

Originalarbeit

Rechtschreibtraining in schulischen Fördergruppen – Ergebnisse einer Evaluationsstudie in der Primarstufe

G. Schulte-Körne, W. Deimel und H. Remschmidt

Aus der Klinik für Kinder- und Jugendpsychiatrie und -psychotherapie der Philipps-Universität Marburg

Zusammenfassung: *Fragestellung:* Ziel der vorliegenden Studie ist die schulische Evaluation eines Förderprogramms (Marburger Rechtschreibtraining), welches sich bereits in der außerschulischen Förderung als wirksam erwiesen hat. Als Kontrollbedingung fungierte das bereits in der Schule vorhandene Förderprogramm.

Methode: Insgesamt 37 Kinder der 2. und 3. Klassen wurden von den Lehrerinnen als rechtschreibschwach benannt und wurden in die Studie eingeschlossen. Die Kinder wurden in Kleingruppen zusätzlich zum regulären Unterricht zweimal wöchentlich eine Schulstunde lang gefördert.

Ergebnisse: Nach zwei Jahren verbesserten sich die Förderkinder unabhängig von der Methode signifikant in ihrer Lese- und Rechtschreibfähigkeit. Dieser Effekt zeigte sich auch im Lehrerurteil und nach Einschätzung durch die Eltern. Hingegen zeigte sich keine Veränderung der schulbezogenen emotionalen Befindlichkeit. Auch die nicht geförderte Vergleichsgruppe konnte sich verbessern.

Schlussfolgerung: Die Förderung rechtschreibschwacher Kinder in schulischen Fördergruppen ist effektiv, wenngleich die Tatsache, dass sich auch die Kinder der Vergleichsgruppe verbessern konnten, Fragen aufwirft. Die Auswahl der Methode hatte in unserer Studie keinen Einfluss auf das Ergebnis. Eine Reihe von potenziell wirksamen Einflüssen (z.B. Stichprobenselektion) konnte in unserer Studie nicht systematisch erfasst werden und sollte in weiteren Arbeiten untersucht werden.

Schlüsselwörter: Lese-Rechtschreibstörung, Förderunterricht, Evaluation

Summary: *Spelling training in school based intervention groups – Results of an evaluation trial in secondary school*

Objective: This study aims to evaluate in a school-based tutoring setting a training program for spelling (Marburger Rechtschreibtraining) that has proven effective in non-school settings. A spelling training program already in use at the school serves as the control condition.

Methods: A total of 37 second- and third-graders rated by their teachers as spelling disabled participated in the study. In addition to their regular lessons, the children received two added lessons in small tutoring groups each week.

Results: The skills of the children in the tutoring program had increased significantly two years later regardless of the method used. This effect was confirmed both by tests as well as by teachers' and parental reports. However, the children's emotional attitudes towards school failed to change significantly. The control group that had received no tutoring improved as well.

Conclusions: Tutoring spelling disabled children in small groups is an effective method for improving their reading and spelling abilities. Nevertheless, the fact that the skills of children in the control group without any tutoring also improved raises a number of questions. The choice of method in our study had no effect on the outcome. Our study was unable to systematically evaluate a number of potential influences (such as sample selection); these should be investigated further.

Key words: Dyslexia, therapy, evaluation

1. Einleitung

Im Beschluss der Kultusministerkonferenz vom 20. April 1978 sind für Schüler, die die Ziele des Lese- und Rechtschreibunterrichts der Jahrgangsstufe 2 noch nicht erreicht haben, sowie für Schüler der Jahrgangsstufe 3 und 4, deren Leistungen im Lesen und/oder Rechtschreiben über einen Zeitraum von mindestens drei Monaten schlechter als ausreichend bewertet werden, besondere Fördermaßnahmen vorgesehen. Die Umsetzung dieser Fördermaßnahmen kann klassenintern oder klassenübergreifend in Fördergruppen erfolgen. Die Förderung sollte ab der 2. Klasse im Umfang von maximal zwei Wochenstunden in Gruppen von vier bis acht Schülern erfolgen. Es wird davon ausgegangen, «dass durch die Förderung in den Jahrgangsstufen 1 bis 6 Schwierigkeiten im Lesen und Rechtschreiben im Wesentlichen behoben sind».

Leider liegen nur sehr wenige Therapiestudien mit deutschsprachigen¹ lese-rechtschreibgestörten Kindern vor, in denen eine Evaluation der zahlreich vorhandenen Therapiekonzepte (Übersicht bei Mannhaupt, 1994 und Roth & Warnke, 2001) vorgenommen wird. Insbesondere liegen kaum Studien vor, die schulische Förderung evaluieren.

- 1 Da es Hinweise darauf gibt, dass die jeweilige Muttersprache einen moderierenden Einfluss auf den Schriftspracherwerb hat (Landerl, Wimmer & Frith, 1997), können Ergebnisse aus z.B. englischsprachigen Ländern nicht ohne weiteres auf deutschsprachige Länder übertragen werden. Ferner sind die nationalen Schulsysteme zum Teil so unterschiedlich, dass Ergebnisse der innerschulischen Förderung sinnvollerweise nicht ohne Einschränkungen auf das deutsche Schulsystem übertragen werden können. Deshalb werden im Folgenden ausschließlich Arbeiten mit deutschsprachigen Kindern zitiert.

Von pädagogischer Seite wird oft vertreten, dass Lese-Rechtschreibschwierigkeiten «in erster Linie [...] ein pädagogisches Problem» darstellen (Naegele & Valtin, 2001, S. 10). Daraus folgt, dass der Förderunterricht meistens lediglich eine Intensivierung des normalen Deutschunterrichts darstellt. Diese Annahme wird durch Berichte über schulische Förderprojekte bestätigt (Hassmann-Kube, 1986; Schäffer, 1986), bei denen eher organisatorische Überlegungen (z.B. Förderunterricht vs. Förderklassen) und weniger die konkreten Inhalte im Vordergrund stehen. Ergebnisse dieser Förderungen werden in der Regel nicht mitgeteilt. Bastian et al. (1988) teilen zwar Ergebnisse von Schulleistungstests hinsichtlich des Schreibens und Lesens mit, eine statistische Auswertung findet aber nicht statt. Immerhin lässt sich anhand der mitgeteilten Werte ablesen, dass die geförderten Klassen im Mittel besser sind. Ob dieser Effekt signifikant ist, bleibt offen. May (2001) beschreibt sogenannte «lernförderliche Bedingungen» des schulischen Förderunterrichts. Erfolgreiche Klassen zeichnen sich demnach durch ein hohes Ausmaß an direkter Zuwendung durch die Lehrerin aus. Weitere förderliche Bedingungen nach May (2001) sind direktives Lehrerverhalten und eine erhöhte individuelle Zuwendung an die Förderkinder auf Kosten der übrigen Kinder.

In einer Übersicht über österreichische Studien kommen Klicpera & Gasteiger-Klicpera (1995) zu der Auffassung, dass zum Teil positive Effekte des Förderunterrichts nachgewiesen werden konnten, allerdings die Ergebnisse deutlich hinter den Erwartungen zurückbleiben. Einen Sonderfall stellen sogenannte Förderprogramme dar (s. hierzu Klicpera & Gasteiger-Klicpera 1995), die, wie der Name bereits andeutet, den Lehrern ein standardisiertes Programm für den Unterricht zur Verfügung stellen. Aller-

dings fehlt auch für die Förderprogramme der klare Wirksamkeitsnachweis (Klicpera & Gasteiger-Klicpera, 1995). Die Autoren betonen auch, dass es einen eklatanten Mangel an wissenschaftlichen deutschsprachigen Studien über die Wirksamkeit von schulischen Fördermaßnahmen gibt. Zu dieser Einschätzung gelangt auch Mannhaupt (1994). In einer Meta-Analyse untersuchte er Evaluationsstudien, die in den davor liegenden 15 Jahren veröffentlicht wurden. Nur 12 der 30 Studien entsprachen jedoch überhaupt methodischen Mindestkriterien. In praktisch allen Therapiestudien wird das Training außerschulisch, d.h. nicht von den Lehrern, durchgeführt. Es liegen kaum Studien vor, die die Anwendbarkeit eines Trainingsprogramms im schulischen Setting untersuchen, oder schulische und außerschulische Fördermaßnahmen in ihrer Effektivität vergleichen.

Tacke et al. (1987) verglichen konventionellen Förderunterricht mit zwei Varianten eines Regeltrainings bei rechtschreibschwachen Hauptschülern der 5. Klasse. Förderunterricht und Training wurden von den Lehrern durchgeführt. Die Trainingsgruppen hatten gegenüber der Kontrollgruppe (Fördergruppe) nach einem Jahr keine signifikant verbesserte Rechtschreibung (Prozentrang Rechtschreibtest). Die Autoren stellten außerdem fest, dass die Motivation zur Mitarbeit bei den Trainingsgruppen (im Gegensatz zur Förderunterrichtsgruppe) extrem gesunken war. Dies zeigt, dass eine didaktisch und graphisch geschickte Umsetzung des Trainings von entscheidender Bedeutung ist; die eigentliche Fragestellung der Studie bleibt aufgrund dieses offensichtlichen Motivationsdefizites der Schüler aber unbeantwortet.

Büttner (1980) verglich die Resultate schulischer und privater Fördermaßnahmen bei 116 «Legasthenikern» der 2. bis 6. Klassen. 93% dieser Kinder erhielten irgendeine Art von Förderung, z.T. auch mehrere Arten gleichzeitig. 68% dieser Kinder nahmen am schulischen Förderunterricht teil. Eine entscheidende Schwäche der Studie ist, dass der Erfolg der Maßnahmen lediglich anhand Angaben der Eltern erfasst werden konnte. Demnach hatten sich 75% der Kinder mindestens leicht verbessert. Einen Einfluss der Art der Förderung auf den Erfolg fand der Autor nicht.

Reuter-Liehr (1993) verglich ihr Trainingsprogramm mit einer unbehandelten Kontrollgruppe in der Sekundarstufe. Die «Behandlung» der Kinder erfolgte in der Schule; aber nicht durch die Lehrer, sondern durch Projektmitarbeiter. Die Zuordnung zu den Lerngruppen erfolgte nicht ausschließlich nach quantitativer Rechtschreibleistung, sondern nach der Anzahl phonologischer Fehler. Die Kinder in den Lerngruppen waren daher gemessen am Rechtschreib-T-Wert nicht stark beeinträchtigt (mittlere T-Werte zu Beginn 44–48). Eine um den Regressionseffekt (durch Extremgruppenselektion) bereinigte Auswertung der Behandlungseffekte wird nicht mitgeteilt. Diese Ein-

wände beeinträchtigen die ansonsten positive Bilanz dieser Studie.

Die oben erwähnte Arbeit von Tacke et al. (1987) ist nach unserem Wissensstand die einzige, die einen Vergleich von einem Rechtschreibtraining mit Förderunterricht versucht. Leider sind die Ergebnisse wegen der beschriebenen Motivationsprobleme nicht aussagekräftig. Eine Recherche im April 2002 in den Datenbanken Medline, PsycLit und PsynDEX mit den Stichworten «spelling AND training AND school» erbrachte keine weiteren Arbeiten zu diesem Thema, das Gleiche gilt für eine pädagogische Datenbank (FIS-Bildung). Nach unserem besten Wissen müssen wir somit davon ausgehen, dass eine wissenschaftlich fundierte Studie zum Einsatz eines Rechtschreibtrainings im schulischen Förderunterricht bisher nicht veröffentlicht wurde.

Das Marburger Rechtschreibtraining (Schulte-Körne & Mathwig, 2001) wurde in zwei Untersuchungen evaluiert. Die Autoren fanden einen bedeutsamen Therapieeffekt bei Einsatz des Trainings bei Zweit-, Dritt- und Viertklässlern (Schulte-Körne et al., 1997, 1998). Die Eltern der Kinder wurden einmal im Monat an Elternabenden von zwei Therapeuten angeleitet. An diesen Abenden wurden Handlungsanleitungen gegeben und das aktuelle Lern- und Übungsmaterial besprochen. Die Eltern führten das Programm zwei- bis dreimal wöchentlich 20 Minuten lang durch. Die effektive Trainingszeit lag in der Woche durchschnittlich bei 20 bis 30 Minuten. Zusammenfassend zeigte sich nach dem ersten Trainingsjahr eine bedeutsame Leistungsverbesserung in dem geübten Regelbereich. Allerdings war ein signifikanter Transfer-Effekt auf nicht geübte Rechtschreibproblembereiche nicht festzustellen: Der mittlere T-Wert im Rechtschreibtest betrug zu Beginn der Studie 37.1 ± 5.1 und nach einem Jahr 35.9 ± 9.4 . Erst nach zwei Jahren stieg dieser Wert signifikant auf 40.6 ± 7.8 an.

Das Marburger Rechtschreibtraining wurde auch in einer Kurzzeitintervention überprüft (Schulte-Körne et al., 2001). Über einen Zeitraum von drei Monaten erhielten zehn lese-rechtschreibschwache 2. bis 4.-Klässler zweimal pro Woche jeweils eine Stunde Förderung. Im Gegensatz zu einem in einer parallelisierten Gruppe durchgeführten Phonologietraining konnten sich die Kinder der Regeltrainingsgruppe dabei in ihrer Rechtschreibung signifikant verbessern.

Eine sehr wichtige bisher noch unbeantwortete Frage ist allerdings, ob ein Regeltraining im Förderunterricht von den Lehrerinnen selbst erfolgreich angewandt werden kann. Auf der Basis einer bereits bestehenden Kooperation mit einer südhessischen Grundschule wurde in dieser Studie das Marburger Rechtschreibtraining im schulischen Förderunterricht über zwei Jahre von den Lehrerinnen angewendet. In einer Kontrollbedingung wurden

Schüler mit dem normalen Förderunterricht der Schule unterrichtet.

Es wurden folgende Hypothesen überprüft:

Hypothese 1: Die Rechtschreibleistung (Mittlerer Rechtschreib-T-Wert bezogen auf die Klassennorm) der Trainingsgruppe ist nach zwei Jahren Training im Vergleich zum Ausgangswert verbessert.

Hypothese 2: Die Rechtschreibleistung (Mittlerer Rechtschreib-T-Wert bezogen auf die Klassennorm) der Trainingsgruppe ist nach zwei Jahren Training besser als die Rechtschreibleistung der Kontrollgruppe.

2. Methode

2.1 Allgemeiner Versuchsplan und Stichprobe

Der Förderunterricht im Rahmen des Projekts begann mit dem Schuljahr 1999/2000, also im Sommer 1999. Die Lehrerinnen der Grundschule Höchst i. Odw. benannten die nach ihrer Einschätzung schlechtesten 40 Rechtschreiber der zweiten und dritten Klassen. Neben der schlechten Rechtschreibung galten dabei noch folgende Selektionsbedingungen: keine unkorrigierte Seh- oder Hörschwäche², Familie einsprachig deutsch³, IQ mindestens im Normbereich (s.u.).

Zu Beginn des Schuljahres erfolgte eine Gruppentestung, in der ein Intelligenztest (je nach Alter Culture Fair Intelligence Test CFT 1, Weiß & Osterland, 1997; Culture Fair Intelligence Test CFT 20, Weiß, 1987), ein klassenadäquater Rechtschreibtest (Weingartener Grundwortschatz-Rechtschreib-Test WRT1+, Birkel, 1995 bzw. Weingartener Grundwortschatz Rechtschreib-Test WRT2+, Birkel 1994) und ein Lesetest (Salzburger Lese-Rechtschreibtest SLRT, Landerl, Wimmer & Moser, 1997) durchgeführt wurden.

Dann erfolgte eine zufällige Aufteilung in Rechtschreibtraining (RT) und Förderunterricht (FU)⁴. Diese beiden Gruppen wurden dann hinsichtlich Geschlecht,

Rechtschreibleistung, IQ und Klassenstufe verglichen, und es wurden einige Schüler der jeweils anderen Gruppe zugeordnet, um eine möglichst gute Gleichverteilung hinsichtlich dieser vier Variablen zu erreichen. Völlig gelang dies aus schulorganisatorischen Gründen nicht⁵, aber die orientierend durchgeführten Tests auf Gruppenunterschiede (t-Tests) sind eindeutig nicht im Bereich einer Signifikanz. Tabelle 1 gibt eine Übersicht über die relevanten Variablen.

Es erfolgte dann eine Einteilung in Kleingruppen von etwa fünf Kindern pro Gruppe. Insgesamt entstanden so zehn Gruppen. Die tatsächliche Gruppengröße war etwas höher, da zusätzlich zu den im Projekt erfassten Kindern weitere Kinder an den Fördermaßnahmen teilnahmen. Es waren dies Kinder, bei denen Förderungsbedarf festgestellt worden war, die aber nicht in die Studie passten (i.d.R. weil sie Ausländer waren)⁶.

Wie zu erwarten gab es im Laufe der zwei Jahre Ausfälle. Zwei Kinder wechselten die Schule bzw. die Familien zogen weg, drei Kinder wiederholten ein Schuljahr. Am Ende der Studie bestand die RT-Gruppe aus 21 Kindern, und die FU-Gruppe aus 16 Kindern.

Es wurden zwei Interventionen miteinander verglichen, das Marburger Rechtschreibtraining und der bisherige schulische Förderunterricht. In der Kontrollgruppe wurde das in der Schule bestehende Förderprogramm wie bisher verwendet. Dieses stützte sich auf zwei Computerprogramme: «Budenberg Wortdiktat» (<http://www.budenberg.de>) und «Uniwort» (Traeger-Verlag, <http://www.etverlag.de>). Alle Gruppen starteten mit dem Budenberg-Programm, und gingen dann zu «Uniwort» über. Im

Tabelle 1: Geschlecht, Klassenstufe, Rechtschreib- und Leseleistung (T-Werte) sowie IQ für RT (Rechtschreibtraining) und FU (Förderunterricht)

	RT	FU	p-Wert
% Jungen	55	55	0,98
% Erstklässler	55	45	0,54
Rechtschreibtest (MW ± SA)	40,1 ± 5,4	41,5 ± 6,0	0,42
Lesen (Zeit HW), (MW ± SA)	44,7 ± 10,8	42,0 ± 10,4	0,41
IQ (MW ± SA)	99,3 ± 9,5	101,7 ± 9,2	0,41
n	22	20	

2 Nach dem Eindruck der Lehrer entschieden, bei Auffälligkeiten Veranlassung zur fachärztlichen Untersuchung.

3 Nach dem Eindruck und durch Elterngespräch mit der Klassenlehrerin entschieden.

4 Im Folgenden werden diese Begriffe und Abkürzungen verwendet:

Projektgruppe (PG): alle geförderten Kinder, n = 37

Rechtschreibtrainingsgruppe (RT): n = 21

Förderunterrichtsgruppe (FU): n = 16

Vergleichsgruppe (VG): nicht geförderte Kinder, Rest der Jahrgänge, n = 98 (jeweils 49 aus den dritten und 49 aus den vierten Klassen).

5 Die Kinder sind nicht beliebig den Gruppen zuteilbar, da der Förderunterricht mit dem normalen Stundenplan koordiniert werden muss. Praktisch bedeutet dies, dass eine Fördergruppe meistens aus Kindern aus einer oder zwei Klassen besteht.

6 Das war eine Konzession an die realen Bedingungen in einer Schule. Es war ethisch nicht vertretbar, Kinder mit Förderbedarf nicht zu fördern, und wegen der Stundenpläne konnten die Nicht-Studienkinder nur auf die einzelnen Gruppen verteilt werden.

Förderunterricht wurden keinerlei Rechtschreibregeln vermittelt. Zwar waren die verwendeten Wortlisten thematisch jeweils einem Schwerpunkt gewidmet, jedoch fehlte die Abgrenzung zu alternativen Schreibweisen. Zum Beispiel wurden Wortlisten zur Mitlautverdopplung verwendet, diese enthielten jedoch ausschließlich Wörter, bei denen eine Mitlautverdopplung vorkommt. Informationen darüber, wann eine Mitlautverdopplung *nicht* vorkommt, konnten die Kinder daraus nicht entnehmen. In diesem Sinne kann auch nicht von «implizitem Regellernen» gesprochen werden.

Die Kinder in der RT-Gruppe wurden nach dem «Marburger Rechtschreibtraining» (Schulte-Körne & Mathwig, 2001) unterrichtet.

Das absolvierte Pensum lässt sich bei den PC-Programmen, also für die FU-Gruppen, kaum quantifizieren. Die RT-Gruppen hatten nach den zwei Jahren meistens Kapitel 10 erreicht; eine Gruppe war komplett mit allen 12 Kapiteln fertig geworden. Es kann also geschlossen werden, dass das Marburger Rechtschreibtraining in zwei Schuljahren knapp bewältigt werden kann.

2.2 Konkrete Durchführung und Messdesign

RT und FU wurden von zwei Lehrerinnen durchgeführt. Beide Lehrerinnen hatten etwa den gleichen Anteil RT und FU. Aus schulorganisatorischen Gründen kam es dabei zu gewissen Veränderungen (im Laufe des Projektes kam eine dritte Lehrerin hinzu, die Gruppenzusammensetzungen innerhalb der RT- und FU-Gruppen änderte sich).

Vom Umfang her waren zwei Wochenstunden Förderung vorgesehen. Innerhalb der Stunden wurde grund-

schultypisch zunächst etwas gespielt, dann fand etwa 20 Minuten Förderung statt, und dann wurde wieder etwas gespielt. Die Inhalte der Förderung und auch der Umfang des tatsächlich stattfindenden Unterrichts wurden von den Lehrerinnen auf speziellen Erfassungsbögen festgehalten. Dort wurde auch notiert, welche Kinder z.B. krankheitsbedingt an bestimmten Tagen fehlten.

Beide Projektgruppen bestanden zu vergleichbaren Teilen aus Kindern von zwei verschiedenen Klassenstufen. Bei Studienbeginn war etwa die Hälfte der Kinder in der zweiten Klasse und die andere Hälfte in der dritten Klasse.

Zusätzlich zur Erfassung der Lese- und Schreibleistungen sollten Lehrer, Eltern und Kinder Erfolg und Zufriedenheit mit dem Training einschätzen. Dazu wurde ein Fragebogen verwendet, der jeweils am Ende eines Schuljahres, und auch zu Studienbeginn ausgefüllt wurde. Darin waren auch Fragen zur emotionalen Situation und zur Selbsteinschätzung enthalten.

Bei der Entwicklung des Fragebogens wurde auf den im Marburger Eltern-Kind-Training verwendeten Fragebogen zurückgegriffen (Schulte-Körne et al., 1997, 1998), der für die schulische Anwendung modifiziert wurde.

Der Kinderfragebogen enthält Fragen zur Einschätzung der Leistung im Schreiben, Lesen und Rechnen, eine Einschätzung darüber, wie gern das Kind schreibt, liest und rechnet, sowie Fragen zur emotionalen Situation, z.B. «Ich traue mich, etwas an die Tafel zu schreiben.» Die Elternversion ist weitgehend der Kinderversion vergleichbar, allerdings gibt es weniger und andere Items zur emotionalen Situation (z.B. ob das Kind gern zur Schule geht). Von den Lehrern wurden Leistungseinschätzungen auf der Notenskala erfragt, und zwar für die Bereiche Recht-

Table 2: Erhobene Daten in den Projektgruppen

Zeitpunkt		0 8/99	1 3/00	2 6/00	3 1/01	4 6/01
Jüngere Teilgruppe:	Klassenstufe	1	2	2	3	3
	Testverfahren Rechtschreibung	WRT1+ ¹	WRT2+ ²	WRT2+	WRT2+	DRT3 ³
Ältere Teilgruppe:	Klassenstufe	2	3	3	4	4
	Testverfahren Rechtschreibung	WRT2+	WRT3+ ⁴	WRT3+	DRT4 ⁵	GRT4+ ⁶
Rechtschreibtest		x	x	x	x	x
Lesetest		x	x	x	x	x
Fragebogen Eltern		x		x		x
Fragebogen Kinder		x		x		x
Fragebogen Lehrer		x		x		x

¹ Birkel 1995

² Birkel 1994

³ Müller 1997. Für einen Teil der Stichprobe wurde hier auch der WRT3+ durchgeführt, Vergleich s.u.

⁴ Birkel 1994a

⁵ Grund et al. 1994

⁶ Birkel 1990

Tabelle 3: Erhobene Daten in der Vergleichsgruppe

Zeitpunkt	0 8/99	1 3/00	2 6/00	3 1/01	4 6/01
Klassenstufe der jüngeren Teilgruppe	1	2	2	3	3
Klassenstufe der älteren Teilgruppe	2	3	3	4	4
Rechtschreibtest			x		x
Fragebogen Eltern	x		x		x
Fragebogen Kinder	x		x		x
Fragebogen Lehrer	x		x		x

schreibung, Lesen, Rechnen und «allgemeine kognitive Leistungsfähigkeit».

Außerdem wurden auch die «Vergleichskinder» (Vergleichsgruppe = jeweils der Rest des Jahrgangs, VG) untersucht. Die Rechtschreibleistung konnte bei dieser Gruppe am Ende des ersten und am Ende des zweiten Jahres erfasst werden. Zum Ausgangszeitpunkt war eine Erfassung aus organisatorischen Gründen leider nicht möglich. In der Vergleichsgruppe wurden – wie in den Projektgruppen – nur einsprachig deutsche Kinder berücksichtigt. Integrationskinder wurden ebenfalls ausgeschlossen. Idealerweise sollte die Vergleichsgruppe damit die nicht-rechtschreibschwachen deutschen Kinder der Jahrgänge enthalten.

Die Tabellen 2 und 3 geben eine Übersicht über die erhobenen Daten. Es gab insgesamt 5 Messzeitpunkte. «0», «2» und «4» sind die Hauptmessungen, jeweils am Schuljahresende, und zwar zu Beginn, während, und am Ende der Studie⁷. «1» und «3» sind Zwischenmessungen mit reduziertem Instrumentarium zum Halbjahreswechsel. Zugunsten einer einfachen Darstellung wird im Folgenden auf diese Nomenklatur Bezug genommen.

7 Die Erstmessung fand am Ende des Schuljahres vor Projektbeginn statt.

3. Ergebnisse

3.1 Zur Selektion der Projektgruppen

Es wurden mit Absicht nicht die Rechtschreibtests zur Selektion der Förderkinder verwendet. Dies geschah aus zwei Gründen. Zum einen entspricht dies der normalen Situation in einer Schule, zum anderen verringert dies den Regressionseffekt (Regression zur Mitte), da das Selektionskriterium nicht identisch ist mit dem später verwendeten Maß zur Verlaufsmessung.

Vom Zeitpunkt 0 (Projektbeginn) liegen leider keine Rechtschreibtests der Vergleichsgruppe vor, daher soll hier eine vergleichende Betrachtung zum Zeitpunkt 2 (nach einem Jahr) erfolgen. Tabelle 4 zeigt die Verteilungen von Projekt- und Vergleichsgruppe.

Es fällt auf, dass VG und PG den Schwerpunkt im unteren Durchschnittsbereich haben.

Es verwundert, dass die VG nicht besser ist, da die schlechtesten Rechtschreiber ja nicht mehr dabei sein sollten⁸. De facto weist die Mehrzahl der geförderten Schüler eine Rechtschreibleistung im Normbereich auf, und gleichzeitig kam die Mehrzahl der unterdurchschnittlich schreibenden Schüler nicht in die Förderung.

Nach Rücksprache mit der Schule scheint dieses Ergebnis folgendermaßen erklärt werden zu können: Kinder mit Schwächen ausschließlich im Rechtschreiben wurden eher als entwicklungsverzögert eingeschätzt («das wächst sich aus»). Diese Kinder wurden nicht für die Förderung vorgeschlagen. Kinder, die eher allgemein nicht mitkamen (= eher nicht teilleistungsgestört), wurden dagegen als Förderkinder vorgeschlagen, da ihre Defizite nicht im normalen Unterricht behebbar waren. Außerdem gibt es Kinder, bei denen die Lehrer äußere Umstände für die schwache Rechtschreibung verantwortlich machten (schlechtes Umfeld, nicht engagierte Eltern, generell häufiges Fehlen). Auch bei diesen wäre nach der Lehrereinschätzung Förderunterricht nicht indiziert.

Es bleibt also festzuhalten, dass die tatsächliche Selektion nicht dem entspricht, was vom wissenschaftlichen Standpunkt her eigentlich hätte erfolgen sollen. Die Lehrerinnen waren angehalten worden, nur nach der einge-

8 Wie später dargestellt wird, hatten sich die Projektkinder zu diesem Zeitpunkt noch nicht bedeutsam verbessert, so dass die Gruppen sich eigentlich deutlicher unterscheiden sollten.

Tabelle 4: Anzahl Kinder mit Rechtschreibwerten in den einzelnen T-Wert-Bändern

T-Wert RST zum Zeitpunkt 2	≤ 30	31–40	41–50	51–60	61–70	Gesamt
Projektgruppe	3	9	20	5	0	37
Vergleichsgruppe	1	14	45	35	2	98

schätzten Rechtschreibung zu selegieren. Tatsächlich gingen aber Annahmen in die Selektionsentscheidung mit ein, die sich kaum überprüfen lassen. Die Rechtschreibleistung selbst war dabei nur ein Teil der Entscheidung. Man gewinnt den Eindruck, dass Kinder mit einer Teilleistungsstörung, d.h. spezifischen Schwierigkeiten im Schreiben und/oder Lesen, eher *nicht* in die Förderung kamen.

Diese Effekte waren erst im Nachhinein erkennbar.

3.2 Vergleich Projektgruppe – Vergleichsgruppe

Von der Vergleichsgruppe liegen Rechtschreibtests nur von Ende des ersten und zweiten Jahres vor (Zeitpunkte 2 und 4). t-Tests zwischen den Gruppen (Rechtschreibtests, VG vs. PG) sind erwartungsgemäß jeweils höchst signifikant ($p < 0,001$). Tabelle 5 gibt einen Überblick über den Verlauf im Schreiben und Lesen.

Es fällt auf, dass von Zeitpunkt 2 zu 4 auch die Werte der Vergleichsgruppe ansteigen, und zwar noch kräftiger als die der Projektgruppe. Dies ist möglicherweise auf die durchgeführten Tests zurückzuführen, die eventuell nicht gleich schwierig sind.

Bei einem Teil der Kinder aus der jüngeren Klasse wurde zum Projektende nicht nur der DRT3, sondern auch der WRT3+ durchgeführt (s. Tab. 6).

Diese Zahlen sprechen dafür, dass WRT2+, WRT3+ und GRT4+ etwa das gleiche Schwierigkeitsniveau haben, aber der DRT3 deutlich leichter ist. Demnach (unter Weglassung der DRT3-Daten) wäre dann die jüngere Gruppe von vornherein schlechter gewesen, und beide

Tabelle 6: Vergleich der Rechtschreibtests in der Vergleichsgruppe

	Zeitpunkt 2	Zeitpunkt 4
Jüngere	45,0 ± 8,0 (WRT2+)	49,1 ± 7,5 (WRT3+) 53,7 ± 7,4 (DRT3)
Ältere	50,3 ± 6,6 (WRT3+)	54,2 ± 7,5 (GRT4+)

Gruppen hätten sich in dem einen Jahr um jeweils ca. 4 T-Punkte verbessert.

3.3 Vergleich der zwei Fördergruppen

3.3.1 Testdaten

Die Tabellen 7 und 8 zeigen die Entwicklung der Rechtschreib- und Lesetestwerte für die beiden Fördergruppen (RT und FU), sowie die Effektstärken des Vergleichs Ausgangsmessung – Endmessung.

Tabelle 9 zeigt die p-Werte der Signifikanztests auf Veränderung für Rechtschreib- und Lesetestwerte. Dafür wurden jeweils die Differenzen zwischen letzter und erster Messung gebildet, und diese dann gegen Null getestet. Die p-Werte in Tabelle 9 sind einseitig und nicht alpha-korri-

Tabelle 9: p-Werte der Signifikanztests auf Veränderung für Rechtschreib- und Lesetests

	RT	FU
Rechtschreibtest	0,005	0,012
Lesen	0,020	0,011

Tabelle 5: T-Werte für Rechtschreibtests und Lesen (PG), sowie Rechtschreibtests (VG)

Zeitpunkt	0	1	2	3	4
PG Lesen	43,1 ± 11,3	46,3 ± 9,3	46,2 ± 10,1	48,8 ± 10,9	48,5 ± 10,9
PG RST	40,8 ± 6,0	35,5 ± 5,9	42,2 ± 7,9	43,4 ± 6,7	44,6 ± 7,0
VG RST	–	–	47,4 ± 7,8	–	54,0 ± 7,4

Tabelle 7: Rechtschreib-T-Werte (MW +STA) und Effektstärken (Ende – Anfang) für die beiden Fördergruppen

Zeitpunkt	0	1	2	3	4	Effektstärke (Zeitp. 4–0)
RT (n = 21)	39,8 ± 5,4	34,5 ± 5,2	40,8 ± 8,1	42,5 ± 7,0	43,6 ± 7,4	0,70
FU (n = 16)	42,1 ± 6,5	36,9 ± 6,7	44,1 ± 7,5	44,6 ± 6,3	45,8 ± 6,4	0,57

Tabelle 8: Lesetest-T-Werte (MW +STA) und Effektstärken (Ende – Anfang), Untertest *Häufige Wörter*

Zeitpunkt	0	1	2	3	4	Effektstärke (Zeitp. 4–0)
RT (n = 21)	42,6 ± 12,1	46,2 ± 9,8	46,2 ± 11,4	49,3 ± 11,7	47,7 ± 12,3	0,42
FU (n = 16)	43,7 ± 10,6	46,6 ± 8,8	46,1 ± 8,5	48,2 ± 10,2	49,7 ± 8,8	0,57

giert. Wenn man eine Alpha-Korrektur für alle vier durchgeführten Tests anwendet, so bleiben die kursiv gesetzten Werte unterhalb des Grenzwertes von 5%.

Bis auf das Lesen in der RT-Gruppe bleiben also alle Werte signifikant. Die Tatsache, dass trotz praktisch punktgleicher Veränderung und im Mittel vergleichbarer Streuung die p-Werte der Rechtschreibtests für RT und FU unterschiedlich ausfallen, reflektiert wohl die unterschiedlichen Gruppengrößen (21 vs. 16).

In Abbildung 1 sind die Daten für die Rechtschreibung zusätzlich nach Klassenstufen aufgeteilt. Für alle Untergruppen zeigt sich der charakteristische V-förmige Verlauf (Schulte-Körne et al., 1998), d.h. bei der ersten Wiederholungsmessung haben sich die Kinder im Mittel verschlechtert. Danach steigen die Werte einigermaßen kontinuierlich bei jeder Messung an.

Für das Lesen gibt es keinen Abfall der Werte bei der ersten Wiederholungsmessung. Hier steigen die Werte von Messzeitpunkt zu Messzeitpunkt leicht an.

3.3.2 Eltern- und Kinderurteil

Bei den folgenden tabellarischen Darstellungen von Fragebogendaten wurden die Items jeweils dichotomisiert, da die Verteilungen dies nahe legten. Es wurde jeweils die positivste Kategorie beibehalten, und die restlichen Kategorien wurden zu einer zusammengefasst. Tabelle 10 zeigt zunächst die Daten der Eltern.

In der RT-Gruppe gibt es am Ende deutlich mehr mit als «gut» eingeschätzter Rechtschreibung, während der Wert in der FU-Gruppe unverändert bleibt. Beim Lesen verbessern sich beide Gruppen, die RT-Gruppe allerdings stärker (wobei hier aber auch der Ausgangswert niedriger ist). Beim Rechnen ist es umgekehrt, hier gibt es eine leichte Verbesserung bei der FU-Gruppe und eine leichte Verschlechterung bei der RT-Gruppe.

Relativ zur Vergleichsgruppe sind sowohl RT als auch FU im Rechtschreiben und Lesen deutlich schlechter.

Tabelle 11 zeigt die korrespondierenden Werte der Kinder.

Hier gibt es durchgängig Verschlechterungen. Möglicherweise sind die Kinder nicht in der Lage, ihre Leistungen realistisch einzuschätzen. Trotz dieses generellen Trends kann man festhalten, dass im Durchschnitt zu beiden Testzeitpunkten RT und FU sich hinsichtlich aller drei Leistungen schlechter als die Vergleichsgruppe einschätzen.

3.3.3 Lehrerurteil

Von den Lehrern gibt es die Bewertung von Rechtschreibung, Lesen, Rechnen und «allgemeine kognitive Leistungsfähigkeit» (= Intelligenz) auf der Notenskala (Tabelle 12).

Es zeigen sich in beiden Gruppen leichte Verbesserungen im Rechtschreiben und Lesen, wobei diese Verbes-

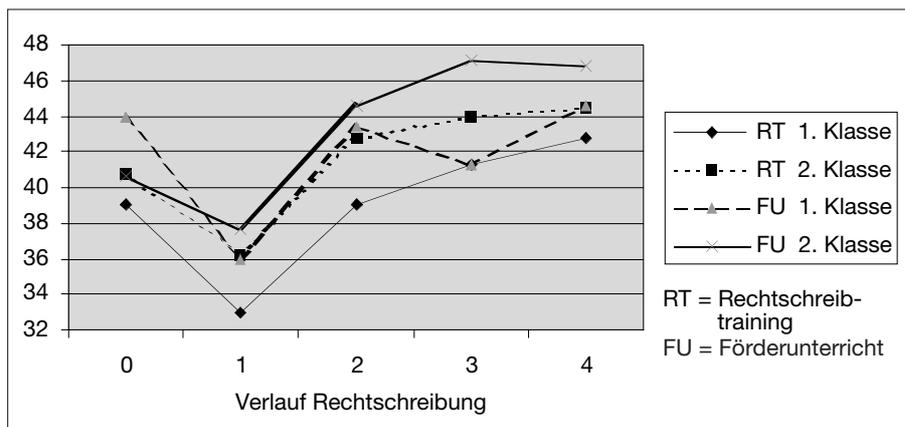


Abbildung 1: Verlauf der Rechtschreibleistung (T-Werte) getrennt für die beiden Projektgruppen nach Klassenstufe. Auf der Abszisse sind die Testzeitpunkte abgetragen.

Tabelle 10: Fragebogendaten der Eltern, jeweils % «gut» (RT = Rechtschreibtraining; FU = Förderunterricht; VG = Vergleichsgruppe)

	Zeitpunkt 0			Zeitpunkt 4		
	RT	FU	VG	RT	FU	VG
Rechtschreiben des Kindes	20	25	75	30	25	74
Lesen des Kindes	35	50	84	67	69	84
Rechnen des Kindes	75	62	83	71	75	89

Tabelle 11: Fragebogendaten der Kinder, jeweils % «gut»

	Zeitpunkt 0			Zeitpunkt 4		
	RT	FU	VG	RT	FU	VG
Rechtschreiben	35	38	67	24	25	54
Lesen	70	63	89	57	44	66
Rechnen	60	56	90	38	50	65

Tabelle 12: Lehrereinschätzung der Leistungen auf der Notenskala

	Zeitpunkt 0			Zeitpunkt 4		
	RT	FU	VG	RT	FU	VG
Rechtschreiben	3,9 ± 1,0	3,9 ± 0,8	2,3 ± 0,8	3,7 ± 1,2	3,2 ± 0,8	2,2 ± 0,9
Lesen	3,3 ± 1,0	3,1 ± 0,6	2,0 ± 0,7	2,9 ± 0,6	2,6 ± 0,8	2,0 ± 0,7
Rechnen	2,9 ± 0,7	3,1 ± 0,7	2,1 ± 0,8	3,0 ± 0,7	2,8 ± 0,5	2,2 ± 0,7
Intelligenz	3,0 ± 0,7	3,2 ± 0,5	2,2 ± 0,8	3,0 ± 0,5	2,6 ± 0,6	2,2 ± 0,8

serungen in der FU-Gruppe etwas höher ausfallen. Im Rechnen und in der eingeschätzten Intelligenz gibt es erwartungsgemäß kaum Veränderungen, bis auf eine erhöhte Intelligenzeinschätzung in der FU-Gruppe am Ende der Studie.

Die Werte der Vergleichsgruppe bleiben überraschend stabil; zwischen den beiden Zeitpunkten liegen immerhin zwei Jahre. Der Abstand der Fördergruppen zur Vergleichsgruppe bleibt trotz der leichten Verbesserung erheblich.

3.3.4 Signifikanztests

Die eingeschätzten Leistungen im Rechtschreiben und Lesen (von Eltern, Kindern und Lehrern) wurden anschließend auf signifikante Veränderung getestet. Dafür wurden jeweils die Differenzen zwischen letzter und erster Messung gebildet, und diese dann gegen Null getestet. Die p-Werte in Tabelle 13 sind einseitig und nicht alpha-korrigiert. Wenn man eine Alpha-Korrektur für alle 12 durchgeführten Tests anwendet, so bleiben nur die kursiv gesetzten Werte unterhalb des Grenzwertes von 5%. Eine signifikante Verbesserung kann damit nur in der FU-Gruppe im Lehrerurteil beim Lesen und Schreiben, sowie für die RT-Gruppe im Elternurteil beim Lesen festgestellt werden.

Tabelle 13: Signifikanztests Fragebögen

	Rechtschreiben		Lesen	
	RT	FU	RT	FU
Eltern	0,0195	n.s.	0,001	n.s.
Kinder	n.s.	n.s.	n.s.	n.s.
Lehrer	n.s.	0,0005	0,038	0,002

3.4 Vorhersage bzw. Korrelate der Verbesserung

Mit den *Fragebogendaten* (Eltern und Kinder) wurden orientierende Mittelwertvergleiche durchgeführt. Überzufällige (also statistisch bedeutsame) Zusammenhänge zwischen den Fragebogendaten und dem Trainingserfolg gab es nicht.

Es kann daher gefolgert werden, dass eine Vorhersage des Trainingserfolges aufgrund der Fragebogeneinschätzungen von Eltern und Kindern nicht möglich ist.

Die *Unterrichtsquote* der Kinder wurde durch Lehrerprotokolle erfasst. Über den gesamten Projektzeitraum hinweg wurde die prozentuale Stundenzahl des Förderunterrichts berechnet. Dabei wurde der tatsächlich pro Kind durchgeführte Unterricht in Beziehung gesetzt zu der maximal möglichen Stundenzahl (pro Jahr: 40 Unterrichtswochen * 2 Wochenstunden). Nicht stattgefundener Unterricht kann eine Reihe von Ursachen haben: Wochen vor oder nach den Ferien, Lehrerin krank oder auf Weiterbildung, Projektwochen, Schulfeste und schließlich auch Krankheit der Kinder oder sonstige Gründe der individuellen Nichtteilnahme.

In beiden Gruppen gibt es eine durchschnittliche Unterrichtsquote von rund 75% (RT: 74,6 ± 5,8; FU: 76,2 ± 3,6). Das bedeutet, dass etwa drei Viertel des theoretisch möglichen Förderunterrichts auch tatsächlich stattfanden.

Überraschend hoch für ein so unspezifisches Maß ist die Korrelation mit dem Erfolg in der Rechtschreibung (Differenz Endmessung minus Anfangsmessung): $r = 0,47$.

Es wurde schließlich versucht, den Trainingserfolg aufgrund der «harten Daten» (Unterrichtsquote und Tests) vorherzusagen. Als Erfolgsmaß wurde wieder die Differenz zwischen End- und Anfangsmessung im Rechtschreibtest verwendet

Tabelle 14: Korrelationen mit RST-Differenz (Projektgruppen)

IQ	Unterrichtsquote	RST	Lesen
0,05	0,47	-0,32	0,27

Tabelle 14 zeigt die Korrelationen mit den Ausgangsmessungen im Lesen und Rechtschreiben, Unterrichtsquote und Intelligenz.

Es zeigt sich, dass die Unterrichtsquote deutlich am stärksten mit dem Trainingserfolg korreliert. Die Ausgangsleistungen im Rechtschreiben und Lesen sind von geringerer Bedeutung. Der IQ ist praktisch unbedeutend.

Zur Vorhersage des Trainingserfolges wurde eine lineare Regression berechnet. Es wurde ein schrittweises Vorgehen verwendet (s. Tab. 15). Der IQ spielt erwartungsgemäß keine Rolle für den Erfolg. Die aufgeklärten Varianzen (Tab. 15) spiegeln etwa die Höhe der einzelnen Korrelationen (Tab. 14) wider.

3.5 Emotionale Entwicklung

Die Eltern wurden gefragt, wie gern die Kinder schreiben und lesen, wie gern sie zur Schule gehen und ob sie Schulangst haben. Letzteres Item wird umgekehrt gepolt dargestellt, so dass hohe Werte bei allen Items positiv zu werten sind.

Die Items wurden wieder dichotomisiert, d.h. es wird die positivste Kategorie vs. alle anderen Kategorien dargestellt.

Bei den Elterndaten (Tab. 16) gibt es für Schreiben und Lesen am Anfang deutliche Unterschiede zwischen Projekt- und Vergleichskindern. Am Ende hat sich dies für das Schreiben nicht verändert, aber beim Lesen haben die Projektkinder aufgeholt, d.h. sie lesen nach Auskunft der Eltern genau so gern wie die Vergleichskinder. Der Teil der Kinder, die gern zur Schule gehen, bleibt stabil und liegt für die Projektkinder unterhalb des Anteils der Vergleichsgruppe. Bezüglich der Schulangst gibt es eine negative Entwicklung, d.h. der Anteil der «schulangstfreien» Kinder geht in der PG zurück und liegt am Ende unterhalb der VG.

Auch die entsprechenden Variablen des Kinderfragebogens wurden dichotomisiert. Es wurde gefragt, wie gern die Kinder schreiben und lesen, ob sie sich trauen, etwas

Tabelle 15: Aufgeklärte Varianzen (R^2) der Regressionen des Erfolgs im Rechtschreiben auf IQ, Unterrichtsquote, Ausgangsmessung im Rechtschreibtest (RST) und im Lesetest, $p_{krit} = 0,1$

IQ	Unterrichtsquote	RST	Lesen	Gesamt
–	0,22	0,08	0,17	0,47

Tabelle 16: Elternangaben (in %) zur Einschätzung der emotionalen Situation des Kindes

	Zeitpunkt 0		Zeitpunkt 4	
	Projektgruppe	Vergleichsgruppe	Projektgruppe	Vergleichsgruppe
Schreiben gern	27	65	32	63
Lesen gern	43	79	68	68
gern zur Schule	68	87	70	84
keine Schulangst	92	89	76	89

an die Tafel zu schreiben bzw. etwas laut in der Klasse vorzulesen und ob sie glauben, dass die anderen Kinder alles besser können.

Die Übersicht zu den Items des Kinderfragebogens ist in Tabelle 17.

Hier gibt es – wie bei den Leistungseinschätzungen der Kinder – einen allgemeinen Abwärtstrend, d.h. die Einschätzungen fallen im Mittel am Ende schlechter aus als am Anfang. Bezüglich des Schreibens und Lesens liegt die PG zu beiden Zeitpunkten unterhalb der VG. Bei den drei übrigen Items liegt die PG ebenfalls unterhalb der VG.

3.6 Bewertung der Förderung

Abschließend soll noch betrachtet werden, wie die Förderung bei Eltern und Kindern ankam. Tabelle 18 zeigt die Einschätzung der Kinder.

Lediglich 63 bis 76% der Kinder fanden die jeweilige Förderung «gut». In der RT-Gruppe ist der Anteil von

Tabelle 17: Kinderangaben (in %) zur Einschätzung von emotionaler Situation, Lesen und Rechtschreiben

	Zeitpunkt 0		Zeitpunkt 4	
	Projekt	Vergleich	Projekt	Vergleich
Schreibe gern	47	68	41	65
Lese gern	72	87	65	79
Traue mich, an die Tafel zu schreiben	50	65	27	42
Traue mich, laut vorzulesen	64	70	41	68
Andere können nicht alles besser	31	39	24	41
Keine Angst, was Falsches zu sagen	36	70	30	46

Tabelle 18: Kinder, Einschätzung «Wie gut gefällt dir der Förderunterricht», in %

%	gut	weniger gut	schlecht	sehr schlecht
RT	76	5	5	14
FU	63	31	6	0

«gut» zwar höher als in der FU-Gruppe, dafür gibt es aber mehr Einschätzungen in den Kategorien «schlecht» und «sehr schlecht» zusammen.

Bei den Eltern war die Zustimmung größer. Zur Frage, wie wichtig der Förderunterricht sei, antworteten lediglich zwei aus der RT-Gruppe (10%) und einer aus der FU-Gruppe (6%) «nicht so wichtig». Das heißt, in beiden Gruppen betrug die Zustimmung über 90%.

Die Eltern wurden am Ende auch gefragt, wie sich die Leistungen ihrer Kinder verändert hatten (Tab. 19 und 20). Beim Schreiben kommen beide Methoden etwa gleich gut weg; in der RT-Gruppe ist der Anteil «deutlicher» Verbesserungen zwar höher als in der FU-Gruppe, dafür gibt es aber auch mehr «unveränderte» Einschätzungen. Die Verbesserungen im Lesen sind nach Elternmeinung größer als die im Schreiben. Hier ist die Bewertung in der RT-Gruppe etwas besser: Über die Hälfte der Kinder haben sich nach Einschätzung der Eltern «deutlich verbessert».

4. Diskussion

Im Sinne unserer Hypothesen lassen sich die Ergebnisse (im Hinblick auf die durchgeführten Rechtschreibtests) wie folgt zusammenfassen:

1. Schulischer Förderunterricht bei rechtschreibschwachen Schülern führt zu einer Verbesserung der Rechtschreibleistung (und auch der Leseleistung) relativ zur Normgruppe.
2. Die beiden von uns angewandten Methoden unterscheiden sich nicht hinsichtlich ihrer Wirksamkeit.

Zu Punkt 1 ist anzumerken, dass die Ergebnisse wegen der Werte der Vergleichsgruppe schwer einzuordnen sind. Letztere hat sich von Testzeitpunkt 2 zu Testzeitpunkt 4,

Tabelle 19: Elterneinschätzung der Verbesserung der Rechtschreibleistung gegenüber dem Beginn der Förderung

%	deutlich verbessert	leicht verbessert	unverändert	verschlechtert
RT	45	35	20	0
FU	33	66	0	0

Tabelle 20: Elterneinschätzung der Verbesserung der Leseleistung gegenüber dem Beginn der Förderung

%	deutlich verbessert	leicht verbessert	unverändert	verschlechtert
RT	45	35	20	0
FU	33	66	0	0

also innerhalb eines Jahres, um 4 T-Wertpunkte verbessert.

Zwei mögliche Ursachen für die Verbesserung der Vergleichsgruppe müssen diskutiert werden: erstens könnte es sein, dass das Ergebnis ein Artefakt ist, was durch unterschiedlich «schwere» Normen der verwendeten Tests zustande gekommen ist. Zweitens könnte es sein, dass durch das hohe Engagement der hauptsächlich für den Förderunterricht zuständigen Lehrerin insgesamt eine Verbesserung des Deutschunterrichts stattgefunden hat. In den letzten Jahren wurden regelmäßige Rechtschreibtests am Schuljahresende bei allen Schülern eingeführt, so dass die Lehrerinnen jeweils recht objektiv zurückgemeldet bekamen, was die Schüler bei ihnen gelernt hatten und auch, welche Schüler rechtschreibschwach sind. Möglicherweise führte dies dazu, dass speziell diese Schüler gezielt gefördert wurden. Außerdem könnte ganz allgemein durch den Fokus der Schule die Qualität der Schriftsprachvermittlung verbessert worden sein.

Bereits 1997 bis 1999, vor der Etablierung des Modellprojektes, wurde von uns an dieser Schule ein Jahrgang u.a. mit Rechtschreibtests untersucht, und zwar jeweils zum Ende der 2., 3. und 4. Klasse, mit den Tests WRT2+, WRT3+ und GRT4+. Die resultierenden Klassenmittelwerte (T-Werte) waren 45,1 (WRT2+), 42,9 (WRT3+) und 42,4 (GRT4+). Hier kam es also sogar zu einer leichten Verschlechterung. Auch dies spricht eher dafür, dass die von uns beobachtete Verbesserung der jetzigen Vergleichsgruppe auf die speziellen Bedingungen dieser Schule zurückzuführen ist.

Was die Testnormen angeht, so scheint es in der Tat so zu sein, dass der DRT3 leichter ist als der WRT3+. Dies ist zunächst erwartungswidrig, da der DRT3 älter ist als der WRT3+, allerdings sind die Normen des DRT3 aktueller als die des WRT3+.

In unseren Daten können wir den DRT3 nicht durch den WRT3+ ersetzen, da nicht alle Kinder der Projektgruppe beide Tests gemacht haben.

Zur zweiten Hypothese fällt das Ergebnis klar aus: es gibt praktisch keinen Unterschied zwischen den beiden Interventionsmethoden; die RT-Gruppe verbessert sich um 3,8 T-Punkte und die FU-Gruppe um 3,7 T-Punkte. Beide Verbesserungen sind signifikant, ein Vergleich zwischen den Gruppen erübrigt sich wegen der praktisch identischen Werte. Die Effekstärken fallen – auch wegen der differierenden Ausgangsstandardabweichungen – leicht unterschiedlich aus (0,70 vs. 0,57 für die Rechtschreibung).

Neben der offensichtlichen Interpretation, nämlich dass beide Methoden gleich wirksam sind, gibt es einige weitere Aspekte:

- Die Stichprobenselektion: Beide Projektgruppen bestanden im Mittel aus nicht besonders schwer betrof-

fenen Kindern. Der mittlere T-Wert im Rechtschreibtest betrug bei Studienbeginn 40,8, lag also knapp im Normbereich.

- Durch die spezielle Selektion der Lehrerinnen wurden Schüler mit Teilleistungsstörung im engeren Sinn, also Defiziten bevorzugt in der Rechtschreibung, eher nicht in die Fördergruppen aufgenommen.
- Da es sich beim Förderunterricht um eine in der Schule etablierte Methode, beim Rechtschreibtraining aber um eine neue und von außen in die Schule getragene Methode handelte, sind Auswirkungen ähnlich derer eines Versuchsleitereffektes denkbar.

Die Auswirkungen dieser Faktoren sind unklar und es muss jedenfalls für diese spezielle Konstellation festgehalten werden, dass die beiden Interventionen gleich wirksam sind. Ein Vorteil eines regelgeleiteten Vorgehens im Vergleich zum «herkömmlichen» bloßen visuellen Trainieren (das ist das Konzept der Computerprogramme) findet sich nicht.

Für beide Interventionsmethoden (unabhängig von der Klassenstufe der Kinder) findet sich eine charakteristische «Lernkurve»: es kommt bis zur ersten Zwischenmessung zu einer Verschlechterung der Rechtschreibung, und danach steigen die Werte an. Dies gilt lediglich nicht für die jüngeren FU-Kinder, bei denen es von Messzeitpunkt 2 zu 3 zu einer leichten Verschlechterung kommt, die dann aber zum 4. Messzeitpunkt wieder ausgeglichen ist. Ursache für den Verlauf im ersten Jahr der Förderung ist möglicherweise, dass die Kinder einen längeren Zeitraum benötigen, um sich die neue Lernstrategie anzueignen.

Die Leistungseinschätzungen von Eltern, Kindern und Lehrern fallen recht uneinheitlich aus. Die Kinder schätzen Schreiben und Lesen am Ende der Förderung schlechter ein als am Anfang, was vermutlich nur bedeutet, dass Kinder im frühen Grundschulalter ihre Leistung nicht realistisch einschätzen können. Bei den Elternangaben finden sich leichte Verbesserungen im Lesen und Schreiben, außer im Schreiben für die FU-Gruppe. Bei den Lehrereinschätzungen ist aber gerade der Wert für die Rechtschreibung der FU-Gruppe im Vergleich zu den anderen Werten relativ stark angestiegen (0,7 Einheiten auf der Notenskala). Man bekommt insgesamt den Eindruck, dass zwar ganz allgemein eine leichte Verbesserung bei Eltern und Lehrern wahrgenommen wird, dass diese aber sehr unspezifisch ist.

Bei den emotionalen Items ist eine zusammenfassende Darstellung ebenfalls nur schwer möglich. Deutliche Verbesserungen der Förderkinder in Relation zur Vergleichsgruppe können aber nicht festgestellt werden.

Eine wichtige Frage ist, ob sich die Verbesserung der Rechtschreibleistung anhand irgendwelcher der erfassten Variablen vorhersagen lassen. Damit könnten dann even-

tuell die Kinder beschrieben werden, die besonders gut (oder eben nicht) von einer der Interventionen profitieren.

Auch bei dieser Frage nach Prädiktoren der Verbesserung waren die Fragebogendaten nicht hilfreich. Bezüglich der «harten Daten», also der Werte von Rechtschreib- und Lesetests, IQ und der Unterrichtsquote, ergab sich zunächst mal ein triviales Ergebnis, dessen Ausprägung aber doch überrascht: Je mehr Unterricht stattfindet (innerhalb der von uns untersuchten Spanne), desto höher fällt die Verbesserung in der Rechtschreibung aus. Die anderen berücksichtigten Faktoren sind von vergleichsweise geringer Bedeutung. Der Zusammenhang zum Ausgangsniveau in der Rechtschreibung ist (allerdings in den Gruppen unterschiedlich stark) negativ, d.h. die Kinder mit der schlechtesten Ausgangsleistung erzielten die höchsten Verbesserungen. Auch dieses Resultat ist wenig überraschend, da sich hier auch zwei technische Faktoren niederschlagen: die Ausgangsmessung ist negativ in der Verbesserung (RST4-RST0) enthalten, daher ist eine negative Korrelation zur Ausgangsmessung zu erwarten. Zum anderen ist aufgrund der Selektion der Gruppen auch eine Regression zur Mitte zu erwarten, die die schlechtesten Rechtschreiber der Gruppen stärker betrifft. Ob darüber hinaus das Ausgangsniveau auch noch inhaltlich eine Rolle spielt, ist kaum zu entscheiden. Die Ergebnisse sprechen außerdem dafür, dass der IQ unter den hier realisierten Bedingungen keinen Einfluss auf das Ausmaß der Verbesserung hat. Bei der Einschätzung dieses Resultats ist allerdings zu berücksichtigen, dass Kinder mit IQ < 85 von vornherein aus der Studie ausgeschlossen worden waren. Rechtschreibschwache Kinder mit Intelligenz im Normbereich profitieren somit unabhängig vom jeweiligen individuellen IQ gleichermaßen von den Förderangeboten.

Die Förderung kommt bei der Mehrheit der Kinder und der Eltern gut an und wird auch für wirksam gehalten.

In der vorliegenden Studie wurde bei möglichst hoher ökologischer Validität die Förderung rechtschreibschwacher Kinder im Förderunterricht evaluiert. Dies bedeutet, dass die Bedingungen der Förderung relativ dicht an dem liegen, was in ähnlicher wenn auch vermutlich weniger systematischer Form in Schulen ohne wissenschaftliche Begleitung stattfindet. Es kann daher zunächst allgemein geschlossen werden, dass diese Art von Förderung effektiv ist, wenn die verwendeten Interventionsformen fundiert sind. Es kann ebenfalls gefolgert werden, dass die pro Kind tatsächlich stattfindende Zahl von Förderstunden der wichtigste Prädiktor für Erfolg ist. Bei den realisierten Bedingungen kann auch festgehalten werden, dass die verwendete Methode zumindest bei uns sich nicht auf die erreichte Verbesserung ausgewirkt hat. Es sind verschiedene Ursachen für diesen Befund denkbar. Ein Effekt des Konzepts der hohen ökologischen Validität war,

dass die Selektion der Kinder nicht zu einer Förderung von im engeren Sinne teilleistungsgestörten Kindern führte. Es ist denkbar, dass daher eine Interventionsmethode, die speziell für teilleistungsgestörte Kinder entwickelt wurde (Marburger Rechtschreibtraining), ihre Stärken in diesem Setting gar nicht ausspielen konnte. Es ist ebenfalls denkbar, dass die Lehrerinnen sich stärker mit «ihrer» Methode identifizierten und es daher zu einer Art Versuchsleitereffekt kam. Schließlich ist natürlich auch denkbar, dass die Methoden sich generell in ihrer Wirksamkeit nicht unterscheiden.

Die Höhe des erreichten Effektes (3,8 T-Wertpunkte für die Projektgruppen) mag zunächst gering erscheinen. Dieser Wert entspricht jedoch Effektstärken von 0,57 (FU) und 0,7 (RT). Diese Werte liegen nach der allgemein akzeptierten Nomenklatur von Cohen (1988) zwischen den Vergleichswerten «mittel» (0,5) und «hoch» (0,8).

Die Interpretation der Ergebnisse wird deutlich erschwert durch die Verbesserung der Vergleichsgruppe. Eine wirklich schlüssige Erklärung gibt es nicht. Am wahrscheinlichsten sind entweder Normprobleme der verwendeten Tests, oder die ganze Schule wurde durch das Modellprojekt quasi wachgerüttelt, und dies hat zu einer allgemeinen Verbesserung des Schriftsprachunterrichts geführt. Weitere Studien zu diesem Problem sind dringend erforderlich.

Eine sehr sinnvolle Ergänzung wäre eine Replikation bei weitgehend unverändertem Versuchsplan, wobei allerdings nur rechtschreibgestörte bzw. lese-rechtschreibgestörte Kinder nach ICD-10 (Schulte-Körne et al., 2001) in die Förderung aufgenommen würden. Eine weitere sehr sinnvolle Ergänzung ist die Evaluation des «normalen» Unterrichts, da kontrovers diskutiert wird, ob rechtschreibschwache Kinder überhaupt in eigenen Gruppen gefördert werden sollten, oder ob nicht eine Förderung im Klassenverbund («Binnendifferenzierung») sinnvoller ist.

Danksagung

Die vorliegende Studie wurde mit finanzieller Unterstützung der deutschen Forschungsgemeinschaft durchgeführt, der an dieser Stelle gedankt werden soll. Ferner gilt unser Dank der Schulleiterin der Grundschule Höchst, Frau M. Hoffmann, der Schulamtsdirektorin, Frau S. Thormählen, sowie den Lehrerinnen an der Grundschule Höchst, G. Richter und E. Zentgraf. Außerdem danken wir den Kindern und Eltern der Grundschule Höchst für ihre Mitarbeit bei diesem Projekt.

Literatur

- Bastian, U.; Buschmann, A.; Buschmann, R.; Jötten, B.; Richter, A.: *LRS-Förderung von der Vorklasse bis zum 4. Schuljahr: Bericht über einen Schulversuch*. In Dummer, L. (Hrsg.) Legasthenie. Bericht über den Fachkongress 1988, 247–271, 1988. Bundesverband Legasthenie e.V.
- Birkel, P.: *Weingartener Grundwortschatz Rechtschreib-Test für erste und zweite Klassen (WRT1+)*. Hogrefe, Göttingen, 1995.
- Birkel, P.: *Weingartener Grundwortschatz Rechtschreib-Test für zweite und dritte Klassen (WRT2+)*. Hogrefe, Göttingen 1994.
- Birkel, P.: *Weingartener Grundwortschatz Rechtschreib-Test für dritte und vierte Klassen (WRT3+)*. Hogrefe, Göttingen 1994a.
- Birkel, P.: *Grundwortschatz Rechtschreibtest für 4. und 5. Klassen (GRT4+)*. Beltz, Weinheim und Basel, 1990.
- Büttner, M.: *Die Erfolge schulischer und privater Fördermaßnahmen bei lese-rechtschreibschwachen Schülern*. In Dummer-Smoch, L.: Legasthenie. Bericht über den Fachkongress 1980. Bundesverband Legasthenie e.V., 1980.
- Cohen, J.: *Statistical Power Analysis for the Behavioral Sciences*. Erlbaum, New York, 1988.
- Grund, M.; Haug, G.; Naumann, C.L.: *DRT4. Diagnostischer Rechtschreibtest für 4. Klassen*. Beltz, Weinheim 1994.
- Hassmann-Kube, G.: *Die Freiburger LRS-Klassen*. In Dummer, L. (Hrsg.) Legasthenie. Bericht über den Fachkongress 1986, 289–291, 1986. Bundesverband Legasthenie e.V., 1986.
- Klicpera, C.; Gasteiger-Klicpera, B.: *Psychologie der Lese- und Schreibschwierigkeiten*. Beltz, Weinheim 1995.
- Landerl, K.; Wimmer, H.; Frith, U.: The impact of orthographic consistency on dyslexia: A German-English comparison. *Cognition* 63, 315–334, 1997.
- Landerl, K.; Wimmer, H.; Moser, E.: *SLRT – Salzburger Lese- und Rechtschreibtest*. Verlag Hans Huber, Bern 1997.
- Mannhaupt, G.: Deutschsprachige Studien zu Intervention bei Lese-Rechtschreibschwierigkeiten: Ein Überblick zu neueren Forschungstrends. *Zeitschrift für Pädagogische Psychologie* 8, 123–138, 1994.
- May 2001: *Welche schulische Förderung ist effektiv? Was leistet Förderunterricht und wie könnte er noch erfolgreicher sein?* In: G. Schulte-Körne (Hrsg.): Legasthenie: erkennen, verstehen, fördern. Beiträge zum 13. Fachkongress des Bundesverbandes Legasthenie 1999. Verlag Dr. Dieter Winkler, Bochum 2001.
- Müller, R.: *DRT3. Diagnostischer Rechtschreibtest für 3. Klassen*. Beltz, Weinheim, Basel 1997.
- Naegele, I.M., Valtin, R.: *LRS – Legasthenie in den Klassen 1–10*. Handbuch der Lese-Rechtschreibschwierigkeiten Band 1. Beltz, Weinheim 2001.
- Reuter-Liehr, C.: Behandlung der Lese-Rechtschreibschwäche nach der Grundschulzeit: Anwendung und Überprüfung eines Konzeptes. *Zeitschrift für Kinder- und Jugendpsychiatrie*, 21, 135–147, 1993.
- Roth, E.; Warnke, A.: Therapie der Lese-Rechtschreibstörung. *Kindheit und Entwicklung* 10, 87–96, 2001.

- Schäffer, K.: *Möglichkeiten und Grenzen der Förderung von lese-rechtschreibschwachen Grundschulern in einer Klein-klasse*. In Dummer, L. (Hrsg.) *Legasthenie. Bericht über den Fachkongress 1986*, 292–296. Bundesverband Legasthenie e.V., 1986.
- Schulte-Körne, G.; Mathwig, F.: *Das Marburger Rechtschreibtraining*. Winkler Verlag, Bochum 2001.
- Schulte-Körne, G.; Schäfer, J.; Deimel, W.; Remschmidt, H.: *Das Marburger Eltern-Kind-Rechtschreibtraining*. *Zeitschrift für Kinder- und Jugendpsychiatrie* 25, 151–159, 1997.
- Schulte-Körne, G.; Deimel, W.; Remschmidt, H.: *Das Marburger Eltern-Kind-Rechtschreibtraining – Verlaufsuntersuchung nach 2 Jahren*. *Zeitschrift für Kinder- und Jugendpsychiatrie* 26, 167–173, 1998.
- Schulte-Körne, G.; Deimel, W.; Hülsmann, J.; Seidler, T.; Remschmidt H.: *Das Marburger Rechtschreib-Training – Ergebnisse einer Kurzzeitintervention*. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 29, 7–15, 2001.
- Tacke, G.; Nock, H.; Staiber, W.: *Rechtschreibförderkurse in der Schule: Wie erfolgreich sind sie, und welche Faktoren tragen zu Leistungsverbesserungen bei?* *Zeitschrift für Pädagogische Psychologie* 1, 45–52, 1987.
- Weiß, R. H.: *Grundintelligenztest Skala 2 CFT 20*. Hogrefe, Göttingen, 1987.
- Weiß, R. H.; Osterland, J.: *Grundintelligenztest Skala 1 CFT 1*. Hogrefe, Göttingen, 1997.

PD Dr. Gerd Schulte-Körne

Klinik für Kinder- und Jugendpsychiatrie und -psychotherapie
der Philipps-Universität Marburg
D-35033 Marburg