probandengruppen aufgeführt.

Tabelle 4: Mittelwerte und Standardabweichungen der Probandengruppen im Vortest und in den curricularen Nachtests

	Interventionsgruppe		Kontrollgruppe	
	M	SD	M	SD
Vortest	10.53	4.08	10.60	4.29
Nachtest	10.13	3.91	9.51	4.06
Follow-up	10.89	3.82	10.32	3.96

Im Vortest erzielten beide Gruppen ähnliche Resultate. Der Mittelwertsunterschied im curricularen Test nach der Intervention von 0.62 blieb bis zum Follow-up mit 0.57 nahezu stabil.

7.2 Vorhersage der Leistungsentwicklung

In Tabelle 5 sind die acht Modelle zur Vorhersage der Mathematikleistung zu T2 dargestellt.

Als Kovariaten wurden auf der Individualebene neben den relevanten mathematischen Vorkenntnissen zu T1 die Familiensprache, der beruflicher Status der Eltern

sowie das Geschlecht des Kindes, seine nonverbale Intelligenz und das Textverstehen eingeführt. Auf der Klassenebene wurde als Dummy-Variable der Einfluss der Intervention (Referenz: Kontrollgruppe) untersucht. Modell M1 analysiert den Zusammenhang zwischen den mathematischen Vorkenntnissen zu T1 und der Leistung zu T2. Modell M2 prüft den Einfluss der Intervention auf diese Beziehung. In den nächsten Modellen M3 bis M7 wurde untersucht, inwieweit dieser Effekt durch die Hinzunahme weiterer Prädiktoren beeinflusst wird. Modell M8 berücksichtigt alle Prädiktoren simultan.

Der Effekt der Intervention auf Klassenebene blieb über die sieben Modelle hinweg nahezu konstant. Der Regressionskoeffizient variierte zwischen B = .16 und B = .18. und war statistisch signifikant (einseitiger Signifikanztest). Auffällig ist, dass der Einfluss des Geschlechts von ursprünglich B = .18 in Modell M3 im Gesamtmodell M8 auf B = .31 anstieg. Die restlichen Level 1 - Koeffizienten waren im Modell M8 alle statistisch signifikant und unterschieden sich gegenüber den Modellen M4 bis M7 nur in einem geringen Ausmaß.

Ein ähnliches Befundmuster wie zwischen T1 und T2 war auch für die Voraussagen zum Messzeitpunkt T3 zu registrieren (vgl. Tabelle 6). Die Werte der Koeffizienten der Interventionsvariablen veränderten sich kaum. Der Interventionseffekt war wiederum in allen Modellen statistisch signifikant (einseitiger Signifikanztest). Der Einfluss des Geschlechts war im Gesamtmodell M8 mit B = .40 wiederum bemerkenswert hoch, verglichen mit Modell M3 (B = .28). Einen geringeren Effekt als zu T2 hatten zu T3 die Variablen des familiären Hintergrundes (Familiensprache, Berufsstatus).

Die Effektstärken für die Intervention variierten in den einzelnen Modellen zwischen $\Delta = 0.22$ und $\Delta = 0.26$. Für beide Messzeitpunkte konnten zudem keine Interaktionseffekte zwischen den Vorkenntnissen der Schüler und der Intervention festgestellt werden. Schwächere und stärkere Schüler profitierten somit von der Intervention in ähnlichem Ausmaße. Derselbe Befund konnte für das Geschlecht der Schüler registriert werden.

Voraussagen auf die abhängigen Variablen „Kompetenzerleben“ und „Interesse am Fach“ wurden ebenfalls untersucht. Es konnten jedoch keine statistisch signifikanten Effekte registriert werden.

Tabelle 5: Befunde (Regressionskoeffizienten) aus Mehrebenenanalysen zur Vorhersage der Mathematikleistung zum Zeitpunkt T2

		M1			M2		
		B	SE	p	B	SE	p
Level 1	Vorkenntnisse T1	0.69	0.03	<0.000	0.69	0.03	<0.000
	Geschlecht (m)						
	Erstsprache Deutsch						
	ISEI						
	CFT						
	Textverständnis ELFE						
Level 2	Intervention Rallye				0.16	0.08	0.032 a
	R2	0.48			0.48		

Anmerkungen: a der gerichteten Hypothese entsprechend, p-Werte für einseitigen Signifikanztest

Tabelle 5: Fortsetzung

		M3			M4		
		B	SE	p	B	SE	p
Level 1	Vorkenntnisse T1	0.68	0.03	<0.000	0.67	0.03	<0.000
	Geschlecht (m)	0.18	0.04	<0.000			
	Erstsprache Deutsch				0.17	0.09	0.060
	ISEI						
	CFT						
	Textverständnis ELFE						
Level 2	Intervention Rallye	0.16	0.08	0.031 a	0.17	0.08	0.026 a
	R2	0.49			0.48		

Anmerkungen: a der gerichteten Hypothese entsprechend, p-Werte für einseitigen Signifikanztest

Tabelle 5: Fortsetzung

		M5			M6		
		B	SE	p	B	SE	p
Level 1	Vorkenntnisse T1	0.67	0.03	<0.000	0.59	0.03	<0.000
	Geschlecht (m)						
	Erstsprache Deutsch						
	ISEI	0.08	0.03	0.017			
	CFT				0.21	0.04	<0.000
	Textverständnis ELFE						
Level 2	Intervention Rallye	0.16	0.09	0.034 a	0.18	0.08	0.023 a
	R2	0.48			0.52		

Anmerkungen: a der gerichteten Hypothese entsprechend, p-Werte für einseitigen Signifikanztest

Tabelle 5: Fortsetzung

		M7			M8		
		B	SE	p	B	SE	p
Level 1	Vorkenntnisse T1	0.64	0.03	<0.000	0.49	0.04	0.000
	Geschlecht (m)				0.31	0.06	0.000
	Erstsprache Deutsch				0.15	0.07	0.046
	ISEI				0.07	0.03	0.048
	CFT				0.23	0.03	0.000
	Textverständnis ELFE	0.14	0.03	<0.000	0.14	0.03	0.000
Level 2	Intervention Rallye	0.16	0.08	0.028 a	0.16	0.07	0.016 a
	R2	0.50			0.56		

Anmerkungen: a der gerichteten Hypothese entsprechend, p-Werte für einseitigen Signifikanztest

Tabelle 6: Befunde (Regressionskoeffizienten) aus Mehrebenenanalysen zur Vorhersage der Mathematikleistung zum Zeitpunkt T3

		M1			M2		
		B	SE	p	B	SE	p
Level 1	Vorkenntnisse T1	0.68	0.02	0.000	0.68	0.02	0.000
	Geschlecht (m)						
	Erstsprache Deutsch						
	ISEI						
	CFT						
	Textverständnis ELFE						
Level 2	Intervention Rallye				0.16	0.09	0.034 a
	R2	0.45			0.45		

Anmerkungen: a der gerichteten Hypothese entsprechend, p-Werte für einseitigen Signifikanztest

Tabelle 6: Fortsetzung

		M3			M4		
		B	SE	p	B	SE	p
Level 1	Vorkenntnisse T1	0.66	0.03	0.000	0.66	0.02	0.000
	Geschlecht (m)	0.28	0.06	0.000			
	Erstsprache Deutsch				0.08	0.07	0.297
	ISEI						
	CFT						
	Textverständnis ELFE						
Level 2	Intervention Rallye	0.16	0.08	0.037 a	0.18	0.09	0.023 a
	R2	0.47			0.45		

Anmerkungen: a der gerichteten Hypothese entsprechend, p-Werte für einseitigen Signifikanztest

Tabelle 6: Fortsetzung

		M5			M6		
		B	SE	p	B	SE	p
Level 1	Vorkenntnisse T1	0.67	0.03	0.000	0.60	0.02	0.000
	Geschlecht (m)						
	Erstsprache Deutsch						
	ISEI	0.02	0.03	0.556			
	CFT				0.16	0.03	0.000
	Textverständnis ELFE						
Level 2	Intervention Rallye	0.17	0.09	0.031 a	0.17	0.09	0.029 a
	R2	0.45			0.47		

Anmerkungen: a der gerichteten Hypothese entsprechend, p-Werte für einseitigen Signifikanztest

Tabelle 6: Fortsetzung

		M7			M8		
		B	SE	p	B	SE	p
Level 1	Vorkenntnisse T1	0.63	0.03	0.000	0.51	0.03	0.000
	Geschlecht (m)				0.40	0.07	0.000
	Erstsprache Deutsch				0.05	0.07	0.519
	ISEI				0.01	0.03	0.772
	CFT				0.17	0.03	0.000
	Textverständnis ELFE	0.12	0.04	0.001	0.14	0.03	0.000
Level 2	Intervention Rallye	0.16	0.08	0.030 a	0.17	0.08	0.022 a
	R2	0.47			0.52		

Anmerkungen: a der gerichteten Hypothese entsprechend, p-Werte für einseitigen Signifikanztest

8 Diskussion

Das Ziel der vorliegenden Studie war es zu klären, ob der Lernerfolg bei Problemlöseaufgaben im Mathematikunterricht von Fünftlässlern durch ein Gruppenrallye, bzw. ein STAD-Programm (Student teams achievement divisions) in bedeutsamer Weise gefördert werden kann. Zudem wurde der Frage nachgegangen, welche Rolle wichtigen Merkmalen der Schüler dabei zukommt. Dies betraf das Ge-

schlecht, die kognitiven Voraussetzungen sowie Indikatoren des familiären Hintergrundes.

Für die beiden untersuchten Messzeitpunkte (Posttest und Follow-up) konnte im Vergleich zur Kontrollgruppe ein moderater, aber stabiler Wissensvorsprung der Experimentalgruppe registriert werden. Dieser Befund gilt für das Gesamtmodell sowohl unter Kontrolle von einzelnen als auch sämtlicher in der Regressionsgleichung eingesetzten Schülermerkmale. Diese Drittvariablen haben die Effekte der Intervention somit kaum beeinflusst. Die Tatsache, dass zwischen Posttest und Follow-up in den beiden Vergleichsgruppen ein deskriptiver Wissenszuwachs zu verzeichnen war, kann auf zwischenzeitliche Entwicklungsprozesse im mathematischen Denken dieser Alterskohorte und auf Lernerfolge im nachfolgenden Unterricht zurückgeführt werden. Die im Posttest festgestellte Differenz zwischen den Vergleichsgruppen blieb im Follow-up-Test trotzdem erhalten. Dies ist ein Beleg für die Stabilität der in der Rallye erzielten Leistungsfortschritte.

Der moderat hohe Wert der statistisch durchwegs signifikanten Effektstärken bedeutet nicht, dass solche Wirkungen als wenig bedeutsam einzuschätzen wären. Vor allem in Anbetracht der relativ kurzen Dauer der Intervention sind die erzielten Ergebnisse beachtenswert. Die Implementation eines didaktischen Konzepts wird sowohl von den Lehrpersonen als auch von den Schülern subjektiv verarbeitet bzw. rekonstruiert. In den 15 Klassen wurden somit 15 Varianten durchgeführt, was den Lehrpersonen im Rahmen möglicher Spielräume auch ausdrücklich zugestanden worden war. Ihre angewandten Maßnahmen auf der Makroebene wie der Anteil kooperativen Lernens oder das Berechnungsverfahren bei den Lernkontrollen konnten mittels Treatmentcheck relativ zuverlässig kontrolliert werden. (So scheint die Gewährung von Bonuspunkten für die Schlussprüfung beispielsweise die Ergebnisse in den betreffenden zehn Klassen nicht wesentlich beeinflusst zu haben. Die Leistungen in den fünf Klassen ohne Bonus; aber ebenfalls mit Veröffentlichung der Gruppenresultate, lagen sogar etwas höher.) Hingegen war nicht eindeutig zu klären, inwieweit es den Lehrkräften gelungen war, auf der Mikroebene die Grundidee der Kooperation tatsächlich ins Zentrum zu rücken. Dazu wären aufwändige Beobachtungsstudien notwendig. Kooperative Unterrichtsformen lassen sich in ihrer Effektivität jedoch vor allem dann steigern, wenn die Schüler auf der Mikroebene über ausreichende Kommunikationsfähigkeiten verfügen, z.B. wenn sie Anleitungen, ihre Arbeitsprozesse inhaltlich zu strukturieren und im Ablauf zu steuern, tatsächlich in die Tat umsetzen (Lipowsky, 2006). Das Training dieser Kompetenzen ist allerdings eine längerfristige Aufgabe. Dies ist auch deshalb der Fall, weil Anregungen zur Kooperation mit kulturellen westeuropäischen Mustern konkurrieren, die das individuelle Lernen favorisieren (Kyndt et al., 2013). Im vorliegenden Fall wurden die Lehrpersonen angeregt, in der verfügbaren Zeit-

spanne, die dem Treatment vorausging, entsprechende Übungen einzusetzen (vgl. oben Workshop 2). Infolgedessen ist bei der Beurteilung der Effektstärken davon auszugehen, dass eine Entwicklung zur Kooperation zwar eingeleitet werden konnte, diese aber keineswegs als besonders fortgeschritten anzusehen ist. Auch eine mehrfache Durchführung von Gruppenrallyes könnte diese Vorbehalte nicht entkräften. Nach wie vor bliebe offen, ob die inkrementellen positiven Effekte der Intervention aufgrund der eingespielten Kooperation zunehmen oder beispielsweise infolge Gewöhnungseffekten sogar abnehmen würden.

Ein weiterer Grund, die erzielten Effekte angemessen zu gewichten, liegt im Verhältnis zum geleisteten Aufwand. Für die Durchführung der Intervention konnten keine außerordentlichen Mittel eingesetzt werden. Genutzt wurden bestehende Fortbildungsressourcen. Wenn mit einem eher kleinen Aufwand (drei Workshops von je drei Stunden Dauer) eine Veränderung bewirkt werden kann, dann ist ein solches Ergebnis auch von praktischer Bedeutung (vgl. Prentice & Miller, 1992), zumal der Fortbildungsaufwand durch eine Verlängerung der Intervention vermutlich nur geringfügig höher ausfallen würde.

Ein Blick auf die moderierenden Faktoren auf der Individualebene wird wohl am ehesten der Tatsache Beachtung schenken, dass die Leistungsentwicklung der Mädchen bei gleichzeitiger Kontrolle der übrigen individuellen Merkmale ungünstiger verlief als diejenige der Knaben.

Die Deutlichkeit des Effektes überrascht. Es ist anzunehmen, dass dabei individuelle Persönlichkeitsmerkmale auf der einen Seite und außerschulische Voraussetzungen des Kompetenzerwerbs zusammenspielten (Brehl et al., 2011). Nachanalysen mit zusätzlich eingesetzten Messinstrumenten ergaben, dass das Interesse der Mädchen an Mathematik tatsächlich geringer war. Ihr fachspezifisches Selbstkonzept lag ebenfalls tiefer, die geäußerte Prüfungsangst dagegen höher. Diese Eigenschaften und Einstellungen dürften in Interaktion mit dem sozialen Umfeld gelernt worden sein, wobei im vorliegenden Fall offen bleiben muss, welche sozialisationsbedingten Aspekte im Einzelnen eine Rolle gespielt haben. Weil keine Interaktion zwischen der Intervention und dem Geschlecht feststellbar war, kann davon ausgegangen werden, dass bei diesen sozialisationsbedingten Merkmalen durch kurzfristige Interventionen kaum Veränderungen zu bewirken sind.

Erwartungsgemäß standen die nonverbale Intelligenz und auch das Verstehen von Texten im Zusammenhang mit dem Lernerfolg. Ein geringerer Einfluss war bei den beiden Merkmalen des familiären Hintergrundes festzustellen. Der Effekt der Familiensprache erwies sich zum Zeitpunkt T2 knapp als nicht signifikant. Zum Zeitpunkt T3 war er deutlich geringer als zu T2. Dabei ist zu beachten, dass 95 % der Kinder angaben, in der Schweiz geboren zu sein. Ein ähnlicher Rückgang war für

den sozioökonomischen Status zu verzeichnen. Ein statistisch signifikanter Effekt wie in M5 zu T2 war zum Zeitpunkt T3 nicht mehr nachzuweisen. Zwischen den beiden Messzeitpunkten können kognitive Entwicklungsprozesse eingesetzt haben, die den Einfluss des familiären Hintergrundes gemildert haben. Fehlende Interaktionen zwischen der Intervention und diesen Variablen auf der Individualebene bedeuten, dass die Schüler unabhängig von ihren individuellen Voraussetzungen vom Treatment gleichermaßen profitieren konnten. Dieser Befund spricht für den Einsatz der Methode im Unterricht.

Einschränkend ist zu erwähnen, dass sich die Intervention auf ca. vier Wochen zu konzentrieren hatte. Dies ist etwa ein Drittel der minimalen Zeitspanne, die in den Studien beansprucht worden war, die Slavin (2008) berücksichtigt hat. In solchen Fällen sind jedoch auch die lokalen schulkulturellen Bedingungen in Rechnung zu stellen. Lehrpersonen erklären sich nur selten bereit, ihren Unterricht für eine längere Zeitspanne auf eine bestimmte Methode umzustellen. In der Regel setzen sie je nach Lernstoff und Lernprozess unterschiedliche Arrangements ein wie Lernen auf individuellen Wegen, Didaktik des Entdeckens, usw.. Hinzu kommt, dass sich der Wettbewerbscharakter einer Rallye im Laufe der Zeit auch bei den Schülern abschwächen dürfte. Ebenso könnte sich die Motivation in den Gruppen verringern, wenn die Rollen der Mitglieder (als Erklärende oder Erklärungen verlangende) längerfristig an ein bestimmtes Skript gekoppelt sind.

Wenn die Implementation einer Methode für die Schüler eine neuartige Umweltsituation darstellt, können zudem Novitätseffekte auftreten. Wir gehen jedoch davon aus, dass eine Rallye keine künstliche, von der normalen Unterrichtssituation deutlich abweichende Ausnahme darstellt. Gruppenarbeiten werden auch im üblichen Unterricht durchgeführt. In der Rallye werden sie intensiviert und durch ein Anleitungsskript strukturiert. Akzentuiert werden darin die individuelle Verantwortlichkeit und die gegenseitige Unterstützung, die im Unterricht auch sonst eine Rolle spielen. Auch formative Lernkontrollen sind den Schülern bekannt. Neu ist lediglich die Berechnungsgrundlage für das Gruppenergebnis. Zu Beginn des Schuljahres wurde zudem bereits eine vorbereitete prototypische Einheit durchgeführt (s. Workshops 1). Die wöchentliche Wiederholung der Methode dürfte Novitätseffekte ebenfalls mildern. Zudem wurde mit dem Lehrmittel gearbeitet, das auch sonst im Unterricht verwendet worden ist.

Bei der Interpretation der Ergebnisse bleibt zudem zu beachten, dass Problemlöseaufgaben dominierten, die hauptsächlich in sprachlicher Form und bildlichen Darstellungen vorlagen und daher auch einen verbalen Austausch in den Gruppen erforderlich machten. Komplexe Anforderungen wurden gewählt, weil sich diese Aufgabenstellungen für den Lernerfolg beim kooperativen Lernen als günstig erwiesen haben (Kirschner, Paas, Kirschner & Janssen, 2011) und weil sie verlangen,

dass mathematische Fähigkeiten in unterschiedlichen Kontexten eingesetzt werden (vgl. „Mathematical Literacy“). Damit Schlüsse auf das Fach Mathematik insgesamt gezogen werden können, wären zusätzlich auch Effekte auf innermathematische Beziehungen zu prüfen, beispielsweise die Beherrschung mathematischer Operationen wie diejenige der schriftlichen Addition usw. (s. Tabelle 1: Mevarech, 1985). Den Kern der Mathematical Literacy bildet jedoch das Anwenden mathematischer Begriffe in unterschiedlichen Situationen. Für die Entwicklung dieser zentralen mathematischen Kompetenzen sind unsere Ergebnisse von Bedeutung.

Zukünftige Forschung könnte mittels qualitativen Analysen zudem untersuchen, welchen Einfluss gruppendynamische Prozesse haben und ob sich Mädchen und Jungen in den Gruppen unterschiedlich verhalten. Aufgrund fehlender statistisch signifikanter Zusammenhänge der Methode mit gängigen Merkmalen wie Kompetenzerleben oder Fachinteresse könnte die Analyse auf neuere psychologische Erkenntnisse, z. B. über die Strukturierung des Selbst im Klassenraum ausgedehnt werden. Die Mindset-Forschung von Dweck (2007) beispielsweise legt nahe, kritisch zu prüfen, ob Gruppenrallyes durch den Fokus auf Verbesserung und gegenseitiger Unterstützung einen Wechsel von einem „fixed mindset“ zu einem „growth mindset“ fördern, oder ob auch unbeabsichtigte Nebenwirkungen in Kauf genommen werden, z. B. wenn Individuen durch den Vergleich mit anderen eher dazu tendieren, eigene Fähigkeiten als fixiert anzusehen.

Im Hinblick auf die praktische Umsetzung erstaunt, wie dünn die empirische Basis hinter den didaktischen Ratschlägen ist, die aus ihnen abgeleitet werden. Im vorliegenden Fall erhielten die Lehrpersonen begleitenden Support. Die Effekte waren trotzdem eher moderat. Eine realistische Einschätzung dieser Methode ist auch deshalb angezeigt, weil sie differenzierte planerische Vorkehrungen verlangt, die für eine gelingende Durchführung entscheidend sein können (vgl. Niggli, 2013). Didaktische Empfehlungen, so macht es den Anschein, lassen sich zwar mit guten Gründen vorbringen. Bei der Umsetzung treffen sie jedoch auf die jeweiligen Umstände in einer Schulklasse.

Literatur

Affolter, W., Amstad, H., Doebeli, M. & Wieland, G. (2009). Schweizer Zahlenbuch 5. Zug: Klett und Balmer Verlag.

Blum, W., Neubrand, M., Ehmke, T., Senkbeil, M., Jordan, A., Ulfig, F. & Carstensen, C. H. (2004). Mathematische Kompetenz. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leitner, M. Neubrand, R. Pekrun, H. G. Rolff, J. Rost & U. Schiefele (Hrsg.), Pisa 2003. Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs (S. 47-92). Münster: Waxmann.

Bonsen, M., Frey, K. A. & Bos, W. (2008). Soziale Herkunft. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 141-156). Münster: Waxmann.

Borsch, F. (2010). Kooperatives Lehren und Lernen. Stuttgart: Kohlhammer.

Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K. & Walther, G. (2004). IGLU: Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich. Münster: Waxmann.

Brehl, T., Wendt, H. & Bos, W. (2011). Geschlechtsspezifische Unterschiede in mathematischen und naturwissenschaftlichen Kompetenzen. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 203-230). Münster Waxmann.

Brophy, J. (2004). Motivating Students to Learn. Mahwah NJ, London: Lawrence Erlbaum Ass. Publ..

Chavez-Lopes, O. (2003). From the textbook to the enacted curriculum: Textbook use in the middle school mathematics classroom. Doctoral Dissertation, University of Missoury - Columbia. Verfügbar unter: http://zeta.math.utsa.edu/~hvz231/dissertation/dissertation.pdf [02.08.2014].

Dennen, V.P. & Hoadley, C. (2013). Designing collaborative learning thraugh computer support. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chann & A. O'Donnel (Ed.), The International Handbook of Collaborative Learning (p. 389 – 402). New York & London: Routledge.

Dweck, C. S. (2007). Mindset. The new psychology of success. New York: Ballantine.

Ertl, B. & Mandl, H. (2006). Kooperationsskripts. In H. Mandl & F. F. Helmut (Hrsg.), Handbuch Lernstrategien (S. 271-281). Göttingen: Hogrefe.

Ganzeboom, H. B. G., De Graaf, P. M., Treiman, D. J. & De Leeuw, J. (1992). A standard international socioeconomic index of occupational status. Social Science Research, 21(1), 1-56.

Ginsburg-Block, M. D., Rohrbeck, C. A. & Fantuzzo, J. W. (2006). A meta analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. Journal of Educational Psychology, 98(4), 732-749.

Glassmann, P. (1989). A study of cooperative learning in mathematics, writing, and self-esteem by gender, race, and ability group. Dissertation, Hofstra University, Hempstead, NY. Verfügbar unter: http://dissexpress.umi.com/dxweb/results.html?QryTxt=&By=&Title=&pubnum=8923028 [04.09.2014].

Gräsel, C. (2011). Die Kooperation von Forschung und Lehrer/-innen bei der Realisierung didaktischer Innovationen. In W. Einsiedler (Hrsg.), Unterrichtsentwicklung und Didaktische Entwicklungsforschung (S. 88-104). Bad Heilbrunn: Klinkhardt.

Haag, N., Heppt, B., Stanat, P., Kuhl, P. & Anand, H. (2013). Second language learners performance in mathematics: Disentangling the effects of academic language features. Learning and Instruction, 28(1), 24-34.

Hasselhorn, M. & Gold, A. (2013). Pädagogische Psychologie. Erfolgreiches Lehren und Lernen. Stuttgart: Kohlhammer.

Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. New York: Routledge.

Howie, S. (2005). System-Level Evaluation: Language and other background factors affecting methematics achievement. Prospects: Quarterly Review of Comparative Education, 35(2), 175-186.

Huber, A. A. (2008). Kooperatives Lernen - kein Problem. Effektive Methoden der Partner- und Gruppenarbeit. Leipzig: Ernst Klett.

Huber, G. L. (1985). Lernen in Schülergruppen. Tübingen: Deutsches Institut für Fernstudien an der Universität Tübingen.

Huber, G. L. (1997). Self-regulated learning by individual students. In D. Stern & G. L. Huber (Eds.), Active learning for students and teachers. Report from eight countries (pp. 137-158). Frankfurt/M: Lang/OECD.

Johnson, D. W. & Johnson, R. T. (1999). Learning together and alone: Cooperative, competitive and individualistic learning. Boston, MA: Allyn and Bacon.

Johnson, D. W., Johnson, R. T. & Stanne, M. E. (2000). Cooperative learning methods: A meta-analysis. Verfügbar unter: http://www.ccsstl.com/sites/default/files/Cooperative%20Learning%20Research%20.pdf [02.02.2015.]

Kintsch, W. & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. Psychological Review, 92(1), 109-129.

Kirschner, F., Paas, F., Kirschner, P. A. & Janssen, J. (2011). Differential effects of problem-solving demands on individual and collaborative learning outcomes. Learning an Instruction, 21(4), 587-599.

Kollar, I., Fischer, F. & Hesse, F. W. (2006). Collaboration scripts - a conceptual analysis. Educational Psychology Review, 18(2), 159-185.

Kollar, I., Ufer, S., Reichersdorfer, E., Vogel, F. & Fischer, F. (2014). Effects of collaboration scripts and heuristic worked examples on the acquisition of mathematical argumentation skills of teacher students with different levels of prior achievement. Learning and Instruction, 32, 22-36.

Kotsopoulos, D. (2010). When collaborative is not collaborative: Supporting student learning through self-surveillance. International Journal of Educational Research, 49(4-5), 129-140

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E. & Dochy, R. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? Educational Research Review, 10, 133-149.

Kyttälä, M. & Letho, J. E. (2008). Some factors underlying mathematical performance: The role of visuospatial memory and non-verbal intelligence. European Journal of Psychology of Education, 23(1), 77-94.

Lehmann, R. H., Peek, R. & Gänsfuss, R. (2011). LAU 5 Aspekte der Lernausgangslage in der Lernentwicklung - Klassenstufe 5. In Behörde für Schule und Berufsbildung (Hrsg.), LAU - Aspekte der Lernausgangslage und der Lernentwicklung (S. 15-119). Münster: Waxmann.

Lenhard, W. & Schneider, W. (2006). ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler. Göttingen: Hogrefe.

Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In C. Allemann-Ghionda & E. Terhart (Hrsg.), Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern (S. 47-79). Weinheim: Beltz.

Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), Pädagogische Psychologie (S. 69-105). Berlin Heidelberg: Springer.

Little, R. J. A. & Rubin, D. B. (2002). Statistical analysis with missing data. New York: Wiley.

Lott Adams, T. & McKoy Lowery, R. (2007). An analysis of children's strategies for reading mathematics. Reading & Writing Quarterly, 23(2), 161-177.

Lou, J., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B. & d'Appolonia, S. (1996). Within-class grouping: A meta-analysis. Review of Educational Research, 66(4), 423-458

Mevarech, Z. R. (1985). The effects of cooperative mastery learning strategies on mathematics achievement. Journal of Educational Research, 78(3), 372-377.

Mevarech, Z. R. (1991). Learning mathematics in different mastery environments. Journal of Educational Research, 84(4), 225-231

Nathan, M. J., Kintsch, W. & Young, E. (1992). A theory of algebra word problem comprehension and its implications for the design of learning environments. Cognition and Instruction, 9(4), 329-389.

Niggli, A. (2013). Didaktische Inszenierung binnendifferenzierter Lernumgebungen. Bad Heilbrunn: Klinkhardt.

Pietsch, M. & Krauthausen, G. (2006). Mathematisches Grundverständnis von Kindern am Ende der vierten Jahrgangsstufe. In W. Bos & M. Pietsch (Hrsg.), KESS 4 - Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen (S. 143-164). Münster: Waxmann.

Prentice, D. A. & Miller, D. T. (1992). When small effects are impressive. Psychological Bulletin, 112(1), 160-164.

Prichard, J. S., Bizo, L. A. & Stratford, R. J. (2011). Evaluating the effects of team-skills training on subjective workload. Learning and Instruction, 21(3), 429-440.

Primi, R., Ferrao, M. & Almeida, L. S. (2010). Fluid intelligence as a predictor of learning. A longitudinal multilevel approach applied to math. Learning and Individual Differences, 20(5), 446-451.

Raudenbush, S. W. & Bryck, A. S. (2002). Hierarchical linear models: Applications and data analysis methods. Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryck, A. S., Cheong, Y. F. & Congdon, R. (2004). HLM 6: Hierarchical linear and nonlinear modeling. Chicago, IL: Scientific Software International.

Roberts, G. & Bryant, D. (2011). Early mathematics achievement trajectories: English-language learner and native English-speaker estimates, using the Early Childhood Longitudinal Survey. Developmental Psychology, 47(4), 916-930.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W. & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic Review. Journal of Educational Psychology, 95(2), 240-257.

Rotering-Steinberg, S., Koppensteiner, W. & Tesar, E. (1988). Kooperative Organisationsformen für Erwachsene. Forum Pädagogik, 2(91-99).

Slavin, R. E. & Lake, C. (2008). Effective programs in elementary mathematics: A best evidence synthesis. Review of Educational Research, 78(3), 427-515

Slavin, R. E. (1984). Team assisted individualization: Cooperative learning and individualized instruction in the mainstreamed classroom. Remedial and Special Education, 5(6), 33.42.

Slavin, R. E. (1988). Student team learning. An overview and practical guide. Washington, DC: NEA Professional Library, National Education Association.

Slavin, R. E. (1994). Student team-achievement divisions. In S. Sharan (Ed.), Handbook of cooperative learning methods (pp. 43-69). Westport CT: Greenwood Press.

Slavin, R. E. (1995). Cooperative learning. Theory, research, and practice. Boston: Allyn and Bacon.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know and what we need to know. Contemporary Educational Psychology, 21(1), 43-69.

Slavin, R. E., Hurley, E. A. & Chamberlain, A. (2003). Cooperative learning and achievement: Theory and research. In W. M. Reynolds & G. E. Miller (Eds.), Handbook of Psychology (pp. 177-198). Hoboken, NJ: Wiley.

Springer, L., Stanne, M. E. & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. Review of Educational Research, 69(1), 21-51.

Staub, F. & Reusser, K. (1995). The role of presentational structures in understanding and solving mathematical word problems. In C. A. Weaver, S. Mannes & C. R. Fletcher (Ed.), Discourse comprehension. Essays in the honor of Walter Kintsch (pp. 285-305). Hillsdale, NJ: Erlbaum.

Steinmayr, R., Ziegler, M. & Traube, B. (2010). Do intelligence and sustained attention interact in predicting academic achievement? Learning and Individual Differences, 20(1), 14-18.

Suyanto, W. (1998). The effects of student teams-achievement divisions on mathematics achievement in Yogyakarta rural primary schools. Houston: University of Houston.

Tarim, K. & Akdeniz, F. (2008). The effects of cooperative learning on Turkish elementary student's mathematics achievement and attitude towards mathematics using TAI and STAD methods. Educational Studies in Mathematics, 67(1), 77-91.

Tolar, T. D., Fuchs, L., Cirino, P. T., Fuchs, D., Hamlett, C. L. & Fletcher, J. M. (2012). Predicting development of methematical word problem solving across the intermediate grades. Journal of Educational Psychology, 4(4), 1083-1093.

Traub, S. (2004). Unterricht kooperativ gestalten. Bad Heilbrunn: Klinkhardt.

Tymms, P., Merrell, C. & Henderson, B. (1997). The first year at school: A quantitative investigation of the attainment and progress of pupils. Educational Research and Evaluation, 3(2), 101-118.

Veenmann, S., Kenter, B. & Post, K. (2000). Cooperative learning in Dutch primary classrooms. Eudcational Studies, 26(3), 281-302.

Vista, A. (2013). The role of reading comprehension in maths achievement growth: Investigating the magnitude and mechanism of the mediating effect on maths achievement in Australian classrooms. International Journal of Educational Research, 62(1), 21-35.

Vukovic, R. K. (2012). Mathematics difficulty with and without reading difficulty: Findings and implications from a four-year longitudinal study. Exceptional Children, 78(3), 280-300.

Wahl, D. (2004). Die Gruppenrallye. In A. A. Huber (Hrsg.), Kooperatives Lernen - kein Problem (S. 86-95). Leipzig: Klett.

Wai, J., Cacchio, M., Putallaz, M. & Makel, C. (2010). Sex differences in the right tail of cognitive abilities: A 30 years examination. Intelligence, 38(4), 412-423.

Weber, H. S., Lu, L., Shi, J. & Spinath, F. M. (2013). The roles of cognitive and motivational predictors in explaining school achievement in elementary school. Learning and Individual Differences, 25(1), 85-92.

Weiss, R. H. (1998). Grundintelligenztest Skala 2: CFT 20. Braunschweig: Westermann.

Wellenreuther, M. (2014). Lehren und Lernen – aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen im Unterricht. Hohengehren: Schneider Verlag.

Zimowski, M. F., Muraki, E., Mislevy, R. J. & Bock, R. D. (2003). BILOG-MG 3 for Windows: Multiple-group IRT analysis and test maintenance for binary items [Computer software] Skokie, IL: Scientific Software International, Inc.

A quasi-experiment examining the Student Team-Achievement Division (STAD): Does the method accomplish what it promises in mathematics education?

The present quasi-experimental intervention study examined effects of a Student Team-Achievement Division (STAD) in mathematics. The sample consisted of 497 5th graders from the German-speaking part of the Swiss canton of Fribourg. 15 classes were randomly assigned either to the intervention or the control group. Multilevel analyses provided the following main results: when controlling for gender, social origin, intelligence and reading comprehension at the individual level, the students who were taught in a STAD-class obtained somewhat higher performance scores than the control group. The control variable with the greatest influence was gender. Girls tended to have lower scores. No interaction effects were found between the control variables and the intervention, thus all students seemed to benefit to a similar extent from the intervention. The effect sizes were rather moderate, so that further inquiries would be justified. This would help with a differentiated assessment of the published didactic recommendations in regards to using the Student Team-Achievement Division for mathematics.

Keywords: cooperative learning – Mathematics performance – STAD – student-team-achievement division

Autoren

Prof. Dr. Christian Wandeler, California State University, Department of Curriculum and Instruction), Fresno,
Prof. Dr. Alois Niggli,
Dr. Caroline Villiger,
Marcel Aebischer, Praxisdozent PH,
Philippe Leopold, M. A., Pädagogische Hochschule Freiburg, Freiburg (Schweiz).
Korrespondenz an: nigglia@edufr.ch