

# "Schulanfang auf neuen Wegen"



## Abschlussbericht zum Modellprojekt

Arbeitskreis Wissenschaftliche Begleitung „Schulanfang auf neuen Wegen“

Mai 2006



Baden-Württemberg

MINISTERIUM FÜR KULTUS, JUGEND UND SPORT

## Impressum

Herausgeber: Ministerium für Kultus, Jugend und Sport Baden-Württemberg  
Schlossplatz 4 (Neues Schloss), 70173 Stuttgart

Verantwortlich: Christa Engemann,  
Ministerium für Kultus, Jugend und Sport Baden-Württemberg

Redaktion: Rolf Thiel

Arbeitskreis Wissenschaftliche Begleitung „Schulanfang auf neuen Wegen“:

- Prof. Dr. Marcus Hasselhorn (Leiter der Abt. Pädagogische Psychologie und Entwicklungspsychologie an der Universität Göttingen), Ausführungen zu: Kognitiver Entwicklungsstand und Zusammenhang von kognitiven Variablen, Einschätzungen von Schülervariablen durch Lehrerinnen/Lehrer mit der Schulleistung der Schüler und Entwicklung ihrer Motivation (Lernfreude), bezogen auf die einzelnen Modelle
- Prof'in Dr. Margareta Götz (Grundschulpädagogik, bis 1999 PH KA) und Prof'in (em.) Dr. Elisabeth Neuhäus-Siemon (Grundschulpädagogik), beide Universität Würzburg): Interpretation der erhobenen Daten zur Unterrichtsgestaltung bei den A- und B-Modellklassen und den Vergleichsklassen Erfassung der Einstellungen zum Modellversuch bei den daran teilnehmenden Lehrerinnen und Lehrern
- Dipl.-Päd. Hubert Kucher (Landesinstitut für Schulentwicklung): Allgemeine Angaben zur wissenschaftlichen Begleituntersuchung (Ziele, Organisation, Stichprobenbeschreibungen), Redaktion
- Dipl.-Psych. Ingeborg Leinenbach (Oberpsychologierätin, Schulpsychologischen Beratungsstelle, Stuttgart),
- Dr. Bernd Lehmann (Direktor des GHS-Seminars in Sindelfingen) Einführung und Ziele der wissenschaftlichen Begleituntersuchung, Unterrichtsgestaltung der Erzieherinnen/Sozialpädagoginnen in den Grundschulförderklassen
- Dipl.-Psych. Rolf Thiel (Leiter des Arbeitskreises, Psychologiedirektor, Landesinstitut für Schulentwicklung): Entwicklung des Beobachtungsbogens für Schüler an Grundschulen durch die Lehrerinnen/die Lehrer (BSG-L) und des Einstellungstests zum Sozialverhalten in Gruppen (SVG); Ergebnisse, bezogen auf die einzelnen Modelle und auf den Entwicklungsverlauf der Einstellung zum Sozialverhalten der Schüler.

Erscheinungsjahr: Mai 2006



# **Abschlussbericht zum Modellprojekt**

Arbeitskreis  
Wissenschaftliche Begleitung „Schulanfang auf neuen Wegen“

# Inhaltsverzeichnis

Vorwort	4
I. Ziele und Rahmenbedingungen	7
I.1 Ziele des Schulversuchs	7
Bildungspolitische Zielsetzungen	7
Schulorganisatorische Zielsetzung	7
Pädagogische Zielsetzung	7
I.2 Ziele der wissenschaftlichen Begleituntersuchung	8
I.3 Rahmenbedingungen	9
Beschreibung der Modelle	9
Organisatorische Positionierung der Begleituntersuchung	10
Durchführung der wissenschaftlichen Begleituntersuchung	11
II. Untersuchungen und Untersuchungsergebnisse (Kohorte 1 und Kohorte 3)	12
II.1 Konzept, Versuchsplan und Stichprobenbeschreibung (Kohorte 1)	12
Versuchsschulen im Schulversuch „Schulanfang auf neuen Wegen“	12
II.2 Intensiverhebungen an den 8 Intensivschulen (Kohorte 1)	15
Erster Messzeitpunkt der Intensiverhebung im Schuljahr 1997/98	15
Zweiter Messzeitpunkt der Intensiverhebung im Schuljahr 1998/99	16
II.3 Hauptuntersuchung Kohorte 3 (Klassenstufe 1 bis 4) Konzept, Versuchsplan und Stichprobenbeschreibung	17
Versuchsschulen im Schulversuch „Schulanfang auf neuen Wegen“	17
Versuchsplan	20
II.4 Schülerverhalten aus Sicht der Lehrerinnen und Lehrer (Kohorte 1)	23
Bogen zur Schülerbeobachtung in den ersten Grundschulklassen (BSG-L)	23
II.5 Entwicklungsverläufe	29
Vorbemerkungen	29
Kognitive Entwicklung	29
Motivationale Entwicklung	34
Soziales Verhalten und soziale Einstellungen	37
Einstellungstest: SVG – Sozialverhalten in Gruppen (Klassen 1 und 2)	37
Einstellungstest: SVG-4.1 – Sozialverhalten, Soziale Kompetenz und Schulzufriedenheit (Klasse 4)	48

---

II.6 Schulische Leistungen	66
Schulische Leistungen gemessen mit standardisierten Verfahren	66
Abschneiden in den landesweiten Vergleichsarbeiten in den Fächern Deutsch und Mathematik am Ende von Kl. 3	73
Leistungsentwicklung von potentiell leistungsstarken Schülern und von Risikoschülern in Abhängigkeit vom Modell	74
II.7 Noten in den Fächern Deutsch und Mathematik in den Klassen 2 bis 4 und Bildungsempfehlung Ende Kl. 4	79
Zeugnisnoten Kl. 2 bis Kl. 4	79
Alter und Bildungsempfehlung, aufgebrochen nach Modell (Stand Juli 03)	80
II.8 Unterrichtsgestaltung	83
Unterrichtsbezogene Ziele der wissenschaftlichen Begleitung	83
Anlage und Durchführung der Hauptuntersuchung	83
Ergebnisse der Hauptuntersuchung	85
Zusammenfassung und Diskussion der Ergebnisse der Hauptuntersuchung	99
Anschlussuntersuchung in der dritten und vierten Jahrgangsstufe	102
II.9 Unterrichtsgestaltung Grundschulförderklassen	105
Kooperation zwischen den beteiligten Pädagogen	110
Organisationsformen	110
Übertritt von Kindern der Grundschulförderklasse in die 1. Klasse der Grundschule	111
Einstellungen der teilnehmenden Erzieherinnen/Sozialpädagoginnen	112
Diskussion B1-Modell	115
II.10 Die Elternmeinung zu den Modellen	117
Einführung	117
Methodische Vorbemerkung	117
Ergebnisse	117
III. Bilanz der Ergebnisse über die verschiedenen Modelle in der Eingangsstufe der Grundschule und Konsequenzen für die Bewertung der Ergebnisse	122
III.1 Eingangsvoraussetzungen der Schüler	122
Kognitive Entwicklung	122
Motivationale Entwicklung	122
III.2 Soziales Verhalten und soziale Einstellungen	123
III.3 Schulische Leistungen Kl. 1 bis Kl. 4	123
III.4 Noten in den Fächern Deutsch und Mathematik (Kl. 2 bis Kl.4)	124
III.5 Das Durchschnittsalter der Schüler am Ende von Kl. 4	124

III.6	Bildungsempfehlung und Modell	125
III.7	Unterrichtsgestaltung	125
	In der Schuleingangsstufe (Klassenstufe 1 und 2)	125
	Anschlussuntersuchung in der dritten und vierten Jahrgangsstufe	125
III.8	Unterrichtsgestaltung Grundschulförderklasse	126
III.9	Die Elternmeinung zu den Modellen	126
III.10	Generelles Fazit	126
	Literatur	129

## Vorwort

Mit diesem Bericht legt der „Arbeitskreis Wissenschaftliche Begleitung“ die Ergebnisse der mehrjährigen Begleituntersuchung des Projekts „Schulanfang auf neuen Wegen“ vor.

Das gleichnamige Symposium im Mai 1996 war der Ausgangspunkt für weit reichende pädagogische und organisatorische Veränderungen in der Eingangsstufe der Grundschule in Baden-Württemberg. Mit dem Projekt „Schulanfang auf neuen Wegen“ werden Lösungsansätze für Probleme der Einschulungspraxis und der Eingangsstufe erprobt, die bislang zu hohen Zurückstellungsquoten und folglich zu hohem durchschnittlichem Einschulungsalter geführt haben.

Der baden-württembergische Modellversuch gilt im Spektrum der in der letzten Dekade eingeleiteten länderspezifischen Reformmaßnahmen als der elaborierteste und als pädagogisch höchst anspruchsvoller Reformversuch.

Die Ergebnisse der wissenschaftlichen Begleituntersuchung sind geeignet, neue Wege in möglichst vielen Grundschulen des Landes zu etablieren, die einen kindgerechten Start in das Schulleben erleichtern.

Das Projekt „Schulanfang auf neuen Wegen“ war breit angelegt. Es gab mehrere Modellvarianten, die unter den Abkürzungen A1, A2, B und C bekannt geworden sind.

Der „Arbeitskreis Wissenschaftliche Begleitung Schulanfang auf neuen Wegen“ wurde 1997 beauftragt, Auswirkungen der Varianten A1, A2 und B

auf den Lern- und Entwicklungsverlauf der teilnehmenden Kinder,  
auf die Einstellungen der Eltern, Lehrerinnen und Lehrer sowie der Erzieherinnen,  
auf die Unterrichtsgestaltung der Lehrerinnen und Lehrer und evtl. Erzieherinnen

zu untersuchen. Zu diesem Zwecke wurde der Vergleich mit Schulen geplant, die weiterhin nach der herkömmlichen Praxis mit der Einschulung und dem Anfangsunterricht verfahren. Über einen Grundschuljahrgang hinweg, beginnend in Klasse 1 bis zum Ende der Klasse 4, also insgesamt über 4 Schuljahre, sollten die Kinder regelmäßig individuell untersucht, die Lehrerinnen und Lehrer sowie die Eltern der Kinder der in die wissenschaftliche Begleituntersuchung einbezogenen Klassen befragt werden.

Im Verlauf der Begleituntersuchung musste das Untersuchungsdesign aus Kostengründen reduziert und die Anzahl der einbezogenen Schulen resp. Klassen sowie die diagnostisch relevanten Einzeluntersuchungen der Kinder deutlich eingeschränkt werden. Diese Veränderungen wirken sich auf die Aussagekraft der erhobenen Daten aus; darauf wurde bereits in den beiden Vorgängerberichten hingewiesen, auch im Abschlussbericht wird bei der Darlegung der Befunde ausdrücklich darauf eingegangen.

Den ersten Bericht – er umfasste zwei Teilbände - legte der Arbeitskreis dem Ministerium für Kultus, Jugend und Sport im Juli 2000 vor (vgl. Arbeitskreis, 2000).

Zu den Inhalten: Der Bericht behandelt hauptsächlich Ergebnisse, die bei Erhebungen, Befragungen und Untersuchungen der Schüler in Kl. 1 und 2 (und der Befragungen ihrer Lehrerinnen und Lehrer) des ersten von der Begleituntersuchung erfassten Schülerjahrgangs, der sog. Kohorte 1, gewonnen wurden.

Hierzu mussten vom Arbeitskreis einige Verfahren neu konstruiert bzw. aus Teilen schon bestehender Verfahren zusammengestellt werden. Besonders aufwändig gestaltete sich die

Feststellung des kognitiven Entwicklungsstands der Kinder, weil hier aus methodischen Gründen auf eine Individualtestung nicht verzichtet werden konnte.

All diese neu konstruierten oder neu zusammengestellten Verfahren wurden zunächst auf ihre Tauglichkeit hin an einer Stichprobe von acht Schulen („Intensivschulen“) mit 220 Kindern erprobt.

Diesen Ergebnissen konnte wissenschaftsmethodisch zunächst nur ein „Trend-“ bzw. „Hypothesencharakter“ zuerkannt werden. Ihre Bestätigung bzw. ihre Modifikation erhielten sie in der Hauptuntersuchung, die 41 Schulen mit ca. 1000 Kindern umfasste.

Zu anderen Untersuchungsbereichen, z.B. der Unterrichtsgestaltung der Lehrerinnen/Lehrer oder der Einstellung der Schüler zum Sozialverhalten in Gruppen konnten bereits im Zwischenbericht erste Verlaufsergebnisse einer großen Stichprobe, die etwa 80 Schulen umfasste, vorgelegt werden.

Im Juli 2002 legte der „Arbeitskreis Wissenschaftliche Begleitung“ dem Ministerium für Kultus, Jugend und Sport den zweiten Bericht<sup>1</sup> vor. Dieser enthält zum einen die Erhebungsergebnisse des Zwischenberichts vom Juli 2000 und zum anderen die Ergebnisse der sog. Hauptuntersuchung, hier die Ergebnisse der Erhebungen des dritten Schülerjahrgangs (Kohorte 3). Es handelt sich dabei um die Darstellung der Ergebnisse, die in der Eingangsstufe (Kl. 1 und Kl. 2) gewonnen wurden. Je nach Untersuchungsbereich konnten Ergebnisse mehrerer Kohorten oder nur aus Kohorte 3 berichtet werden.

Die Hauptuntersuchung begann mit dem Schuljahr 1999/00 an 41 Schulen in Klasse 1. Sie erfasste die Schüler der Kohorte 3, des dritten Schülerjahrgangs dieser Begleituntersuchung. Schwerpunkt bei den Untersuchungen in Kohorte 3 bildeten die mittels Individualtestung an einer ausreichend großen Schülerzahl – insgesamt um die 1000 Schülerinnen und Schüler – gewonnenen Ergebnisse zum kognitiven Entwicklungsstand der Schüler zu Beginn ihres Schulbesuchs in Kl. 1, die Entwicklung des kognitiven Entwicklungsstands zu Beginn der Klasse 2, ebenso die zu denselben Zeitpunkten gemessenen Ausprägungen der motivationalen Entwicklung (Lernfreude). Zusätzlich wurden die ersten standardisierten Verfahren zur Messung bestimmter Aspekte der Schülerleistungen im Schreiben, Lesen und Rechnen in den Klassen 1 und 2 eingesetzt.

Dieser wesentlich größeren Schülerzahl in den genannten Untersuchungsfeldern stand eine erhebliche Reduzierung der Schülerzahl (über 50 %) in anderen Untersuchungsbereichen der Hauptuntersuchung gegenüber. Diese Reduzierung war das Resultat einer aus Kostengründen vorgenommenen Verminderung der Anzahl der in die Begleituntersuchung einbezogenen Schulen und einer Beschränkung bei den Modellerprobungsschulen auf die Erprobungsklassen (relativ viele Grundschulen boten neben den Erprobungsklassen herkömmliche Jahrgangsklassen an). Waren es in Kohorte 1 noch knapp über 80 Schulen, so verblieben nur noch 41 Schulen mit gleichzeitiger - dies wurde soeben dargestellt - Reduzierung der in einer Schule untersuchten Klassenzahl.

So können andere Aussagen, z.B. zum Sozialverhalten der Kinder, zur Beschreibung des Schülerverhaltens durch die Lehrerinnen/Lehrer oder zur Unterrichtsgestaltung der Lehrerinnen/Lehrer der A1-, A2-, B-Erprobungsschulen und der Vergleichsschulen der Kohorte 1 sich auf deutlich größere Schülerstichproben als die in Kohorte 3 beziehen und teilweise zusätzlich noch auf Wiederholungsmessungen.

Alle über die Eingangsstufe hinausreichenden Ergebnisse in Kl. 3 und 4 basieren nur auf Daten der Kohorte 3.

---

<sup>1</sup> Vgl. Vorläufiger Abschlussbericht zur Eingangsstufe der Grundschule



Sämtliche Berichte sind das Gemeinschaftsprodukt des Arbeitskreises. Um dies zu dokumentieren, werden bei der Vorstellung der einzelnen Mitglieder die Arbeitsbereiche angegeben, bei denen sie die Federführung hatten. Im Einzelnen sind dies:

Frau Prof.'in Dr. Margareta Götz (Grundschulpädagogik) und Frau Prof.'in (em.) Dr. Elisabeth Neuhaus-Siemon (Grundschulpädagogik), beide Universität Würzburg:

Entwicklung eines Fragebogens zur Unterrichtsgestaltung für die A- und B-Modellklassen und die Vergleichsklassen, Interpretation der erhobenen Daten zur Unterrichtsgestaltung bei den A-, B-Modellklassen und den Vergleichsklassen, Erfassung der Einstellung zum Modellversuch bei den teilnehmenden Lehrerinnen und Lehrern; Entwicklung eines Elternfragebogens.

Herr Prof. Dr. Marcus Hasselhorn (Leiter der Abt. Pädagogische Psychologie und Entwicklungspsychologie an der Universität Göttingen): Kognitiver Entwicklungsstand und Zusammenhang von kognitiven Variablen; Einschätzungen von Schülervariablen durch Lehrerinnen/Lehrer mit der Schulleistung der Schüler und Entwicklung ihrer Motivation (Lernfreude); jeweils bezogen auf die einzelnen Modelle und gesondert für potentiell leistungsstarke sowie Risikoschüler.

Herr Hubert Kucher M.A. (Landesinstitut für Schulentwicklung):

Allgemeine Angaben zur wissenschaftlichen Begleituntersuchung (Ziele, Organisation, Stichprobenbeschreibungen); Entwicklung von Fragebögen zur Unterrichtsgestaltung der Erzieherinnen / Sozialpädagoginnen in den Grundschulförderklassen und Interpretation; Organisation der Erhebungen; Entwicklung eines Elternfragebogens; Redaktion und Layout des Endberichts.

Herr Dr. Bernd Lehmann (Direktor des Staatlichen Seminars für Didaktik und Lehrerbildung in Sindelfingen) und Frau Dipl.-Psych. Ingeborg Leinenbach (Psychologiedirektorin, Leiterin der Schulpsychologischen Beratungsstelle für die Landeshauptstadt Stuttgart):

Ziele und Rahmenbedingungen des Schulversuchs und der wissenschaftlichen Begleituntersuchung; Entwicklung eines Fragebogens zur Unterrichtsgestaltung der Erzieherinnen / Sozialpädagoginnen in den Grundschulförderklassen; Entwicklung eines Elternfragebogens.

Herr Dipl.-Psych. Rolf-Dietmar Thiel (Leiter des Arbeitskreises, Psychologiedirektor, Landesinstitut für Schulentwicklung):

Entwicklung des Beobachtungsbogens für Schüler an Grundschulen durch die Lehrerinnen/die Lehrer (BSG-L) und des Einstellungstests zum Sozialverhalten in Gruppen (SVG); Auswertungen, bezogen auf die einzelnen Modelle und auf den Entwicklungsverlauf der Einstellung zum Sozialverhalten der Schüler; Entwicklung eines Elternfragebogens; Redaktion und Layout des Endberichts.

Nicht Mitglied des Arbeitskreises, aber doch besonders zu nennen ist

Herr Dipl.-Psych. Peter Sikorski (Psychologiedirektor, Landesinstitut für Schulentwicklung), der für den Arbeitskreis eine Reihe von Auswertungen und Testanalysen vornahm und in die Form von Arbeitsberichten brachte, und, was ebenso wichtig ist, alle Daten in eine Datenbank zusammengespielt hat.

Es soll an dieser Stelle nicht versäumt werden, allen Erzieherinnen, Lehrerinnen und Lehrern zu danken, die trotz teilweise erheblicher Belastung durch die Anforderungen, die die Umsetzung der neuen Modelle in die praktische Unterrichtsarbeit mit sich brachte, dennoch mit Sorgfalt auch die vom Arbeitskreis erbetenen Befragungen, Daten und Schülereinschätzungen durchgeführt bzw. ausgefüllt haben. Ebenso gilt unser Dank den Schulleiterinnen und Schulleitern, die uns bei der Durchführung der Untersuchungen tatkräftig unterstützt haben.

## **I. Ziele und Rahmenbedingungen<sup>2</sup>**

### **I.1 Ziele des Schulversuchs**

#### Bildungspolitische Zielsetzungen

In Baden-Württemberg wird seit mehr als zwei Jahrzehnten im Durchschnitt jedes zehnte Kind zurückgestellt - bei starken regionalen Schwankungen. Die vorzeitigen Einschulungen sind dagegen auf 1,5 Prozent gesunken. Ähnlich oder teilweise noch höher sind die Zurückstellungsquoten in anderen Bundesländern. Das durchschnittliche Einschulungsalter liegt damit im Schj. 1997/98 statistisch gesehen in Baden-Württemberg bei 6 Jahren und 7 Monaten. Viele Kinder haben also ihren siebenten Geburtstag schon hinter sich, wenn sie in die Schule kommen.

Dies hat Auswirkungen auf die Entwicklungsstände von Kindern, auf die der Anfangsunterricht trifft, und folglich auf die schulorganisatorische (konzeptionelle) Gestaltung des Eingangunterrichts und die pädagogischen, didaktischen und methodischen Konzepte für die Grundschule.

Mit diesem Projekt soll das angestiegene durchschnittliche Schuleintrittsalter durch

- den Abbau der hohen Zurückstellungsquote und
- die Erhöhung der Zahl der vorzeitig eingeschulter Kinder

vermindert werden, ohne dass es dabei zu ungünstigen Auswirkungen auf die Entwicklung der betroffenen Kinder kommt.

#### Schulorganisatorische Zielsetzung

Die angestrebte Senkung des durchschnittlichen Schuleintrittsalters setzt einen an den individuellen Voraussetzungen orientierten Schulstart für möglichst viele schulpflichtige Kinder voraus. Im Projekt soll dies angestrebt werden durch

- eine Flexibilisierung des Einschulungszeitpunktes und
- eine Flexibilisierung der individuellen Verweildauer in der Schuleingangsstufe.

#### Pädagogische Zielsetzung

Es gehört zu den bildungspolitischen Anliegen des Schulversuchs, neben schulorganisatorischen Innovationen auch Veränderungen in der didaktisch-methodischen Gestaltung des Unterrichts in der Schuleingangsstufe zu initiieren. Damit ist die Absicht verbunden, der wachsenden Heterogenität der Kinder Rechnung zu tragen.

Bei der Realisierung der Zielsetzungen ist den Varianten der Vorzug zu geben, die eine Separation vermeiden und eine Integration möglichst vieler Kinder in die Eingangsstufe ermöglichen.

---

<sup>2</sup> Bei der Darstellung der Ziele und Rahmenbedingungen stützt sich dieser Bericht auf das Konzeptionspapier des Ministeriums für Kultus, Jugend und Sport Baden-Württemberg, Christa Engemann, Ref. 33: „Schulanfang auf neuen Wegen. Konzeption und operative Planung der Erprobungen an Grundschulen in Baden-Württemberg ab dem Schuljahr 1997/98“, Februar 1997.

## I.2 Ziele der wissenschaftlichen Begleituntersuchung

Die wissenschaftliche Begleitevaluation zu dem Projekt sollte vor allem Aufschluss darüber bringen, welche Konsequenzen die veränderte Schuleingangsstufe im Vergleich zur bisherigen Einschulungspraxis für den Entwicklungsverlauf der betroffenen Kinder hat.

So sollte die Begleituntersuchung u.a. Antworten auf die folgenden Fragestellungen geben:

- Gibt es Unterschiede zwischen Kindern aus den Erprobungsklassen in den Modellen A1, A2 und B und solchen aus den Vergleichsklassen mit der bisher üblichen Schuleingangspraxis, und zwar im Hinblick auf die Entwicklung

*der schulischen Fähigkeiten und Fertigkeiten im Bereich des Lesens, Rechtschreibens, der Mathematik;*

*der schulrelevanten kognitiven Kompetenzen wie phonologische Bewusstheit, Aufmerksamkeit, Arbeitsgedächtnis, analoges Denken;*

*der sozialen Kompetenzen wie zum Beispiel Rücksichtnahme, Hilfsbereitschaft, Toleranz;*

*der Stützfaktoren des schulischen Lernens wie zum Beispiel Lernfreude, Psychomotorik, Instruktionsverständnis, Lern- und Arbeitsverhalten.*

- Unterscheiden sich die Kinder der Erprobungs- und Vergleichsklassen hinsichtlich der Entwicklungs- und Lernverläufe?
- Gibt es Unterschiede in der Unterrichtsgestaltung zwischen Erprobungs- und Vergleichsklassen, was den Umfang und die Qualität der methodisch-didaktischen Ansätze angeht?
- Gibt es Unterschiede im Hinblick auf modellspezifische Einstellungen und Erfahrungen zwischen den am Modellversuch beteiligten Erzieherinnen und Lehrerinnen / Lehrern?
- Gibt es Unterschiede bzgl. der Einstellungen zur Schule zwischen den Eltern der Kinder aus Erprobungs- und Vergleichsklassen?
- Unterscheiden sich die einzelnen Modelle A1, A2 und B voneinander im Hinblick auf die oben genannten Variablen?
- Gibt es modellspezifische Effekte bzgl. des Entwicklungsverlaufs von Teilgruppen von Kindern, die auf Grund ihrer Schuleingangsvoraussetzungen als potentiell leistungsstarke oder als Risikokinder eingestuft werden?

### I.3 Rahmenbedingungen

#### Beschreibung der Modelle

Zur Flexibilisierung des Schuleingangs werden 3 Modelle angeboten: A, B, C. In die wissenschaftliche Begleituntersuchung gehen jedoch nur die Modelle A und B ein. Diese werden im Folgenden kurz skizziert.

#### Die A-Modelle:

In den A-Modellen wird auf Zurückstellung verzichtet, d.h. die Schuleingangsstufe wird so verändert, dass alle Kinder<sup>3</sup> eingeschult werden und eventuelle Entwicklungsrückstände, aber auch beschleunigte kognitive Entwicklungsverläufe („besonders befähigte Kinder“) innerhalb der Eingangsstufe aufgefangen bzw. angemessen unterstützt werden können. Es gibt zwei Varianten:

#### Modell A1: Eingangsstufe mit variabler Verweildauer von ein bis drei Jahren mit jährlicher Einschulung

Die Klassenstufen 1 und 2 werden zu einer einheitlichen jahrgangübergreifenden Eingangsstufe zusammengefasst, alle schulpflichtigen Kinder werden ohne Überprüfung der Schulfähigkeit aufgenommen, Verweildauer in der Eingangsstufe von ein bis drei Jahren je nach individuellem Lernzeitbedarf.

Unter Beibehaltung des jährlichen Einschulungszyklus kann dieses Modell auch in einer einzügigen Grundschule ohne große schulorganisatorische Umstellungen eingeführt werden. Im zweiten Jahr wird die Eingangsgruppe getrennt und durch Aufnahme je der Hälfte der neu einzuschulenden Kinder entstehen zwei parallele jahrgangsgemischte Eingangsstufen. Dafür entfällt die Klassenstufe 2. Diese Trennung der Gruppen ist nur im zweiten Modelljahr nötig. In späteren Jahren findet ein fließender Austausch durch jährliche Aufnahme von Schulneulingen bei gleichzeitiger Abgabe von Kindern nach Klasse 3 statt.

#### Modell A 2: Eingangsstufe mit halbjährlicher Einschulung und variabler Verweildauer

Zusätzlich zu den Gegebenheiten im Modell A1 ist hier ein zweiter Einschulungstermin zum Schulhalbjahr (Februar / März) gegeben. Diese Kinder können dann bis zu 2,5 Jahre in der Schuleingangsstufe bleiben.

Im Vergleich zur einmaligen Einschulung im Modell A1 bedeutet dies:

- Einerseits eine individuelle, kindgerechtere Wahl des Einschulungszeitpunktes, stellt doch der Zeitraum von einem Jahr bei einem 6-jährigen Kind entwicklungspsychologisch eine große Entwicklungsspanne dar.
- Andererseits jedoch eine noch größere Heterogenität der Kindergruppe, eine Veränderung der sozialen Beziehungen in der Gruppe und einen hohen unterrichtlichen Differenzierungsbedarf für die Lehrerinnen und Lehrer.

---

<sup>3</sup> Mit Ausnahme der als förderschulbedürftig anerkannten

### Modell B: Grundschulförderklasse neuen Zuschnitts

Das Modell B setzt eine Grundschule mit einer Grundschulförderklasse voraus. Es stellt eine enge Verzahnung zwischen der schulvorbereitenden Einrichtung Grundschulförderklasse und dem ersten Schuljahr her.

Kennzeichnend hierfür ist

- eine intensive Kooperation des Personals beider Einrichtungen, also zwischen Lehrer/Lehrerin und Erzieherin/Sozialpädagogin;
- diese soll sich niederschlagen im Erstellen gemeinsamer Förderpläne für Kinder aus beiden Einrichtungen,
- im Austausch von Kindern zwischen erster Klasse und Grundschulförderklassen zu gemeinsamen Förderangeboten,
- in einer zeitlich flexiblen Regelung von Übergängen: Kinder der Grundschulförderklasse sollen je nach individuellem Entwicklungsfortschritt gleitend auch während des Schuljahres in die 1. Klasse überwechseln können.

*Zusatz:*

Im Versuchsablauf haben sich zwei B-Modelle herauskristallisiert:

- das gerade beschriebene, das mittlerweile B1-Modell genannt wird,
- das neue, das mittlerweile B2-Modell genannt wird.

Es kann kurz so beschrieben werden:

In Zusammenarbeit von Kindergarten und Grundschule werden Kinder, die mit ziemlicher Sicherheit von der Grundschule zurückgestellt und eine Empfehlung zum Besuch der Grundschulförderklasse erhalten würden, ein halbes Jahr vor Schulbeginn in eine dafür konzipierte Grundschulförderklasse übernommen. Die Erzieherin bzw. Sozialpädagogin, die die Kinder in der Grundschulförderklasse auf die Schule vorbereitet, begleitet ihre „Schützlinge“ noch ein halbes Jahr in der ersten Grundschulklasse. Diese Modellvariante erspart einer Reihe von Kindern ein ganzes Jahr, das sie sonst mit dem Besuch der Grundschulförderklasse alten Stils noch hätten aufwenden müssen.

Leider gibt es nur sehr wenig Schulen, die dieses interessante Modell bislang durchführen. Das Konzept der quantitativ angelegten Begleituntersuchung setzt zur Sicherung von signifikanten Unterschieden bzw. Befunden bestimmte Mindestgrößenordnungen voraus, diese wurden bedauerlicherweise in zu vielen Bereichen nicht erreicht. Deshalb konnte das B2-Modell in diese Begleituntersuchung nicht miteinbezogen werden.

## Organisatorische Positionierung der Begleituntersuchung

Nachdem am 6. Mai 1996 während des vom Ministerium für Kultus, Jugend und Sport Baden-Württemberg organisierten Symposiums Grundschule mit dem Titel „Schulanfang auf neuen Wegen“ Überlegungen zur Weiterentwicklung der Schuleingangsstufe mit Wissenschaftlern, Pädagoginnen und Pädagogen, Vertretern der Schulverwaltung, der Trägerverbände, der Lehrerverbände und der Kirchen diskutiert wurden, richtete Kultusministerin Frau Dr. Annette Schavan Expertengruppen ein, die den künftigen Erprobungsschulen konzeptionelle und organisatorische Hilfestellung geben bzw. deren Ergebnisse wissenschaftlich auswerten sollten.

Das Projekt wurde im Kultusministerium unter der Projektleitung von Frau Ministerialrätin Engemann, Referat 33, die auch für die Projektkonzeption maßgeblich verantwortlich zeichnet, angesiedelt.

### Durchführung der wissenschaftlichen Begleituntersuchung

Die wissenschaftliche Begleituntersuchung zum Projekt „Schulanfang auf neuen Wegen“ umfasste auf Grund von Finanzierungsproblemen (s.u.), länger als ursprünglich geplant, einen Zeitraum von 7 Jahren.

Um zu empirisch gesicherten Ergebnissen zu kommen, bezog die Untersuchung neben Schülern aus Modell- bzw. Erprobungsklassen auch solche von Vergleichsklassen als Kontrollgruppen ein.

Mit Beginn des Schuljahres 1997/98 startete die Erprobungsphase an 88 Schulen mit den Kindern, die im Herbst 1997 eingeschult wurden. Die Schulen erprobten jeweils das A - Modell (A1 oder A2) oder das B-Modell. Vergleichsklassen als Kontrollgruppen kamen hinzu.

Der Arbeitskreis wissenschaftliche Begleitung musste einige Verfahren als Messinstrumentarien erst neu entwickeln. Zu diesem Zweck wurden zu Beginn des Schuljahres 1997/1998 (im Weiteren als Kohorte 1 bezeichnet) in einer ersten Erprobungsphase in einer sog. Intensivtestung an 8 ausgewählten Schulen, den so genannten Intensivschulen, die entsprechenden Verfahren eingesetzt. Bei dieser Intensiverhebung (vgl. Kap. II.1) handelte es sich um umfangreiche individualdiagnostische Untersuchungen zum allgemeinen Entwicklungsstand der Schulanfängerinnen / Schulanfänger.

Bei allen teilnehmenden Schulen kamen Lehrerbeobachtungsbögen und Fragebögen zur Unterrichtsgestaltung zum Einsatz.

Der weitere vom Arbeitskreis geplante Untersuchungsverlauf wurde durch Haushaltskürzungen eingeschränkt: Zum einen wurde die Hauptuntersuchung um ein Schuljahr verzögert; zum anderen musste der Testleitereinsatz für die individualdiagnostischen Untersuchungen reduziert werden. Und zusätzlich führten diese Einsparungen zu einer Anpassung des Untersuchungsdesigns an die Haushaltsmittel. Dies bedeutet, dass nicht alle eingangsgestellten Fragen in dem gewünschten Umfang beantwortet werden können.

Ab dem Schuljahr 1999/2000 wurden in der Hauptuntersuchung, im Weiteren als Kohorte 3 bezeichnet, 41 sog. Stichprobenschulen (16 A1- , 9 A2- , 8 B-Modellschulen und 8 Vergleichsschulen) untersucht.

Für diese Kohorte 3 liegen komplette Ergebnisse für die Klassenstufen 1 und 2 vor.

Leider waren die Klassenstufen 3 und 4 (Schj.01/02 und 02/03) wieder von Haushaltskürzungen betroffen, so dass das geplante Programm erneut gekürzt werden musste; so wurde aus Kostengründen auf die Individualtestung verzichtet.

Besonders dem Engagement von Frau MR'in Engemann ist es trotz der vorher beschriebenen widrigen finanziellen Umstände zu verdanken, dass die Untersuchung wenigstens für einen Jahrgang dennoch bis zum Ende der Grundschulzeit durchgeführt werden konnte.

## **II. Untersuchungen und Untersuchungsergebnisse (Kohorte 1 und Kohorte 3)**

### **II.1 Konzept, Versuchsplan und Stichprobenbeschreibung (Kohorte 1)**

#### Versuchsschulen im Schulversuch „Schulanfang auf neuen Wegen“

Mit Beginn des Schuljahres 1997/98 begannen die Voruntersuchungen mit den Kindern, die im Herbst 1997 eingeschult wurden. Insgesamt 88 Schulen wurden mit Beginn dieses Schuljahres in die wissenschaftlich begleitete Untersuchung einbezogen. Es erproben

26 Schulen die Modellvariante A1,  
9 Schulen die Modellvariante A2,  
38 Schulen das Modell B.  
15 Schulen nehmen als Vergleichsschulen an dem Versuch teil.  
In diesen Anzahlen sind die 8 Intensivschulen enthalten.

Mit Beginn des Schuljahres 1998/99 nehmen insgesamt 78 Schulen an dem Versuch teil.  
Es erproben

24 Schulen die Modellvariante A1,  
9 Schulen die Modellvariante A2,  
35 Schulen das Modell B.  
10 Schulen nehmen als Vergleichsschulen an dem Versuch teil.  
In diesen Anzahlen sind die 8 Intensivschulen enthalten.

**Tabelle 1: Beschreibung der Analytestichproben**

<b>Schülerzahlen (Kohorte 1)</b>												
<b>Erhebung zum kognitiven Entwicklungsstand der Schüler und ihrer schulischen Leistungen</b>												
<i>Nur 8 Intensivschulen</i>												
<b>Schuljahre</b>	<b>Intensiverhebung (kognitiver Entwicklungsstand)</b>			<b>KNUSPEL Lesetest (nur Kl.2)</b>			<b>WRT 1 + , Rechtschreibtest (nur Kl. 2)</b>			<b>Noten Deutsch/Mathematik (nur Kl. 2)</b>		
	N= Gesamt	N= Erprob.	N= Vergl.	N= Gesamt	N= Erprob.	N= Vergl.	N= Gesamt	N= Erprob.	N= Vergl.	N= Gesamt	N= Erprob.	N= Vergl.
Schj. 97/98	219	97	122	-	-	-	-	-	-	-	-	-
Schj. 98/99	219	97	122	184	74	110	186	74	112	177	64	113

Anmerkung: Erprob. = Modell-Erprobungsklassen, Vergl. = Vergleichsklassen an Erprobungsschulen und Vergleichsschulen

<b>Erhebung nicht-kognitiver Schülervariablen (Sozialverhalten, Arbeitsverhalten, Feinmotorik etc.)</b>												
<i>Gesamtstichprobe und Intensivschulen (8 Schulen) getrennt ausgewiesen</i>												
<b>Schuljahre</b>	<b>Sozialverhalten in Gruppen (SVG)</b>						<b>Beobachtungsbogen für Schüler (BSG-L) [nur Schj.97/98]</b>					
	Gesamt	A1-Erprobungs-klassen	A2-Erprobungs-klassen	B-Erprobungs-klassen	A1-, A2-, B-Vergleichs-klassen	Klassen in Vergleichs-schulen	Gesamt	A1-Erprobungs-klassen	A2-Erprobungs-klassen	B-Erprobungs-klassen	A1-, A2-, B-Vergleichs-klassen	Klassen in Vergleichs-schulen
Schj. 97/98	1165	283	222	417	161	82	2432	530	152	865	180	413
Schj. 98/99	1434	396	213	321	128	376	-	-	-	-	-	-



<b>Erhebung nicht-kognitiver Schülervariablen (Sozialverhalten, Arbeitsverhalten, Feinmotorik etc.)</b> [Fortsetzung]						
<i>Gesamtstichprobe und Intensivschulen (8 Schulen) getrennt ausgewiesen</i>						
<i>Intensivschulen</i>	<b>SVG (Sozialverhalten in Gruppen)</b>			<b>Beobachtungsbogen für Schüler (BSG-L)</b>		
<b>Schuljahre</b>	N= Gesamt	N= Erprob.	N= Vergl.	N= Gesamt	N= Erprob.	N= Vergl.
Schj. 97/98	145	69	76	123	59	64
Schj. 98/99	192	79	113	-	-	-

<b>Lehrerbefragungen (Kohorte 1)</b>						
<b>Befragung zur Unterrichtsgestaltung</b>						
<i>Gesamtstichprobe</i>						
<b>Schuljahre</b>	<b>Fragebogen zur Unterrichtsgestaltung</b>					
	Gesamt	A1-Erprobungs- klassen	A2-Erprobungs- klassen	B-Erprobungs- klassen	A1-, A2-, B-Ver- gleichsklassen	Klassen in Ver- gleichsschulen
Schj. 97/98	132	33	11	48 *	19	21
Schj. 98/99	121	33	19	48	13	8

\* Zusätzlich liegen noch 29 ausgefüllte Fragebogen für die Grundschulförderklasse (GFK) vor.

## II.2 Intensiverhebungen an den 8 Intensivschulen (Kohorte 1)

Erster Messzeitpunkt der Intensiverhebung im Schuljahr 1997/98

Aus den 88 Schulen des Untersuchungsjahrgangs 1 wurden im September 1997 acht sog. Intensivschulen ausgesucht, um im Vorlauf zur Flächenuntersuchung die Testverfahren für die Intensiverhebung (Individualdiagnostik) der Schulanfänger/-innen zu erproben. Die Erhebung der Eingangsvoraussetzungen der Schülerinnen und Schüler (N=219) wurde in der Zeit vom 4. bis 24. Oktober 1997 durchgeführt.

Im Untersuchungsjahrgang 1 verteilen sich die 219 getesteten Schulanfänger der 8 Intensivschulen – 3 A1-Versuchsschulen, 2 A2-Versuchsschulen, 3 Vergleichsschulen – wie folgt:  
58 Schülerinnen und Schüler besuchen eine Erprobungsstufe im A1-Modell

39 Schülerinnen und Schüler besuchen eine Erprobungsstufe im A2-Modell

122 Schülerinnen und Schüler besuchen eine erste Klasse der herkömmlichen Art an einer Vergleichsschule (68 an einer Vergleichsschule, 14 an einer Schule mit parallelem A1-Modell, 40 an einer Schule mit gleichzeitig realisiertem A2-Modell).

Dabei kamen folgende Tests zum Einsatz:

Testbatterie Menge und Zahl (TMZ; Winkelmann, 1976). Daraus der

Subtest: Zahlabstraktion (ZA)

Subtest: Zählen von Objekten (ZO)

Subtest: Mengendefinition (MD)

Kaufman-Assessment-Battery for Children (K-ABC; Deutsche Version: Melchers & Preuß, 1993). Daraus der

Subtest: Handbewegungen

Subtest: Zahlen nachsprechen

Subtest: Bildhaftes Ergänzen

Subtest: Räumliches Gedächtnis

Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC; Jansen/Mannhaupt/Marx & Skowronek, 1996 bzw. 1999): Daraus der

Subtest: Pseudowörter nachsprechen

Subtest: Wort-Vergleich-Suchaufgabe

Subtest: Reimen

Subtest: Laute assoziieren

Subtest: Silben segmentieren

Subtest: Laut-zu-Wort

Lernfreude (in Anlehnung an Helmke, 1993)

Lese- und Schreibtests:

Sichtwörter

Lesen: Buchstabenkenntnisse

Zusatztest: Frühleser  
Schreibkenntnisse

Graphomotorische Testbatterie (Rudolf, 1986). Daraus die Subtests

Graphesthesia-Test (GMT II)  
Graphomotorischer Test (GT)

Zweiter Messzeitpunkt der Intensiverhebung im Schuljahr 1998/99

Das Vorhaben der wissenschaftlichen Begleitung, die Eingangserhebung mit Beginn des Untersuchungsjahrgangs 2 (Einschulung: im September 1998) auf alle Erstklässler der teilnehmenden Versuchsschulen auszudehnen, konnte nicht durchgeführt werden, da hierfür keine Finanzmittel zur Verfügung gestellt wurden.

Folglich wurde die zweite Eingangserhebung auf die 8 Intensivschulen begrenzt, und zwar ausschließlich auf die „Zweitklässler“ des Untersuchungsjahrgangs 1.

In der Intensiverhebung 1998 im Oktober 1998 wurden folgende Verfahren verwendet:

Testbatterie Menge - Zahl (TMZ; Winkelmann, 1976). Daraus der

Subtest Zahlabstraktion (ZA)  
Subtest Zählen von Objekten (ZO)

Kaufman-Assessment-Battery for Children (K-ABC; Deutsche Version: Melchers & Preuß, 1993). Daraus der

Subtest: Handbewegungen  
Subtest: Zahlen nachsprechen  
Subtest: Bildhaftes Ergänzen  
Subtest: Räumliches Gedächtnis

Phonologische Bewusstheit (Marx, 1997; Marx & Schneider, 2000). Daraus der

Subtest: Vokalersetzung  
Subtest: Restwortbestimmung  
Subtest: Phonemvertauschung

Kunstwort-Nachsprech-Test (Hasselhorn & Körner, 1997)

Lernfreude (in Anlehnung an Helmke, 1993)

Zur Feststellung der Schulleistungen der „Zweitklässler“ in den Fächern Deutsch und Mathematik wurden im Verlauf des Schuljahres in den 8 Intensivschulen folgende Tests eingesetzt:

WRT 1+ Weingartener Grundwortschatz Rechtschreib-Test für erste und zweite Klassen (Birkel, 1995); Durchführung bis Mitte / Ende Februar 1999.

KNUSPEL-Lesetest (Marx, 1998) mit den Subtests 1 bis 4: Hörverstehen, Rekodierfähigkeit, Dekodierfähigkeit, Leseverstehen; Durchführung bis zum Schuljahresende Juli 1999.

### **II.3 Hauptuntersuchung Kohorte 3 (Klassenstufe 1 bis 4) Konzept, Versuchsplan und Stichprobenbeschreibung**

Versuchsschulen im Schulversuch „Schulanfang auf neuen Wegen“

Mit Beginn des Schuljahres 1999/2000 begannen die Untersuchungen mit den Kindern, die im Herbst 1999 eingeschult wurden. Insgesamt 41 Schulen wurden mit Beginn dieses Schuljahres in die wissenschaftlich begleitete Untersuchung einbezogen. Es erprobten

16 Schulen die Modellvariante A1,  
9 Schulen die Modellvariante A2,  
8 Schulen das Modell B,  
8 Schulen nehmen als Vergleichsschulen an dem Versuch teil.

Mit Beginn des Schuljahres 2000/2001 nahmen weiterhin dieselben Schulen an der sog. Hauptuntersuchung bis zum Schuljahresende 2002/2003 teil.

**Tabelle 2: Beschreibung der Analysestichproben**

<b>Schülerzahlen (Kohorte 3)</b>				
<b>Erhebungen zum kognitiven Entwicklungsstand der Schüler, ihren schulischen Leistungen (WRT; WLLP, Hamlet 3-4, DEMAT , Noten Deutsch und Mathematik, Bildungsempfehlung, Motivation/Lernfreude, Sozialverhalten in Gruppen)</b>				
<b>Schuljahre</b>	<b>Schüler/innen</b>			
	Gesamt <sup>4</sup>	A1- + A2-Erprobungsklassen	B-Erprobungsklassen	Klassen in Vergleichsschulen
Schj. 99/00	1040	564	190	286
÷	÷	÷	÷	÷
Schj. 02/03	1040	564	190	286

<sup>4</sup> Die Schülerzahlen sind auf eine gemeinsame Grundgesamtheit von Schülern hin ausgerichtet, von denen in Kl. 1 bis Kl. 4 Daten vorliegen.

<b>Lehrerbefragungen (Kohorte 3)</b>					
<b>Befragung zur Unterrichtsgestaltung</b>					
<i>Stichprobe der Kohorte 3, 41 Schulen</i>					
<b>Schuljahre</b>	<b>Fragebogen zur Unterrichtsgestaltung</b>				
	Gesamt	A1-Erprobungs- klassen	A2-Erprobungs- klassen	B-Erprobungsklassen	Klassen in Ver- gleichsschulen
Schj. 99/00	70	25	18	11	16
Schj. 00/01	68	24	19	10	15
Schj. 01/02	55	18	12	10	15
Schj. 02/03	50	16	11	9	14

## Versuchsplan

Folgende Erhebungen wurden im Verlauf der Klassen 1 bis 4 in den 41 Stichprobenschulen durchgeführt:

### **Individualerhebungen bei den Schülerinnen und Schülern**

#### Intensiverhebungen:

T1: Beginn 1. Klasse (Oktober 1999) (75 min.)

T2: Beginn 2. Klasse (Oktober 2000) (45 min.)

#### Schulleistungstests in Deutsch (Lesen, Rechtschreiben)

2. Klasse: WRT 1+<sup>5</sup>, WLLP

3. Klasse: Diagnosearbeiten Deutsch Grundschule

4. Klasse: WRT 3+, WLLP, Hamlet 3-4 (Leseverständnistest Teil 1)

#### Schulleistungstests in Mathematik

2. Klasse: DEMAT 1+<sup>6</sup>

3. Klasse: DEMAT 2+, Diagnosearbeiten Mathematik Grundschule

4. Klasse: DEMAT 4<sup>7</sup>

#### Zeugnisnoten in den Fächern Deutsch und Mathematik, jeweils im Versetzungszeugnis am Ende der 2., 3. und 4. Klasse

#### Grundschul- oder Bildungsempfehlung

In Klasse 4 ausgesprochene Empfehlung für eine weiterführende Schulart

#### Schülerbefragung zum Sozialverhalten in der Gruppe (SVG)

1. Klasse: SVG 3.1 (Januar 2000), SVG 3.2 (Juli 2000)

2. Klasse: SVG 2.1 (Juni / Juli 2001)

4. Klasse: SVG 4.1 (Juni / Juli 2003)

### **Befragung der Erprobungslehrerinnen und -lehrer**

Fragebogen zur Unterrichtsgestaltung, differenziert nach den Modelltypen A, B sowie Vergleichsschule, in B-Modellen mit Grundschulförderklassen kamen zum Zeitpunkt T1 zusätzlich B1- und B2-Fragebogen zur Unterrichtsgestaltung der Grundschulförderklassen zum Einsatz

T1: Ende 1. Klasse (Juni / Juli 2000)

T2: Ende 2. Klasse (Juni / Juli 2001)

T3: Ende 3. Klasse (Juni / Juli 2002)

T4: Ende 4. Klasse (Juni / Juli 2003)

#### Bogen Schülerbeobachtung in den ersten Grundschulklassen (BSG-L)

1 Klasse (Juli 2000)

---

<sup>5</sup> Es bedeuten: WRT = Weingartener Grundwortschatz Rechtschreib-Test für erste, zweite, dritte und vierte Klassen, WLLP = Würzburger Leise Leseprobe; SVG = Erhebung zum Sozialverhalten in der Gruppe

<sup>6</sup> DEMAT = Deutscher Mathematiktest für erste, zweite, dritte und vierte Klassen

<sup>7</sup> Der Anfang der 4. Klasse durchgeführte Test DEMAT 3+ bleibt hier unberücksichtigt (Näheres s. S. 71).

### **Fragebogen zur Schulzufriedenheit der Eltern**

Der Fragebogen wurde im Juli 2001 – in Ausnahmefällen im September 2001 – bei den Eltern der sog. Kohorte 3 eingesetzt (Einschulungstermin: September 1999). Die Kinder befanden sich zum Zeitpunkt der Erhebung am Ende der Klasse 2 (T: vor den Sommerferien) oder am Anfang der Klasse 3 (T: nach den Sommerferien).

### **Schulverlaufsdaten**

Im Juni / Juli 2003 wurden von den 41 Schulen der Hauptuntersuchung Angaben zum Schulverlauf der Kohorte 3 (Einschulungstermin: September 1999) erfragt. Im Einzelnen betraf dies Angaben zum

- Schulbesuch von ..... bis
- Verweildauer in der Eingangsstufe
- Vorzeitiger Abgang von der Schule (z.B. Wegzug, Förderschulbesuch)
- Überspringen von Klasse(n)
- Wiederholen von Klasse(n)
- Grundschul-, Bildungsempfehlung und Jahresendnoten in Klasse 4 in Deutsch und Mathematik.

### Überblick zu den bei den beiden Intensiverhebungen verwendeten Testverfahren:

Im Herbst 1999 (ab 11. Oktober bis 10. November) wurde bei den Schulanfängern der 41 Stichprobenschulen die Individualerhebung (Eingangsdagnostik) durchgeführt.

Folgende Verfahren kamen dabei zum Einsatz:

Kaufman-Assessment-Battery for Children (K-ABC; Deutsche Version: Melchers & Preuß, 1993). Daraus der

- Subtest 3: Handbewegungen
- Subtest 5: Zahlen nachsprechen
- Subtest 9: Räumliches Gedächtnis

### Lese- und Schreibtests:

- Sichtwörter
- Lesen-Buchstabenkenntnisse, mit Zusatztest: Frühleser
- Schreibkenntnisse

### Testbatterie Menge und Zahl (TMZ; Winkelmann, 1976):

- Subtest: Zahlabstraktion (ZA)
- Subtest: Zählen von Objekten (ZO)



Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten - BISC: (Jansen / Mannhaupt / Marx & Skowronek, 1996 bzw. 1999). Daraus der

Subtest 2: Reimen (RP)

Subtest 8: Silben segmentieren (SS)

Lernfreude (in Anlehnung an Helmke, 1993)

Kunstwort-Nachsprech-Test (Hasselhorn & Körner, 1997)

Im Herbst 2000 (Oktober / November) kamen bei der Individualerhebung (Eingangsdagnostik) der Schüler, nun in Klasse 2, an den 41 Stichprobenschulen folgende Verfahren zum Einsatz:

Kaufman-Assessment-Battery for Children (K-ABC; Deutsche Version: Melchers & Preuß, 1993). Daraus der

Subtest 3: Handbewegungen

Subtest 5: Zahlen nachsprechen

Subtest 9: Räumliches Gedächtnis

Testbatterie Menge und Zahl (TMZ; Winkelmann, 1976):

Subtest: Zahlabstraktion (ZA)

Subtest: Zählen von Objekten (ZO)

Lernfreude (in Anlehnung an Helmke, 1993)

Kunstwort-Nachsprech-Test (Hasselhorn & Körner, 1997)

Bilder benennen (Stuttgarter Version)

Phonologische Bewusstheit (Marx & Schneider, 2000)

Subtest 1: Pseudowort-Segmentierung

Subtest 2: Vokalersetzung

Subtest 3: Restwortbestimmung

Subtest 4: Phonemvertauschung

## **II.4 Schülerverhalten aus Sicht der Lehrerinnen und Lehrer (Kohorte 1)**

### Bogen zur Schülerbeobachtung in den ersten Grundschulklassen (BSG-L)

#### *Einführung*

#### Ausgangslage

Für die wissenschaftliche Begleitung der Modellversuche „Schulanfang auf neuen Wegen“ ist ein wichtiges Untersuchungsziel, Aussagen über die schulische Leistungsentwicklung bzw. Förderung der schulischen Leistungsfähigkeit von Schülern machen zu können und in einem weiteren Schritt mögliche Abhängigkeiten der Fördereffekte von den jeweiligen Modellvarianten aufzuzeigen. Dazu ist es nötig, möglichst breite Informationen über den einzelnen Schüler zu sammeln. Aussagen über die schulische Leistungsfähigkeit eines Schülers können aus dessen Notenbild entnommen werden. Dies ist jedoch bei Schülern des ersten Schuljahres nicht möglich, weil für sie nur ein sog. Schulbericht vorliegt. Die dort enthaltenen Informationen über den einzelnen Schüler sind bedauerlicherweise aus zwei Gründen nicht für die Zwecke der Begleituntersuchung geeignet:

1. Für eine Weiterverarbeitung müssten die Informationen in eine quantifizierbare Form umcodiert werden. Dieses war mit den Mitteln der Begleituntersuchung nicht zu leisten.
2. Die Schulberichte konzentrieren sich auf bestimmte, den Lehrerinnen/Lehrern wichtige Bereiche, sie sind aber nicht systematisch in dem Sinne, dass sie den Schüler nach einem festgelegten Raster von Dimensionen beschreiben.

Daher konnte es keinen anderen Weg geben, als ein Instrument einzusetzen, das diese systematische und quantitative Schülerbeschreibung ermöglichte.

Auf Grund der Fragestellungen, die aus der Untersuchungskonzeption abgeleitet wurden, und nach der Durchsicht der veröffentlichten Verfahren, entschied sich der „Arbeitskreis Wissenschaftliche Begleituntersuchung“, ein eigenes Verfahren zu konstruieren. Die Endform des Bogens zur Schülerbeobachtung in den ersten Grundschulklassen (BSG-L) setzt sich aus 52 Fragen zu folgenden Bereichen (Skalen) zusammen:

- Feinmotorik (4 Fragen)
- Grobmotorik (4 Fragen)
- Hypermotorik (4 Fragen)
- Emotionale Selbststeuerung (6 Fragen)
- Konzentration (4 Fragen)
- Arbeitsverhalten und Lernmotivation (11 Fragen)
- Instruktionsverständnis (4 Fragen)
- Phantasie und Kreativität (5 Fragen)
- Sozialverhalten (10 Fragen)

Zu jeder Frage war eine 5-stufige Skala vorgegeben (Hypermotorik 4-stufig), anhand derer die Klassenlehrer jeden einzelnen ihrer Schüler beschreiben sollten.

Diesem Zwischenbericht ist ein Materialienband beigelegt. Diesem sind alle Einzelheiten über Aufbau und Testanalysedaten im Bericht „Bogen zur Schülerbeobachtung in den ersten Grundschulklassen (BSG-L) - Konstruktion und Testanalyse“ zu entnehmen.

### *Testanalyse des Schülerbeobachtungsbogens*

#### Vorbemerkung

Für die im vorliegenden Bericht dargestellten teststatistischen Analysen wurden die Antworten so kodiert, dass ein höherer Zahlenwert prinzipiell auch einen günstigeren Entwicklungsstand repräsentiert. Dies gilt auch für den Bereich „Hypermotorik“, d.h., je höher der entsprechende Skalenwert ausfällt, umso weniger hypermotorisch ist das Kind.

Die Analysen basieren auf den Beobachtungen von 142 Lehrern aus 65 Schulen zu insgesamt 2812 Schülern. Allerdings lagen zu den einzelnen Fragen (Items) unterschiedlich viele Beobachtungen vor, sodass sich die Berechnungen auf eine zum Teil stark schwankende Anzahl von Fällen (N) beziehen.

#### Verteilung der Skalen-Rohpunktwerte

In den Abbildungen, die im Materialienband dargestellt werden, wird die Verteilung der Skalen-Rohwerte graphisch als Histogramm mit Normalverteilungskurve dargestellt. Es wird deutlich, dass die einzelnen Skalen mehr oder weniger von der Normalverteilung abweichen. Alle Verteilungen sind rechtsschief, d.h., die Mehrheit der Schüler erreicht höhere Skalenwerte, was für einen altersgemäßen Entwicklungsstand der Schülerschaft spricht. Wenn man nachvollziehen möchte, wie die Grundschullehrerinnen/-lehrer derzeit die Erstklassenschüler sehen, z.B. auch, wie hoch der Prozentsatz der von ihnen als entwicklungsretardiert eingeschätzten ist, bezogen auf die einzelnen Items oder die Skalen, der sei auf den Materialienband zum Zwischenbericht 2000 verwiesen.

#### Angaben zur Trennschärfe und Reliabilität der Skalen

Es ist aus der Angabe der Bandbreite der Trennschärfekoeffizienten der einzelnen Items (Beobachtungen) zu sehen, dass sie sich in Bezug auf die zugehörigen Skalen (Beobachtungsbereiche) als z. T. ungewöhnlich hoch erweisen. Man kann also davon ausgehen, dass das von uns gewählte Verfahren, differenzierte Einstufungsmöglichkeiten vorzugeben, eine zuverlässige Erfassung der entsprechenden Schülermerkmale erlaubt. Dementsprechend hoch fallen auch die Reliabilitätskoeffizienten der Skalen selbst aus. Dies ist bei Skalen mit teilweise so geringer Itemzahl bemerkenswert (siehe Tabellenband). Die hohe innere Konsistenz der Messung kann als Qualitätsmerkmal des Verfahrens gewertet werden.

#### Faktorenanalyse der Items und Interkorrelationen der Skalen

Allerdings werfen einzelne Ergebnissen der Testanalyse die Frage auf, wie diese hohen Werte möglicherweise zu erklären sind. Denn wenn man in die Betrachtung die Ergebnisse der Korrelationsberechnungen (Interkorrelationen der Skalen) mit einbezieht<sup>8</sup>, so finden sich für die hohen Werte folgende Erklärungsansätze:

---

<sup>8</sup> Details siehe Bericht im Tabellenband.

Der Entwicklungsstand der Kinder ist als „ganzheitlich“ zu betrachten. D.h., Kinder, die in einem Bereich einen hohen Entwicklungsstand aufweisen, weisen diesen in der Regel auch in den anderen Bereichen auf. Gleiches gilt für die Kinder mit Entwicklungsrückständen. Die Lehrer neigen dazu, das einzelne Kind im Sinne eines so genannten Hofeffekts „ganzheitlich“ wahrzunehmen. D.h., wenn der Lehrer von dem Kind einen guten Eindruck in einem bestimmten Bereich hat, so überstrahlt dies auch die anderen Bereiche, obwohl das Kind bei objektiver Messung hier schlechter abschneiden würde. Gleiches gilt umgekehrt natürlich auch für negative Eindrücke.

Dass z.B. das *Instruktionsverständnis* hoch mit *Konzentration* (.72) und *Arbeitsverhalten* (.82) korreliert, ist nachvollziehbar. Bei der Korrelation zu *Phantasie und Kreativität* (.77) fällt einem eine Erklärung dagegen schon schwerer.

In diesem Zusammenhang sind auch die Ergebnisse der Faktorenanalyse<sup>9</sup> interessant. So werden z.B. die Items der Skalen *Hypermotorik* und *Emotionale Selbststeuerung* in einem gemeinsamen Faktor abgebildet. Gleiches gilt für die Skalen *Konzentration* und *Arbeitsverhalten*. Auf der anderen Seite werden aber dem Faktor, der von den Items der Skala *Feinmotorik* geprägt ist (M1 bis M4), auch Items wie: „Kann das Kind bestimmte Bastelarbeiten oder graphomotorische Figuren nach 1-2-maligem Vormachen richtig nachmachen?“ (*Instruktionsverständnis*, I3) und: „Hat das Kind beim Malen und Zeichnen viele Einfälle?“ (*Phantasie und Kreativität*, PK4), zugeordnet.

Die genannten Items werden dem Lehrer im Beobachtungsbogen unter der entsprechenden Bereichsüberschrift vorgegeben. Offensichtlich urteilen die Lehrer hier aber weniger nach der eigentlichen Intention der Items, sondern mehr nach ihrer technischen Ausführung.

Bei der Faktorenanalyse<sup>10</sup> handelt es sich um ein statistisches Verfahren, das zum Ziel hat, eine Vielzahl von Informationen (Variablen) zu einigen wenigen prägnanten Dimensionen (Faktoren) zu verdichten. Dabei spielen die Korrelationen der Variablen untereinander im Sinne eines „Ähnlichkeitsmaßes“ eine wichtige Rolle.

Im vorliegenden Fall wurde zunächst eine Faktorenanalyse über alle 52 Items (Beobachtungen) gerechnet. Dabei sollte geprüft werden, inwieweit sich die einzelnen Beobachtungsbereiche (Skalen) im Datenmaterial wiederfinden lassen.

Diese Faktorenanalyse, die sich nach dem Eigenwertkriterium richtet<sup>11</sup>, zeigt, dass dies in etwa gelungen ist. Die Faktorenanalyse findet sechs Dimensionen, die weitgehend deckungsgleich mit den eingangs angeführten Beobachtungsbereichen sind und sich gut interpretieren lassen. Die im Fuß der Tabelle angegebene „Erklärte Gesamtvarianz“ zeigt darüber hinaus auf, dass die gefundenen Faktoren die zu Grunde liegende Datenstruktur zu fast 70 Prozent aufklären, ein Ergebnis, das als relativ gut zu bewerten ist.

### *Zur Validität des Bogens zur Schülerbeobachtung in den ersten Grundschulklassen (BSG-L)*

Wenn ein Testverfahren entwickelt wird, muss untersucht werden, ob es auch tatsächlich das misst, was es zu messen vorgibt.

---

<sup>9</sup> Eine Beschreibung dieses Verfahrens findet sich im nächsten Absatz. Details zu den Einzelergebnissen siehe Bericht im Materialienband zum Zwischenbericht 2000.

<sup>10</sup> Extraktionsmethode: Hauptkomponentenanalyse  
Rotationsmethode: Varimax mit Kaiser-Normalisierung

<sup>11</sup> Anzahl der Faktoren: Eigenwert > 1

Bei den Schülern der sog. Intensivschulen aus der Individualtestung des Entwicklungsstands der Schüler zu Beginn des ersten Schuljahres liegen eine Reihe von Verfahren zur Sicherung der sog. empirischen Validität vor.

Ebenso sind von diesen Schülern die Noten in den Fächern Deutsch und Mathematik vom Ende des zweiten Schuljahres bekannt, darüber hinaus noch Ergebnisse aus zwei standardisierten Schulleistungstests: „WRT 1 + Weingartener Grundwortschatz Rechtschreib-Test, Birkel, 1995“) und „KNUSPEL-L“ (Lesetest, Marx, 1998).

### Zusammenhänge zwischen Testergebnissen aus der sog. Individualtestung<sup>12</sup> und Skalen des BSG-L<sup>13</sup> (siehe Tabellen im Tabellenband)

Es gibt eine Reihe von aufgezeigten Zusammenhängen zwischen einzelnen Skalen und einzelnen Testscores, die unmittelbar einleuchten:

Die hohen Zusammenhänge zwischen der BSG-L -Skala Feinmotorik und den Testscores von „Wörter schreiben“, Graphesthesia Test, Vornamen schreiben und Graphomotorischer Test.

Die hohen Zusammenhänge zwischen den Skalen Konzentration, Arbeitsverhalten und Lernmotivation, Instruktionsverständnis und Phantasie mit z.B. den Testscores von „Pseudowörter nachsprechen“, „Zahlennachsprechen“, „Zahlenabstraktion“ und „Reimen“ etc.

Die plausiblen Zusammenhänge zwischen der BSG-L -Skala Emotionale Selbststeuerung und den Testscores von „Räumliches Gedächtnis“ und „Handbewegungen“, die im Sinne von „reflexives vs. impulsives Denken“ zu verstehen sind.

Ebenfalls sehr plausibel ist der insgesamt niedrige Zusammenhang zwischen den Testscores und den BSG-L -Skalen Grobmotorik, Emotionale Selbststeuerung, Hypermotorik und Sozialverhalten.

Aufmerksamkeit verdienen die hohen Zusammenhänge zwischen dem Test Bielefelder Screening Test - BISC (Risikowert) und bestimmten Skalen des BSG-L. Die durch den BISC als gefährdet ausgewiesenen Schüler, was den Lese-Schreibprozess angeht, werden indirekt durch die Lehrerinnen/Lehrer so charakterisiert:

Instruktionsverständnis (niedrig), Arbeitsverhalten und Lernmotivation (schlecht), Konzentration (schlecht/gestört), Feinmotorik (gering entwickelt), und auch noch Sozialverhalten (weniger gut) und emotionale Selbststeuerung (eher gering ausgeprägt).

Eine Reihe anderer Zusammenhänge sind weniger plausibel (z.B. mit der BSG-L -Skala Grobmotorik); dies kann zum Teil durch die Tendenz zur eher globalen Schülerbeurteilung begründet werden.

### *Zusammenhänge zwischen den Skalen des BSG-L, den Jahresendnoten Klasse 2 und standardisierten Schulleistungstests*

### Zusammenhänge zwischen den Skalen des BSG-L und den Jahresendnoten in Deutsch und in Mathematik, Klasse 2

Am Ende von Klasse 2 liegen mit den Noten in Deutsch und Mathematik zum ersten Mal im Zeugnis eines Schülers der Eingangsstufe numerische Leistungsbeurteilungen vor.

<sup>12</sup> Zur Individualtestung siehe Kap. II.5 Entwicklungsverläufe.

<sup>13</sup> Sortiert nach den jeweils vier höchsten Koeffizienten, weitere Details siehe Tabellenband.

Es ist daher bei den acht Intensivschulen möglich, Zusammenhänge zwischen den Skalenwerten des BSG-L und den Noten der beiden Fächer herzustellen. Wir bedienen uns dazu des Verfahrens der multiplen Regression<sup>14</sup>, wobei die sog. Einflussvariablen die 9 BSG-L - Skalen und die abhängigen jeweils die Noten in Deutsch und Mathematik sind.

Das Ergebnis weist einen sehr engen Zusammenhang zwischen der Schülerbeurteilung der Lehrerinnen/Lehrer mittels des standardisierten BSG-L in Klasse 1 (etwa im 4. Quartal des ersten Schuljahres vorgenommen) und den Jahresendnoten Kl. 2 auf. Die entsprechenden Koeffizienten betragen  $R_{\text{mult}} = .780$  ( $R^2_{\text{corr}} = .49$ ) für Mathematik und  $R_{\text{mult}} = .798$  ( $R^2_{\text{corr}} = .52$ ) für Deutsch (Tabelle siehe Tabellenband). Nahe liegende Schlüsse zur Konstanz des relativen Rangplatzes von Schülern in der Eingangsstufe der Grundschule sind leider durch die relativ geringe Fallzahl nicht zu belegen, sie können aber als Grundlage einer noch zu bestätigenden Hypothese herangezogen werden.

#### Zusammenhänge zwischen den Skalen des BSG-L und den standardisierten Schulleistungstests (WRT 1+Weingartener Rechtschreibtest) und KNUSPEL-Lesetest

Das Ergebnis weist einen sehr engen Zusammenhang zwischen der Schülerbeurteilung der Lehrerinnen/Lehrer mittels des standardisierten BSG-L in Klasse 1 (etwa im 4. Quartal des ersten Schuljahres vorgenommen) und den beiden standardisierten Schulleistungstests auf, die am Ende der Klasse 2 eingesetzt wurden (WRT  $R_{\text{mult}} = .712$ ,  $R^2_{\text{corr}} = .37$ ; KNUSPEL Leseverstehen  $R_{\text{mult}} = .672$ ,  $R^2_{\text{corr}} = .29$ ; KNUSPEL Hörverstehen  $R_{\text{mult}} = .854$ ,  $R^2_{\text{corr}} = .65$ ). Natürlich ist auch hier die Fallzahl zu klein, um aus den aufgezeigten Beziehungen weit reichende Schlüsse zu ziehen. Dennoch kann man besonders die hohe Varianzaufklärung beim KNUSPEL-Untertest Hörverstehen so verstehen, dass die Schülerbeobachtung durch die Lehrerinnen/Lehrer der Eingangsstufe der Grundschule hoch valide sein kann. Voraussetzung ist allerdings, dass diese Schülerbeobachtung anhand eines elaborierten und systematischen Verfahrens geschieht.

#### *Modellspezifische Ergebnisse (nur Intensivschulen)*

In den vergangenen Kapiteln wurde aufgezeigt, dass der BSG-L ein zuverlässiges (reliables) und gültiges (valides) Instrument darstellt. Somit ist es auch möglich, dieses Verfahren zu Vergleichszwecken einzusetzen, um eventuell modellspezifische Besonderheiten erkennen zu können.

Allerdings können aus den Ergebnissen keine weit reichenden Schlüsse gezogen werden. Dies aus folgenden Gründen:  
Die Fallzahlen sind zu klein.

Damit verknüpft ist die Frage, die nicht entschieden werden kann, ob eine Kongruenz der Schülerstichproben gegeben ist.

Bestimmte Verhaltensweisen, z.B. das Sozialverhalten, können bei den Lehrerinnen / Lehrern der Erprobungsklassen möglicherweise stärker im Fokus ihrer Aufmerksamkeit stehen, weil sie eng mit modellspezifischen Elementen verknüpft sind. Die Vermutung ist:

---

<sup>14</sup> Die Ergebniswerte werden durch zwei Koeffizienten ausgedrückt:  $R_{\text{mult}}$ , d.h. multipler Korrelationskoeffizient und  $R^2_{\text{corr}}$ , d.h. stichprobenkorrigierte Angabe der aufgeklärten Varianz (ein Wert von 1.00 bedeutet maximal aufgeklärte Varianz, 0.00 keinerlei aufgeklärte Varianz).

Weil die Lehrerinnen / Lehrer der Erprobungsklassen auf diese beiden Bereiche ihr Augenmerk und ihr Engagement richten, verändert sich ihre Wahrnehmung zum Positiven, da sie auf diesem indirektem Weg ihre Anstrengungen belohnt sehen. Daher kann es sein, dass aufgezeigte Unterschiede mehr durch die einschätzenden Personen und weniger durch die tatsächlichen Schülerunterschiede bestimmt sind.

Es sind relativ wenig signifikante Ergebnisse zwischen den Erprobungsklassen und den Vergleichsschulen aufzuzeigen (siehe Tabelle im Tabellenband). Lediglich in den beiden Skalen Phantasie und Sozialverhalten werden die Schüler der Erprobungsklassen mit einem jeweils höheren Mittelwert beurteilt. Es wäre von den Intentionen der Modellansätze plausibel, dass sich hier erste Förderungserfolge abzeichnen; genauso gut kann es sich um einen Prozess selektiver Wahrnehmung handeln, wie schon weiter oben als Vermutung geäußert wurde.

Die Folgerung ist, dass die Ergebnisse zunächst nur zur Bildung von Hypothesen für den Endbericht herangezogen werden sollten.

### *Zusammenfassung*

Der Bogen zur Schülerbeobachtung (BSG-L) ist konzipiert worden, um auch von den Schülerinnen und Schülern der Klassenstufe 1 eine umfassende Beschreibung zu erhalten, da eine Auswertung der sog. Schulberichte im Rahmen dieser Begleituntersuchung nicht möglich ist.

Die Endform des „Bogens zur Schülerbeobachtung für Schüler der ersten Grundschulklassen“ (BSG-L) setzt sich aus 52 Fragen zu 9 Skalen zusammen. Zu jeder Frage war eine 5- bzw. 4-stufige Skala vorgegeben, anhand derer die Klassenlehrer jeden Einzelnen ihrer Schüler beschreiben sollten.

Die umfangreichen Testanalysen konnten zeigen, dass es sich bei diesem Verfahren um ein zuverlässiges (reliables) und gültiges (valides) Instrument handelt, das auch eine Fülle interessanter Ergebnisse über die Schülersituation (z.B. Anteil der als hypermotorisch eingestuften Schüler) in der Grundschule erbringt (siehe Testbericht im Materialienband, Zwischenbericht 2000).

Vergleicht man die Einschätzungen der Schüler aus den verschiedenen Modellklassen miteinander, so sind relativ wenig signifikante Ergebnisse zwischen den Erprobungs- und den Vergleichsklassen aufzuzeigen (siehe Tabellenband). Lediglich in den beiden Skalen Phantasie und Sozialverhalten werden die Schüler der Erprobungsklassen jeweils günstiger beurteilt. Es kann aufgrund der vorliegenden Datenbasis nicht entschieden werden, ob es sich um die Wahrnehmung erster Fördererfolge handelt oder um eine Art „self-fulfilling prophecy“, weil beide Skalen zentrale Bereiche der Erprobungsansätze erfassen. Auch ist die Fallzahl zu gering, als dass daraus Schlüsse gezogen werden können.

### *Ausblick*

Leider ist die Bearbeitung des BSG-L für die Lehrerinnen und Lehrer mit einem erheblichen Aufwand verbunden, mit der Folge, dass von einer Reihe von Schulen keine Bearbeitung erfolgte. Aus diesem Grund ist eine Kurzfassung erarbeitet worden, die schon im Schuljahr 1999/2000 eingesetzt wurde. Ob mit dieser Kurzfassung ähnlich überzeugende Werte bezüglich der Testkriterien erzielt werden können, werden zukünftige Analysen zeigen.

## II.5 Entwicklungsverläufe

### Vorbemerkungen

Das Ziel der Verminderung des im Laufe der letzten Jahrzehnte stark angestiegenen durchschnittlichen Schuleintrittsalters soll im Rahmen des Projektes "Schulanfang auf neuen Wegen" vor allem durch zwei Kategorien von Maßnahmen erreicht werden: zum einen durch einen Verzicht auf die bisherige Rückstellungspraxis bei gleichzeitiger Flexibilisierung der Verweildauer in der Schuleingangsstufe (evtl. auch des Einschulungszeitpunktes); zum anderen durch eine Integration der bisherigen Klassenstufen 1 und 2 bei gleichzeitiger pädagogischer bzw. didaktisch-methodischer Weiterentwicklung eines differenzierten Anfangsunterrichts.

Auch die wissenschaftliche Begleitevaluation zu diesem Projekt verfolgt vor allem zwei Zielsetzungen. Einerseits soll geprüft werden, welche Konsequenzen die veränderte Schuleingangsstufe im Vergleich zur bisherigen Beschulungspraxis für den Entwicklungsverlauf der betroffenen Kinder hat. Andererseits sollen mögliche Schwachstellen des Projektansatzes identifiziert und möglichst zeitnah rückgemeldet werden, wobei auch hierbei der individuelle Lern- und Entwicklungsverlauf der Kinder die Bewertungsgrundlage bildet.

In einer im Herbst 1997 begonnenen Intensiverhebung wurde eine Pilotstudie realisiert, bei der es zum einen um die Erprobung geeigneter Erhebungsverfahren ging, zum anderen aber bereits auch um einen ersten Vergleich der Entwicklungsverläufe von Kindern, die in herkömmlichen Klassenstrukturen unterrichtet wurden, und solchen, die in den kombinierten Klassen des A-Modells unterrichtet wurden.

Die in der Intensiverhebung realisierten "abhängigen Variablen" wurden so ausgewählt, dass ein möglichst differenziertes Bild der kindlichen Entwicklung mit einem Schwerpunkt auf den besonders schulrelevanten Entwicklungsbereichen nachgezeichnet werden kann.

Im Folgenden werden die wichtigsten Ergebnisse zu den individuellen Entwicklungsverläufen der Kinder der Voruntersuchung in Abhängigkeit davon vorgestellt, ob sie in einer Erprobungs- oder einer Vergleichsklasse unterrichtet wurden. Die Verfahren zur Erfassung von Entwicklungsaspekten wurden bei den 1997 eingeschulerten Kindern zum einen in den ersten Wochen ihres ersten Schuljahres und dann etwa ein Jahr später zu Beginn ihres zweiten Schuljahres in intensiven Einzeluntersuchungen durchgeführt. Die folgenden Auswertungen beziehen sich auf diese Daten der Intensiverhebung, die von insgesamt 219 Kindern weitgehend vollständig vorliegen. 97 dieser Kinder besuchten eine Erprobungsklasse im A-Modell des „Schulanfangs auf neuen Wegen“, 122 besuchten traditionelle Vergleichsklassen, die sich entweder an den gleichen Schulen oder aber an Grundschulen ohne Erprobungsklassen befanden.

### Kognitive Entwicklung

Ein wichtiges Ziel des Schuleingangsunterrichts ist die Vermittlung basaler Fertigkeiten des Schreibens, Lesens und Rechnens. Von daher ist es nicht verwunderlich, dass kognitiven Entwicklungsmerkmalen und Fertigkeiten eine große Bedeutung für den Schulerfolg eingeräumt wird. Die neuere Grundlagenforschung zum Schriftspracherwerb und zum Erwerb rechnerischer Fertigkeiten bietet darüber hinaus ein differenziertes Bild über die relative Bedeutung einzelner kognitiver Funktionsmerkmale beim Erwerb des Lesens (vgl. Scheerer-Neumann, 1997), Schreibens (vgl. Schneider, 1997) und Rechnens bei Kindern (Stern, 1997). An den Ergebnissen dieser Grundlagenforschung orientierte sich die Auswahl der empirisch erfassten kognitiven Funktionsbereiche.



Neuere Metaanalysen zur Schulerfolgsprognose bestätigen die herausragende Bedeutung kognitiver Kompetenzen und Merkmale für eine erfolgreiche schulische Entwicklung (vgl. Wang, Haertel & Walberg, 1993; Helmke & Weinert, 1997). Für das Projekt "Schulanfang auf neuen Wegen" lässt sich daraus das Ziel ableiten, dass sich die zu erprobenden Veränderungen des Anfangsunterrichts nicht negativ auf relevante kognitive Funktionsbereiche auswirken dürfen.

Im Gegensatz zu früheren psychometrisch orientierten Ansätzen wählten wir bei der Erfassung der in ihrer Entwicklung zu beobachtenden kognitiven Merkmale nicht eine allgemeine Kompetenzebene (z. B. Intelligenz), sondern die Ebene näher spezifizierbarer kognitiver Funktionen, die sich für die schulische Entwicklung als besonders relevant erwiesen haben. Nach Sichtung der einschlägigen Literatur fiel unsere Wahl auf die kognitiven Funktionsbereiche phonologische Bewusstheit, Aufmerksamkeit, (phonologisches) Arbeitsgedächtnis und Denken.

### Phonologische Bewusstheit und Aufmerksamkeit

Schon in den 60-er Jahren hatte Kossakowski (1961) die Beobachtung gemacht, dass Kinder mit Problemen beim Schriftspracherwerb auch im Bereich der Phonemanalyse Minderleistungen zeigen. Die Rolle der Einsicht von Kindern in die Phonologie der Sprache ist seit Ende der 70er Jahre unter dem Terminus "phonologische Bewusstheit" näher untersucht worden. Mit Skowronek und Marx (1989) lassen sich dabei zwei verschiedene Kategorien der phonologischen Bewusstheit voneinander unterscheiden. Unter phonologischer Bewusstheit im weiteren Sinne versteht man die Kenntnis und den Umgang mit den lautlichen Aspekten der Sprechsprache (z.B. das Finden von Reimwörtern, die Gliederung von Wörtern nach Sprechsilben). Die phonologische Bewusstheit im engeren Sinne kommt hingegen in sprachlichen Leistungen zum Ausdruck, bei denen explizit mit lautlichen Strukturen operiert werden muss, die weder semantische noch sprechrhythmische Bezüge aufweisen (z. B. das Erkennen von Anfangs- bzw. Endlauten eines Wortes).

Mittlerweile hat sich in Längsschnittstudien die Hypothese empirisch bestätigen lassen, dass Indikatoren der phonologischen Bewusstheit besonders gut geeignet sind, um spätere Lese- / Rechtschreibleistungen vorherzusagen (vgl. Schneider, 1997). Insbesondere für den Schriftspracherwerb sind daher die bei der Einschulung vorhandene phonologische Bewusstheit und ihre Fortentwicklung im Verlauf des ersten Schuljahres von großer Bedeutung. Für den „Schulanfang auf neuen Wegen“ überprüften wir daher, ob zwischen Erprobungsklassen und Vergleichsklassen der Voruntersuchung Unterschiede in den Eingangskompetenzen zur phonologischen Bewusstheit bestehen und vor allem, ob sich im Laufe des ersten Schuljahres Unterschiede in der Fortentwicklung zeigen.

Die Eingangsvoraussetzungen wurden in Individualuntersuchungen mit dem Bielefelder Screening BISC zur Früherkennung von Lese-Rechtschreibschwierigkeiten (Jansen, Mannhaupt, Marx & Skowronek, 1999) erfasst, das u.a. auch Subtests zur Erfassung der phonologischen Bewusstheit im engeren (Laute assoziieren, Laut-zu-Wort Zuordnung) wie im weiteren Sinne (Reimen, Silben segmentieren) enthält.

Neben Aufgaben zur phonologischen Bewusstheit enthält das Bielefelder Screening auch Subtests zur Erfassung von Aufmerksamkeits- und Gedächtnisfunktionen, die ebenfalls für die Zwecke der Evaluation des Projektes geeignet sind. Insbesondere die selektive Aufmerksamkeit hat sich in Metaanalysen als bester Frühindikator späterer Schulleistungsprobleme erwiesen (Horn & Packard, 1985). Und auch in der Bielefelder Längsschnittuntersuchung zur Früherkennung von Lese-Rechtschreibschwierigkeiten in der Grundschule erwiesen sich neben der phonologischen Bewusstheit besonders Variablen aus dem Aufmerksamkeits- und Gedächtnisbereich als prädiktionsstark (vgl. Marx, 1992). Wir hielten

es daher für sinnvoll, auch die Funktionstüchtigkeit im Bereich der Aufmerksamkeit zu betrachten. Zu diesem Zwecke wurde der BISC-Subtest "Wort-Vergleich-Suchaufgabe" (Aufmerksamkeit für visuelle Symbolfolgen) erhoben.

All diese Merkmale der phonologischen Bewusstheit und Aufmerksamkeit wurden kurz nach der Einschulung der Kinder unserer Voruntersuchung im Herbst 1997 in intensiven Einzeluntersuchungen erfasst. Um die Fortentwicklung der phonologischen Bewusstheit analysieren zu können, mussten ein Jahr später erneut Einzeluntersuchungen durchgeführt werden. Das Bielefelder Screening ist allerdings zu Beginn der zweiten Klasse nicht mehr einsetzbar. So führten wir stattdessen einige Subtests einer damals noch in der Entwicklung befindlichen Testbatterie zur Erfassung der phonologischen Bewusstheit im Grundschulalter von Marx und Schneider (2000) durch, bei denen Phonemvertauschungen, Restwortbestimmungen und Vokalersetzungen vorzunehmen waren.

Die Details der von uns mit diesen Daten durchgeführten statistischen Analysen sind dem Tabellenband zu entnehmen.

*Zusammengefasst ergeben sich folgende Befunde: Zu Beginn von Kl.1 bringen die Schüler der A-Modelle eher einen niedrigeren Entwicklungsstand mit als die der B-Modelle und der V-Schulen (dies zeigt das Bielefelder-Screening). Zu Beginn des zweiten Schuljahres, etwa im Herbst 2000, zeigten sich dann bei kovarianzanalytischen Berechnungen in zwei von vier Untertests der phonologischen Bewusstheit signifikante Unterschiede zwischen den A-Modellen und den Vergleichsschulen, die zu Lasten der A-Modelle ausfallen. Dabei muss allerdings der erheblich größere Ausländeranteil in den A-Modellen berücksichtigt werden.*

### (Phonologisches) Arbeitsgedächtnis

Die phonologische Bewusstheit ist jedoch nicht die einzige bedeutsame individuelle kognitive Voraussetzung für das Lernen von Lesen und Schreiben. Ähnlich bedeutsam scheint auch das sprachgebundene Kurzzeitgedächtnis bzw. das phonologische Arbeitsgedächtnis zu sein. Unter Kurzzeit- oder Arbeitsgedächtnis versteht man ein internes kognitives System, das es ermöglicht, mehrere Informationen vorübergehend zu speichern und miteinander in Beziehung zu setzen. Die funktional verfügbare Kapazität dieses Systems unterliegt im Kindesalter nicht nur deutlichen Entwicklungsveränderungen, sondern sie weist auch relativ stabile interindividuelle Unterschiede auf. Diese Unterschiede bleiben nicht ohne Auswirkungen auf den Schriftspracherwerb. Gerade bei Schulanfängern muss man sich den Lesevorgang so vorstellen, dass erhebliche Übersetzungsschritte der Buchstaben in ihre lautlichen Entsprechungen erforderlich sind. Da diese Übersetzungsschritte zunächst noch recht langsam ablaufen, müssen die lautsprachlichen Entsprechungen der schriftlichen Symbole im Arbeitsgedächtnis möglichst lange gespeichert werden, damit das Zusammenziehen von Einzellaute im nächsten Schritt (Rekodieren) möglichst gut gelingen kann. Verfügt ein Kind über ein nur schwaches Arbeitsgedächtnis, so hat es gerade bei der Rekodierung längerer Wörter am Schluss eines Wortes schon wieder vergessen, wie der Wortbeginn lautete.

Da die Funktionstüchtigkeit des Arbeitsgedächtnisses nicht nur eine entscheidende Entwicklungsdeterminante für den Schriftspracherwerb ist, sondern für jegliche Art von intentionalem Wissenserwerb (vgl. Gathercole & Baddeley, 1993; Hasselhorn & Schumann-Hengsteler, 1998), entschlossen wir uns dazu, bei der Evaluation des Projektes "Schulanfang auf neuen Wegen" dem Arbeitsgedächtnis besondere Aufmerksamkeit zu widmen. Einer der Subtests des zur Erfassung der phonologischen Bewusstheit eingesetzten BISC ist z.B. dem Bereich des Arbeitsgedächtnisses zugeordnet. Der Subtest „Pseudowörter

nachsprechen" misst die Qualität des phonetischen Rekodierens im Arbeitsgedächtnis. Da dieser Test vermutlich für viele Kinder zu Beginn der zweiten Klassenstufe zu leicht ist, führten wir ihn in unserer Voruntersuchung nur kurz nach Beginn der Einschulung im Herbst 1997 durch. Im Herbst 1998 ersetzten wir ihn durch den Test zum Kunstwörter nachsprechen von Hasselhorn und Körner (1997), der sehr ähnlich aufgebaut ist, von seinem Schwierigkeitsgrad jedoch bis zum Ende der Grundschulzeit einsetzbar ist. Eingesetzt wurden darüber hinaus drei Subtests der Kaufman-Assessment-Battery for Children (K-ABC von Melchers & Preuß, 1993), mit deren Hilfe eine recht differenzierte Entwicklungsdiagnostik des Arbeitsgedächtnisses erfolgen kann. Der Subtest „Handbewegungen" ermöglicht eine Abschätzung der Leistungsfähigkeit des sprachfreien Arbeitsgedächtnisses, der Subtest „Zahlen nachsprechen" gehört zu den klassischen Indikatoren des phonologischen Arbeitsgedächtnisses und der Subtest „Räumliches Gedächtnis" erlaubt eine Diagnose des visuell-räumlichen Arbeitsgedächtnisses, dem u.a. eine besondere Bedeutung bei der Bewältigung mathematischer Anforderungen zukommt.

Im Tabellenband werden die Details der statistischen Analysen aufgeführt, die Ergebnisse lassen sich folgendermaßen beschreiben:

Das Ausmaß des Entwicklungsfortschrittes der funktionalen Kapazität des Arbeitsgedächtnisses zwischen Herbst 1999 und Herbst 2000 schwankt von Modell zu Modell und von Untertest zu Untertest. Bei dem Kunstwort-Nachsprech-Test und dem K-ABC Untertest Räumliches Gedächtnis zeigen die B-Modellschüler den größten Entwicklungsfortschritt, beim K-ABC-Untertest Zahlennachsprechen gibt es nur zwischen den Schülern der Vergleichsschulen und denen der A-Modelle einen signifikanten Unterschied (V-Modell größer); beim K-ABC Untertest Handbewegungen ist die Entwicklung bei allen Modellen etwa gleich.

*Zusammengefasst gilt daher auch für die Entwicklung des Arbeitsgedächtnisses, dass die Modellvariante, also auch die jahrgangsübergreifende Beschulung im A-Modell des „Schulanfangs auf neuen Wegen“, keinerlei systematischen Einfluss auf die Entwicklung im Verlauf des ersten Schuljahres hat.*

## Denken

Schließlich ist das induktive Denken, also die Fähigkeit allgemeine Regeln oder Gesetzmäßigkeiten zu erkennen bzw. aus Einzelerfahrungen abzuleiten, von besonderer Bedeutung für schulisches Lernen (z. B. Klauer, 1993). Das induktive Denken gehört zu den Kerndimensionen der „Mechanik“ der allgemeinen Intelligenz und wird vor allem in Anforderungen des analogen Denkens erfasst. Im Gegensatz zu tradierten Überzeugungen weiß man heute, dass die Mechanik der Intelligenz nicht nur das Produkt genetisch determinierter Reifungsprozesse ist, sondern auch durch die Beschulung prinzipiell beeinflusst wird (vgl. Hasselhorn & Grube, 1997).

Es schien uns besonders wichtig, Facetten des mathematischen Denkens zu erfassen, von denen angenommen werden darf, dass sie wichtige Entwicklungsvoraussetzungen des Rechnen lernens darstellen. Nach Sichtung der einschlägigen Verfahren fiel unsere Wahl auf zwei Subtests der unveröffentlichten Testbatterie Menge und Zahl von Winkelmann (1976):

Bei der *Zahlabstraktion* soll das Kind jeweils zu einer Menge von Punkten im Wahl-Antwort-Verfahren aus fünf anderen Mengen diejenige herausuchen, die gleich viele Elemente enthält. Die Wahlmöglichkeiten sind so konstruiert, dass das Kind auf der Ebene

der bloßen Anschauung durch die Anordnung der Punkte und / oder der Größe der Punkte zu einer falschen Antwortalternative verführt wird. Der Subtest soll prüfen, inwieweit das Kind in seiner Zahlauffassung von der Anordnung und Größe der Elemente abstrahiert. Beim *Zählen von Objekten* geht es um die Prüfung der Fähigkeit, eine Menge von Punkten oder sonstigen einfachen geometrischen Figuren, die in unterschiedlicher Weise angeordnet sind, richtig zu zählen. Der besondere Schwierigkeitsgrad der Aufgaben besteht darin, trotz z.T. krummliniger, ungeordneter und überlagerter Objektanordnung, jedes Element nur einmal zu berücksichtigen.

Die Details der statistischen Analysen sind wiederum dem Tabellenband zu entnehmen.

*Für die zwei untersuchten Bereiche des Denkens (Zahlabstraktion, Zählen von Objekten) fielen die kovarianzanalytisch errechneten Befunde wie folgt aus: Zum Zeitpunkt der Einschulung zeigten sich bei den Schülern des B-Modells höhere Leistungen im Untertest Zahlabstraktion. Ansonsten waren bei allen Modellen deutliche, voneinander nicht unterscheidbare Entwicklungszuwächse bis etwa Mitte des ersten Schulhalbjahrs von Kl. 2 zu verzeichnen.*

## Motivationale Entwicklung

Neben den kognitiven Funktionen und Fertigkeiten sind in jüngerer Zeit bei der Diskussion erfolgreicher schulischer Sozialisation vermehrt die individuellen motivationalen Bedingungen hervorgehoben worden. Dabei werden bisweilen zwei sehr unterschiedliche Argumentationslinien miteinander verwechselt. Die eine Argumentationslinie bezieht sich auf die Bedeutung interindividueller Unterschiede im Ausmaß schulischer Lernmotivation, die andere auf die Bedeutung allgemeiner intraindividuelle Entwicklungsverläufe in der Lernfreude, mit der eine Altersabhängigkeit der spontanen Anstrengungsbereitschaft in Lernprozessen einhergeht. Während ersterer Aspekt das Problem mit sich bringt, dass die schulische Lernmotivation keine besonders zeitstabile personenspezifische Disposition ist, scheint uns die allgemeine Entwicklungssequenz der Lernfreude im Rahmen der Langzeitevaluation des Projektes "Schulanfang auf neuen Wegen" von besonderem Interesse zu sein (vgl. Hasselhorn, 1996).

Beim derzeitigen Stand der Forschung lässt sich folgendes Bild des allgemeinen intraindividuellen Entwicklungsverlaufes der schulischen Lernmotivation skizzieren: Zum Zeitpunkt der Einschulung ist bei nahezu allen Kindern mit einer sehr hohen Lernmotivation zu rechnen. Kinder dieser Entwicklungsstufe sind nicht nur neugierig, ja geradezu wissbegierig (vgl. Hasselhorn & Mähler, 1998), sondern sie sind auch durch einen scheinbar ungebrochenen Überoptimismus charakterisiert. Schulanfänger glauben, dass sie alles lernen und leisten können, wenn sie sich nur genügend anstrengen. Bereits die Tatsache, dass sie eine Tätigkeit ausführen, empfinden sie als Erfolg. Schon Piaget (1930) hatte bei jüngeren Kindern beobachtet, dass sie nicht zwischen Wünschen und realistischen Erwartungen unterscheiden und im Sinne eines „Wunschdenkens“ sich selbst alles zutrauen. So verwundert es auch nicht weiter, dass mehr als 90 Prozent aller Erstklässler glauben, dass sie selbst zu den besten Schülern ihrer Klasse gehören. Der Überoptimismus von Schulanfängern ist einerseits das Produkt einer kognitiven Unreife, andererseits versetzt er Kinder in einen optimalen motivationalen Zustand für die zu bewältigenden Lernanforderungen: Durch ihr (unrealistisch) positives Selbstbild investieren sie wie von selbst jede Menge Anstrengung in ihre Lernaufgaben (vgl. Bjorklund, Gaultney & Green, 1993).

Etwa im achten Lebensjahr verlieren Kinder diesen Überoptimismus. Zunächst sind es die Mädchen, etwas später auch die Jungen, die die kognitive Unreife überwinden und ihre eigenen Leistungsmöglichkeiten immer realistischer einzuschätzen vermögen. Doch dieser kognitive Realismus muss motivational teuer bezahlt werden: Wer sieht, dass er regelmäßig weniger gute Lernleistungen bringt als die Mehrzahl der Klassenkameradinnen und Klassenkameraden, der wird sich schwer tun, immer wieder von neuem Anstrengung zu investieren. Insofern kann durch die Zurückstellung eines Kindes von der Einschulung unter Umständen der spätere Schulstart motivational sogar erschwert werden. Denn wer schon im Laufe des Erstunterrichts seinen lernwirksamen Überoptimismus abgelegt hat, dem fällt es schwerer, viel Anstrengung in das schulische Lernen zu investieren.

Aufgrund dieses allgemeinen Entwicklungsbildes der schulischen Lernmotivation ist also ein Abbau der Lernfreude ab der zweiten Klassenstufe zu erwarten. Für den „Schulanfang auf neuen Wegen“ hielten wir darüber hinaus die Beantwortung der Frage für zentral, ob sich zwischen den Modellen mit jahrgangsübergreifenden Klassen und denen mit Jahrgangsklassen Unterschiede im Entwicklungsverlauf der schulischen Lernmotivation zeigen. Theoretisch denkbar wäre ja, dass die A-Modelle zwar positive Auswirkungen auf die kognitive Entwicklung von Kindern haben, jedoch durch das tägliche Miteinander mit älteren Schülern, die ihr (unrealistisch) positives Selbstbild bereits überwunden haben, auch ein vorzeitiger Verlust der damit einhergehenden motivationalen Vorteile nicht auszu-

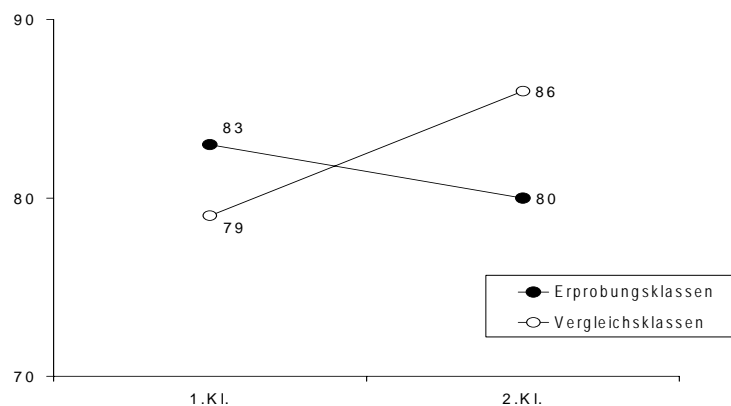
schließen ist. Daher wurde bereits schon die Voruntersuchung so angelegt, dass eine empirische Prüfung des Entwicklungsverlaufes der schulischen Lernfreude in den Erprobungs- und Vergleichsklassen möglich war.

Erfasst wurde die Lernfreude in Anlehnung an ein von Helmke (1993) entwickeltes Verfahren, bei dem die Kinder nach der Beliebtheit verschiedener Lernaktivitäten gefragt werden (z.B. Wie gerne gehst Du zur Schule? Wie gerne liest Du?). Die Antworten erfolgen auf einer 5-stufigen Ratingskala, die den Kindern mit verschiedenen abgestuften „happy / sad faces“ veranschaulicht wird (sehr gern / ziemlich gern / einigermaßen / nicht so gern / gar nicht gern). Aufgrund der Antworten der Kinder wurde ein Gesamtwert der Lernfreude (Addition sämtlicher Einzelskalen) zusammengestellt. Darüber hinaus konnten separate Analysen für die einzelnen Facetten der Lernfreude vorgenommen werden.

Für diesen Bericht lagen uns die Daten vom Beginn der ersten und dem Beginn der zweiten Klasse vor. Hier sollten noch keine bedeutsamen Veränderungen der allgemeinen schulischen Lernfreude auftreten. Wie den Ergebnistabellen im Tabellenband im Einzelnen zu entnehmen ist, fallen die Befunde für den Gesamtwert der Lernfreude erwartungsgemäß aus: Es ergaben sich in dem untersuchten Zeitraum keine nennenswerten Entwicklungsveränderungen, die „Höhe“ der Lernfreude im Zeitraum Beginn Kl. 1 und Beginn Kl. 2 blieb, über alle Schüler gerechnet, konstant.

Allerdings gab es in der Entwicklung modellspezifische Unterschiede: die A-Modelle und die Vergleichsschulen unterschieden sich jeweils signifikant von dem B-Modell. Seine Schüler zeigten eine höhere, eingangskorrigierte Lernfreude zu Beginn von Klasse 2.

Bei der Erprobung der Einzelskalen zur Lernfreude bei den sog. Intensivschulen von Kohorte 1 ergaben sich jedoch vereinzelte Abweichungen von diesem allgemeinen Befundmuster. So zeigten sich für die Lernfreude „Schule“ und für die Lernfreude „Lesen“ Unterschiede zwischen den Untersuchungsgruppen, die allerdings schon zu Beginn des ersten Schuljahres beobachtbar waren und daher nicht als Folge der unterschiedlichen pädagogischen Konzeptionen in den Erprobungs- und Vergleichsklassen interpretiert werden dürfen: Das motivationale Niveau in Bezug auf „Schule“ (allgemein) und „Lesen“ war in den Vergleichsklassen günstiger als in den Erprobungsklassen. Vermutlich ist dies als Zufallseffekt einzustufen, der allerdings dafür sorgt, dass die Kinder der Vergleichsklassen im Vergleich zu den Erprobungsklassen etwas günstigere motivationale Ausgangsniveaus aufzuweisen hatten.



**Abbildung 1: Veränderung der Lernfreude „Schreiben“ (in Prozent durchschnittlich vorhandener Lernfreude) zwischen dem Beginn der 1. und dem Beginn der 2. Klasse**

Ein Hinweis auf unterschiedliche motivationale Veränderungen zwischen Erprobungs- und Vergleichsklassen im Verlauf des ersten Schuljahres ergab sich lediglich für die Skala Lernfreude „Schreiben“ (vgl. Abbildung 1, vorherige Seite).

Hier zeigte sich ein statistisch bedeutsamer Effekt, der dadurch zustande kommt, dass die Lernfreude „Schreiben“ in den Erprobungsklassen im Laufe des untersuchten Jahresabstandes eher abnimmt, während sie in den Vergleichsklassen eher zunimmt.

Für die Skala „Lernfreude Hausaufgaben“, also für die positive Einstellung bzw. Motivation, mit der die Kinder an ihre Hausaufgaben herangehen, fand sich schließlich ein einheitlicher Veränderungseffekt, unabhängig davon, ob die Kinder eine Erprobungs- oder eine Vergleichsklasse besuchten: Die Kinder machten zu Beginn der 2. Klasse zwar immer noch recht gerne Hausaufgaben (Durchschnittswert bei 67%), dieser Wert liegt jedoch deutlich unter dem zu Beginn der 1. Klasse festgestellten Wert (78%).

*Insgesamt lässt sich zur motivationalen Entwicklung festhalten, dass die Kinder das hohe durchschnittliche Motivationsniveau zum Schulbeginn auch noch ein Jahr später zeigen; allerdings nicht ganz unabhängig vom Modell, das sie besuchen. Die Schüler der A-Modelle und der Vergleichsschulen zeigten etwa gleiche Werte, die der B-Modelle gaben im ersten Schulhalbjahr von Klasse 2 eine signifikant höhere Lernfreude an.*

## Soziales Verhalten und soziale Einstellungen

### Einstellungstest: SVG – Sozialverhalten in Gruppen (Klassen 1 und 2)

#### *Einführung*

Von Lehrern und Eltern wurde im Vorfeld des Modellversuchs der Verdacht geäußert, dass sich die Modellansätze im Rahmen von "Schulanfang auf neuen Wegen" negativ auf das Sozialverhalten von Kindern auswirkten, was wiederum die Lernatmosphäre in den Modellklassen ungünstig beeinflussen könnte. Somit ist das Untersuchungsziel die Erhebung des Sozialverhaltens der Schüler der Klassen 1 und 2 der Grundschule bzw. das, was man als gemeinsame Grundlage hinter den einzelnen Verhaltensäußerungen vermuten kann. Mit dem deutschen Fachbegriff wird dies als „Einstellung“ bezeichnet.

Der Umstand, dass Kinder, die eingeschult werden, in der Mehrzahl zwischen sechs und sieben Jahre alt sind, engt die Bandbreite der möglichen Verfahren erheblich ein. Wie die Literatur zeigt, dominiert bei dieser Altersgruppe die mittels systematisch konstruierter Schätzskalen durchgeführte Fremdbeobachtung von Erwachsenen, bei dem vorliegenden Untersuchungsziel meist die von Lehrerinnen/Lehrern. Bedauerlicherweise sind solche Schätzverfahren für den, der die Einschätzungen vornimmt, relativ aufwändig<sup>15</sup>.

Um die Belastung der an dem Modellversuch beteiligten Lehrer möglichst niedrig zu halten, sah sich der „Arbeitskreis Wissenschaftliche Begleitung“ daher gezwungen, besonders auf das Test-Nebenkriterium "Ökonomie"<sup>16</sup> zu achten. Es galt also einen Weg zu finden, wie man auf ökonomische Weise die gesuchten Informationen über das Sozialverhalten von den Schülern der Klassen 1 und 2 selbst bekommen kann.

Dieser Weg bestand darin, ein eigenes Verfahren zu konstruieren. Dieser muss zunächst genauer beschrieben werden, damit die Bedeutung der mit ihm im Rahmen der Begleituntersuchung erzielten Ergebnisse genauer eingeordnet werden kann. Eine detaillierte Beschreibung der Testkonstruktion und Testkennwerte (Itemkennwerte, Reliabilität und Validität) ist im Materialienband zum Zwischenbericht 2000 beigefügt.

#### *Konstruktion*

#### Besonderheiten und Einschränkungen, die sich aus der Zielgruppe des Verfahrens ergeben, und Konsequenzen für die Testkonzeption

Eine ökonomische Befragung von Schülern kann es nur mittels eines vom Schüler selbst zu bearbeitenden Testverfahrens geben. Hier ergibt sich aber die Schwierigkeit, dass Schüler und Schülerinnen der Klassen 1, und das gilt auch noch für einige der Klassen 2, noch unzureichend lesen und schreiben können. Es liegt daher nahe, die Items akustisch zu präsentieren und die vom Schüler ausgewählte Alternative mittels eines für diese Altersgruppe verständlichen optischen Rasters zu markieren.

Um einen hohen Grad an Aufmerksamkeit bei der Testdurchführung zu erreichen, schien es sinnvoll, sich das durch den Entwicklungsstand dieser Kinder geformte "Weltbild" zunutze zu machen, das Kroh<sup>17</sup> prägnant mit "magischer Phase" umschrieben hat.

<sup>15</sup> Eine Zwischenstellung nimmt dabei das DOS (Diagnostisches Soziogramm) von R. Müller (Hogrefe Verlag, 1980) ein, das laut Manual schon in der ersten Klasse eingesetzt werden kann. Allerdings sprechen mehrere Gründe gegen den Einsatz:

1) Normen liegen erst für die Kl. 3 vor. 2) Der Lehrer/die Lehrerin muss jeden einzelnen Schüler befragen. Stichworte der zu befürchtenden Nachteile: soziale Erwünschtheit der Schülerantworten (kaum kontrollierbar), zeitlicher Aufwand für den Lehrer/die Lehrerin (hoch).

<sup>16</sup> Lienert, G. A., *Testaufbau und Testanalyse*, Weinheim, Berlin, Basel: Verlag Julius Beltz, 1961<sup>3</sup>.

<sup>17</sup> Kroh, O., *Entwicklungspsychologie des Grundschulkindes*, Langensalza: Beyer, 1928.



Somit ergaben sich als Planungsrichtlinien für die Konstruktion des Testverfahrens folgende Ansätze:

*Untersuchungsziel:* Auskunft über das Sozialverhalten von Schülern in der Klasse

*Adressaten der Befragung:* Schüler der Klasse 1 und 2

*Testdarbietung:* akustisch

*Protokollierung der gewählten Item-Alternativen:* vom Schüler selbst, Ummalung einer Figur, die die Handlungsalternative symbolisiert, mit einem Stift

*Die Figuren, die die Handlungs-Alternativen symbolisieren, sollten einen hohen Bekanntheitsgrad haben:* hier empfahlen sich Tierfiguren.

*Die Tierfiguren sollten, um eine höhere und positive Identifikationswahrscheinlichkeit zu ermöglichen, möglichst anthropomorph sein:* die Figuren stellen bekleidete, mit Tierköpfen versehene, menschenähnliche Wesen dar, die freundlich und lustig wirken.

*Aufrechterhaltung von Konzentration und Motivation:* Einbettung der Testitems in eine möglichst phantasiereiche, phantastische und spannende Geschichte; die Testfiguren werden farbig gedruckt.

Die Altersgruppe der Schüler bestimmt eine Beschränkung auf wenige Items und Antwortalternativen. Die Zahl der Items wurde auf 13 beschränkt, mit einer Ausnahme werden pro Item nur zwei Antwortalternativen angegeben.

Natürlich ist zu erwarten, dass diese Beschränkungen zu einer deutlichen Varianzverminderung führen und daher zu einer Senkung der internen Konsistenz (Reliabilität) beitragen werden.

Insgesamt ist die Instruktion so aufgebaut, dass vor allem versucht wird, eine "funktionale Standardisiertheit"<sup>18</sup> herzustellen, d.h., möglichst alle Kinder sollen verstanden haben, was beim Test zu tun ist.

### Theoretisches Konzept des SVG

Im Folgenden soll versucht werden, durch die Formulierung einiger Grundannahmen etwas präziser zu fassen, was wir unter „Sozialverhalten“ verstehen und was der SVG misst. (Eine ausführliche Darstellung des theoretischen Konzepts ist dem Materialienband zum Zwischenbericht 2000, Kapitel: Testbericht „Einstellungstest: SVG - Sozialverhalten in Gruppen“ zu entnehmen).

In Anlehnung an die Lerntheorie wird davon ausgegangen, dass das Sozialverhalten eine Resultante eines komplexen Systems von Lernerfahrungen in früheren Interaktionssituationen darstellt. Solche früheren Lernerfahrungen werden zunächst in der Familie gesammelt, dann in Ablösung davon mit anderen Kindern, auch ohne Erwachsenenaufsicht. Im Kindergarten oder ähnlichen Einrichtungen sammeln die meisten Kinder zusätzlich noch Erfahrungen, die der Schulsituation am nächsten kommen.

Soziale Lernerfahrungen finden in der Regel immer dann statt, wenn die Bestrebungen<sup>19</sup> eines Individuums mit denen eines anderen oder mehrerer anderer zusammentreffen. Meist muss das Individuum dann in einem letztlich kognitiven Prozess entscheiden, ob es seine Bestrebungen weiterverfolgen soll oder ob negative Folgen absehbar sind. Kurz, es findet ein Abgleich mit erinnerten Erfahrungen aus ähnlichen situativen Interaktionssequenzen statt. Entscheidend ist, dass dieser Abgleich schnell erfolgen muss. Relativ häufig domi-

---

<sup>18</sup> Siehe dazu z.B. Cronbach, L. J., *Essentials of psychological testing*, New York/London: Harper & Row, 1961 (S.43 ff.)

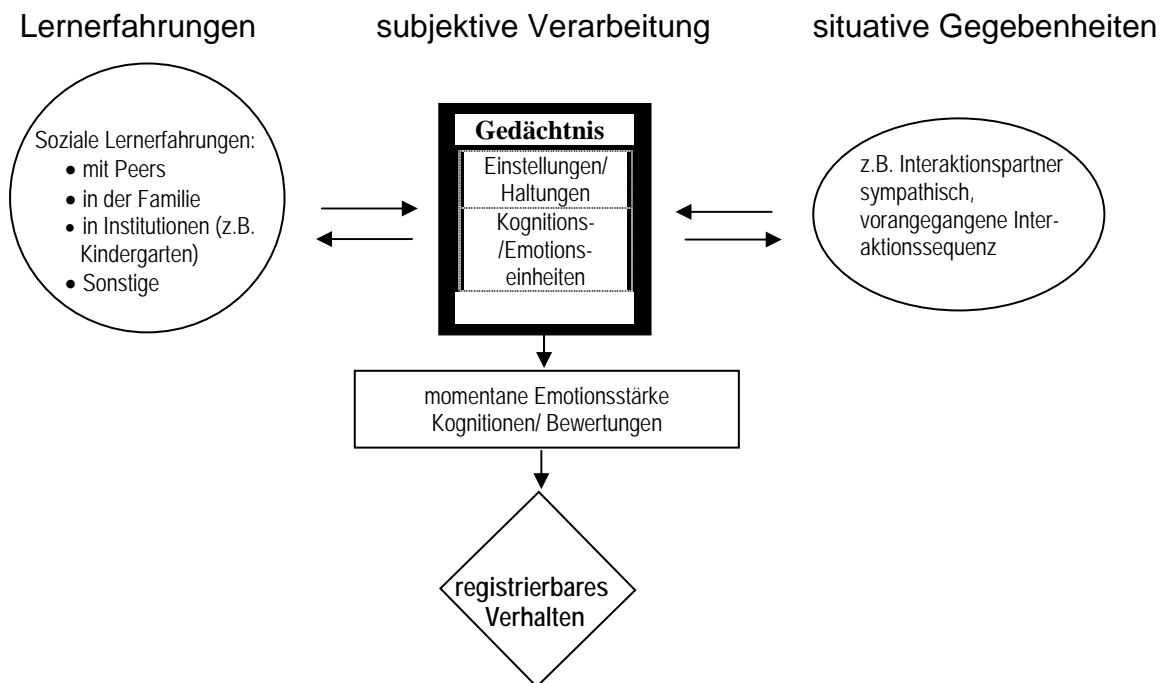
<sup>19</sup> Hier wird mit Absicht dieser unscharfe Begriff verwendet, um nicht zu nahe an die mit bestimmten Konnotationen verknüpften Begriffe wie "Bedürfnisse" oder "drive" etc. heranzukommen, was eine zusätzliche Ausweitung dieses Kapitels bedeuten würde.

niert die emotionale Komponente die kognitive, so dass bestimmte Einstellungen relativ diffus erscheinen, wenn man versucht, sie von Kindern detaillierter zu erfragen.

Mehr noch als bei älteren Kindern und bei Erwachsenen stellt sich in dieser Altersgruppe die Frage, wie eng Einstellung und beobachtbares Verhalten verbunden sind. Aufgrund der geringeren sozialen Lernerfahrung dieser Altersgruppe muss erwartet werden, dass die Situationsparameter erheblich deutlicher Verhalten bestimmen als dies bei älteren Schülern der Fall ist. Das bedeutet allerdings nicht, dass die Messung von Einstellungen wertlos ist, im Gegenteil, denn sie bilden die Keimzelle einer kognitiv/emotionalen Struktur, die durch die sozialen Lernerfahrungen weiter ausgeformt und zunehmend handlungsbestimmend wird. Letztlich muss es als eine erzieherische Aufgabe der Schule angesehen werden, sozial positive Einstellungen der Schüler zu fördern.

In der folgenden schematischen Darstellung wird versucht, die theoretischen Annahmen graphisch zu verdeutlichen.

**Theoretisches Modell über den Zusammenhang der SVG-Messungen mit dem tatsächlichen Sozialverhalten (schwarz geränderter Kasten = SVG-Messungen)**



Die schematische Darstellung verdeutlicht die Annahme, dass das durch andere Personen registrierbare Verhalten stark von situativen Momenten geprägt wird. Allerdings soll sie auch verdeutlichen, dass eine Fixierung auf die systematische Verhaltensbeobachtung das Problem aufwirft, auf welche relativ invariante, steuernde Momente (Einstellungen) die einzelnen Verhaltensweisen zurückgeführt werden können. Die Fehlervarianz wird durch ein solches Vorgehen mit ziemlicher Sicherheit eher erhöht, d.h., die Verhaltensvoraussage wird dadurch nicht verbessert. Lediglich der Aufwand, um solche Verhaltensbeobachtungen zu realisieren, ist erheblich größer als bei einer Einstellungsmessung.

**Als Ergebnis der theoretischen Grundannahmen ergaben sich für die Item-Konstruktion folgende Konsequenzen:**

Die Aufgabe bestand darin, Items zu konstruieren, die

- von den Schüler dieser Altersgruppe verstanden werden konnten,

- geeignet sind, ihre eher kognitiv diffusen/gefühlshaften Einstellungen anzusprechen,
- mehreren (hypothetischen) Dimensionen zugeordnet werden können.

Es wurden aus dem Itempool schließlich 13 Items ausgewählt, die folgenden drei hypothetischen Dimensionen zugeordnet wurden:

Kohäsion (positiv gegen Mitschüler eingestellt, keine Cliquenbildung, keine Aggression gegen Mitschüler, fühlt sich durch Mitschüler akzeptiert) (Items 1-5, 13)

Hilfsbereitschaft/Offenheit gegenüber neuen Sozialbeziehungen (Items 6, 9, 10)

Positive Besetzung des Einstellungskomplexes "Schule" (Schulfreude) (Items 7, 8, 11, 12)

Die Kurzformulierung der Fragen, Testanalyse/Verteilung der Antworten, Antwortbogen und der gesamte Fragenteil sind dem Materialienband zum Zwischenbericht 2000 beigelegt.

### *Testanalyse des SVG*

#### Vorbemerkung

Die in den nachfolgenden Kapiteln dargestellten teststatistischen Analysen zum „Sozialverhalten in der Gruppe“ (SVG) beziehen sich auf zwei verschiedene Versionen des Tests. Der SVG 1.1 wurde im Schuljahr 1997/88, der SVG 2.1 im Schuljahr 1998/99 eingesetzt. Beide Versionen unterscheiden sich in ihrer Kontextgeschichte und in den Tierbildern. Die insgesamt 13 Testitems, d.h. die Aussagen, zu denen die Schüler Stellung beziehen sollen, sind dagegen identisch.

Schüler, die im Schuljahr 1997/98 erstmals an der Untersuchung teilnahmen, werden im Folgenden als „Kohorte 1“ bezeichnet. Bei „Kohorte 2“ handelt es sich um Schüler, die im Schuljahr 1998/99 erstmals an der Untersuchung teilnehmen. (Kinder aus Grundschulförderklassen wurden bei der Bildung beider Kohorten nicht berücksichtigt.)

Für die teststatistischen Analysen standen demnach drei Analysegruppen zur Verfügung, die wie folgt beschrieben werden können:

#### **SVG 1.1 Kohorte 1**

Schüler des Einschulungsjahrgangs 97/98, die mit der Testversion SVG 1.1 im Schuljahr 1997/98 zum ersten Mal getestet wurden.

Alle Schüler (excl. GFK), von denen vollständige Datensätze vorlagen (**N=1158**).

#### **SVG 2.1 Kohorte 1**

Schüler des Einschulungsjahrgangs 97/98, die mit der Testversion SVG 2.1 im Schuljahr 1998/99 zum zweiten Mal getestet wurden.

Teilstichprobe, Schüler, von denen zum Zeitpunkt der Analysen vollständige Datensätze vorlagen (**N=788**).

#### **SVG 2.1 Kohorte 2**

Schüler des Einschulungsjahrgangs 98/99, die mit der Testversion SVG 2.1 im Schuljahr 1998/99 zum ersten Mal getestet wurden.

Teilstichprobe, Schüler (excl. GFK), von denen zum Zeitpunkt der Analysen vollständige Datensätze vorlagen (**N=1014**).

### Verteilung der Antworten

Generell, das ist aus der Verteilung der Antworten zu sehen, sind alle Items rechtsschief, d.h., die überwiegende Mehrzahl der Schüler entscheidet sich für die „sozial positive“ Antwort, in der Regel sind es mindestens über 80 % (siehe dazu auch den Tabellenband). Dies bedeutet, dass die große Mehrzahl der Schüler gerne in die Schule geht und sich zu eher positiven sozialen Einstellungen bekennt.

Aus den Schwierigkeitsindizes kann auch abgelesen werden, dass sich die Einstellung der Schüler zu den befragten Aspekten des Sozialverhaltens in den verschiedenen Analysegruppen als ziemlich stabil darstellt, vor allem auch, was die Rangreihe der Prozentwerte betrifft.

### Itemanalyse (Trennschärfe und Reliabilitätskoeffizienten (innere Testkonsistenz))

Die Trennschärfewerte liegen mit Ausnahme des Items 1 in der Regel in mindestens zwei Analysegruppen zwischen .30 und .55. Sie sind also, den Rahmenbedingungen entsprechend, zufriedenstellend.

Über alle Analysegruppen gesehen, erreichen die Items 4 und 8 relativ niedrige Trennschärfewerte. Tatsächlich unterscheiden sich die beiden Items inhaltlich von allen anderen. Bei Item 4 wird der Schüler als „Objekt“ angesprochen, während er bei den anderen Items „Subjekt“ ist. In Item 8 wird eine Stellungnahme zur Klassenlehrerin verlangt, während alle anderen Items die Schule allgemein bzw. die Klassenkameraden zum Inhalt haben.

Schließt man daher die beiden Items aus den Analysen aus, so hat dies eine Erhöhung der inneren Testkonsistenz zur Folge.

In Anbetracht der geringen Itemzahl und der hohen Tendenz zu positiven Einstellungen (also sehr schiefen Verteilung) sind die Werte als durchaus befriedigend zu bezeichnen.

### Faktorenanalyse<sup>20</sup>

Bei der Konstruktion des Tests zum Sozialverhalten in der Gruppe wurden zur Formulierung der Items die erwähnten und im Testbericht dargestellten hypothetischen Dimensionen herangezogen.

Mit Hilfe der Faktorenanalyse wurde überprüft, inwieweit sich die hypothetischen Dimensionen im vorliegenden Datenmaterial wiederfinden lassen.

Am stabilsten zeigt sich dabei die Dimension „Schulfreude“. Die Dimensionen „Kohäsion“ und „Hilfsbereitschaft“ lassen sich dagegen nur andeutungsweise reproduzieren. Eine Aufteilung des SVG in entsprechende Untertests empfiehlt sich daher nicht. Es bietet sich vielmehr an, von einem Generalfaktor auszugehen und bei künftigen Analysen mit einem Gesamttestpunktwert zu rechnen.

Durch die Faktorenanalysen konnten bis etwa 51% der Gesamtvarianz aufgeklärt werden.

---

<sup>20</sup> Extraktionsmethode: Hauptkomponentenanalyse  
Anzahl der Faktoren: Eigenwert > 1  
Rotationsmethode: Varimax mit Kaiser-Normalisierung

### Verteilung des SVG-Gesamtrohwertes

Die Verteilung des Testrohwerkes ist in allen Analysen (Kohorte 1 zwei Messungen: Schj. 1997/98 und 1998/99 und Kohorte 2 Schj. 1998/99 - siehe dazu den Tabellenband) deutlich rechtsschief und ähnelt einer J-Kurve, d.h., das „Sozialverhalten in der Gruppe“ ist statistisch gesehen nicht normal verteilt. Dies zeigen auch die im Fuß der Tabellen angeführten Mittelwerte und Mediane, die sich dicht bei der maximal zu erreichenden Testpunktzahl (13 bzw. 11) ansiedeln.

### Zur Validität des SVG

Wenn ein Testverfahren neu entwickelt wird, muss zunächst untersucht werden, ob es auch tatsächlich das misst, was es zu messen vorgibt.

Untersuchungen zur Validität sind einfach durchzuführen, wenn ähnliche Verfahren eingesetzt werden können. Dies ist, wie schon dargestellt wurde, kein Weg, der in dieser Begleituntersuchung gegangen werden kann. Erste Anhaltspunkte zur Validität ergeben allerdings einige wenige Items, die aus dem in dieser Begleituntersuchung eingesetzten. „Bogen zur Schülerbeobachtung in den ersten Grundschulklassen (BGS-L)“, Skala „Sozialverhalten“, etwa in die gleiche Richtung zielen. (Näheres ist dem Materialienband zum Zwischenbericht 2000 zu entnehmen.)

### *Ergebnisse*

#### Einführung

Die folgenden Berechnungen geben Auskunft über mögliche Veränderungen der Schülereinstellungen zum Sozialverhalten in der Gruppe im Zeitraum vom ersten Schulhalbjahr Kl. 1 bis Mitte / Ende von „Klassenstufe 2“.

Es können dabei sowohl die Ergebnisse für

- die Gesamtheit der Erprobungs- und Vergleichsschulen (Gesamtstichprobe) Kohorte 1 als auch die für
- die Schülerstichprobe aus den 41 Versuchsschulen, die die Kohorte 3 stellen, aufgeführt werden.

Bei den Fallzahlen ist zu beachten, dass es im Längsschnitt zu Datenausfällen gekommen ist.

Als Folge der rechtsschiefen Verteilung der Rohwerte wurden die Ergebnisse zunächst mit einem robusteren Verfahren, dem sog. T-Test, gerechnet. Um allerdings auch möglichen Wechselwirkungen auf die Spur zu kommen, wird auch das Verfahren der sog. Messwiederholungsvarianzanalyse eingesetzt.

#### Ergebnisse bei den verschiedenen Vergleichsstichproben (Kohorte 1)

#### *Ergebnisse bei der Gesamtstichprobe (alle in die Begleituntersuchung einbezogenen Klassen)*

Anmerkung:

Die detaillierten Einzelergebnisse sind dem Tabellenband zu entnehmen.

### Gruppierungen von Analysegruppen der Gesamtstichprobe nach dem Kriterium I

Im Folgenden werden Ergebnisse mit zwei Gruppierungen innerhalb der „Gesamtstichprobe 1“ vorgestellt. Diese Gruppierungen bestehen aus verschiedenen gebildeten Analysegruppen. Die Gesamtstichprobe 1 wird nach Kriterien des Testzeitpunkt 1 (Kriterium I) gebildet, hier sind die B-Modellklassen Erprobungsklassen.

Die Gesamtstichprobe 1 besteht aus 5 Analysegruppen:

3 Analysegruppen, die jeweils aus den Erprobungsklassen der verschiedenen Modellvarianten (A1-, A2-, B-Modelle) gebildet wurden,

2 Analysegruppen, die jeweils aus den Vergleichsklassen gebildet wurden. Wobei die Vergleichsklassen in den Erprobungsschulen, die eine der drei Modellvarianten durchführen, als gesonderte Gruppe ausgewiesen werden.

Die Gesamtstichprobe 2 besteht aus 3 Analysegruppen:

1 Analysegruppe der A-Modelle,

1 Analysegruppe B-Modell,

1 Analysegruppe Vergleichsklassen (Vergleichsklassen der Erprobungsschulen und Vergleichsschulen).

Dieses Vorgehen bietet sich aus mehreren Gründen an: Die Fallzahlen in den Gruppen werden größer, eine Vergleichbarkeit mit den Ergebnissen der acht sog. Intensivschulen wird möglich, die Einteilung berücksichtigt die „Modellkategorien“.

### Ergebnisse Gesamtstichprobe 1

#### Unterschiede im Querschnittsvergleich

Schuljahr 1997/98

Es bestehen signifikante Unterschiede zwischen dem Sozialverhalten von Schülern in den verschiedenen Analysegruppen. Der niedrigste Wert ist bei der Analysegruppe 2 (A2-Erprobungsklassen) anzutreffen, der höchste bei der Analysegruppe 5 (Vergleichsschulklassen). Wichtig ist dabei festzuhalten, dass die Erprobungsklassen zum Testzeitpunkt 1 noch Jahrgangsklassencharakter aufwiesen.

Schuljahr 1998/99

Die Rangreihe der Unterschiede zwischen den verschiedenen Analysegruppen bleibt mit einer Ausnahme bestehen. (Bei der Analysegruppe 5 ist, im Gegensatz zu den anderen vier Gruppen, ein etwas niedrigerer Mittelwert bei der Messung zum Zeitpunkt 2 zu registrieren). Die Unterschiede zwischen dem Sozialverhalten von Schülern der Klassen in den verschiedenen Analysegruppen Ende des zweiten Schulhalbjahrs der Klasse 2 bleiben also weiterhin signifikant.

#### Unterschiede im Längsschnittsvergleich

Es sind in der Gesamtstichprobe 1 keine signifikanten Veränderungen im Sozialverhalten im Laufe von etwa anderthalb Jahren zu sichern, obwohl der Mittelwert bei allen Analysegruppen sogar noch leicht zunimmt.

Die Wechselwirkung ist nicht signifikant, also sind keine systematischen modellspezifischen „Treatment-Einflüsse“ (hier Wirkungen der anderthalbjährigen Unterrichtung in verschiedenen modellspezifischen Strukturen) auf die Einstellung der Schüler zu erkennen. Insgesamt kann demnach bei der Analyse der Einstellung der Schüler zum Sozialverhalten, wenn eine Einteilung in fünf Analysegruppen vorgenommen wird, mit gewisser Sicherheit der Einfluss des „modellspezifischen Treatments“ als vernachlässigbar klein eingeschätzt werden.

### Ergebnisse Gesamtstichprobe 2

#### Unterschiede im Querschnittsvergleich

##### Schuljahr 1997/98

Es bestehen gerade noch signifikante Unterschiede zwischen dem Sozialverhalten von Schülern in den verschiedenen Analysegruppen. Die noch nicht jahrgangsgemischten Klassen der A-Modelle weisen beim Testzeitpunkt 1 den niedrigsten Wert auf, die Vergleichsklassen den höchsten.

##### Schuljahr 1998/99

Die Rangreihe der Unterschiede zwischen den verschiedenen Analysegruppen bleibt unverändert. Die Unterschiede zwischen dem Sozialverhalten von Schülern der Klassen in den verschiedenen Analysegruppen Ende des zweiten Schulhalbjahrs der Klasse 2 bleiben also bestehen.

#### Unterschiede im Längsschnittsvergleich

Es sind in der Gesamtstichprobe keine signifikanten Veränderungen im Sozialverhalten im Laufe von etwa anderthalb Jahren zu sichern, obwohl der Mittelwert bei allen Analysegruppen sogar noch leicht zunimmt.

Die Wechselwirkung ist nicht signifikant, also sind keine systematischen modellspezifischen „Treatment-Einflüsse“ (hier Wirkungen der anderthalbjährigen Unterrichtung in verschiedenen modellspezifischen Strukturen) auf die Einstellung der Schüler zu erkennen. Insgesamt kann demnach bei der Analyse der Einstellung der Schüler zum Sozialverhalten, auch wenn eine Einteilung in drei Analysegruppen vorgenommen wird, der Einfluss des „modellspezifischen Treatments“ als vernachlässigbar klein eingeschätzt werden.

### Gruppierungen von Analysegruppen der Gesamtstichprobe nach dem Kriterium II

Nach dem Kriterium II zusammengefasste Analysegruppen gehen vom Testzeitpunkt 2 aus. Hier, in der Jahrgangsstufe 2, sind die Klassen des B-Modells strukturell den Vergleichsklassen zuzuordnen. Es wird hier also ein Vergleich von kombinierten Klassen (Erprobungsklassen A1 und A2) und Jahrgangsklassen (Vergleichsklassen, Klassen des B-Modells und Vergleichsschulen) vorgenommen.

#### Unterschiede im Querschnittsvergleich

##### Schuljahr 1997/98

Es bestehen gerade noch signifikante Unterschiede zwischen dem Sozialverhalten von Schülern in den beiden Analysegruppen. Die kombinierten Klassen starten also mit einem etwas niedrigeren Mittelwert.

Schuljahr 1998/99

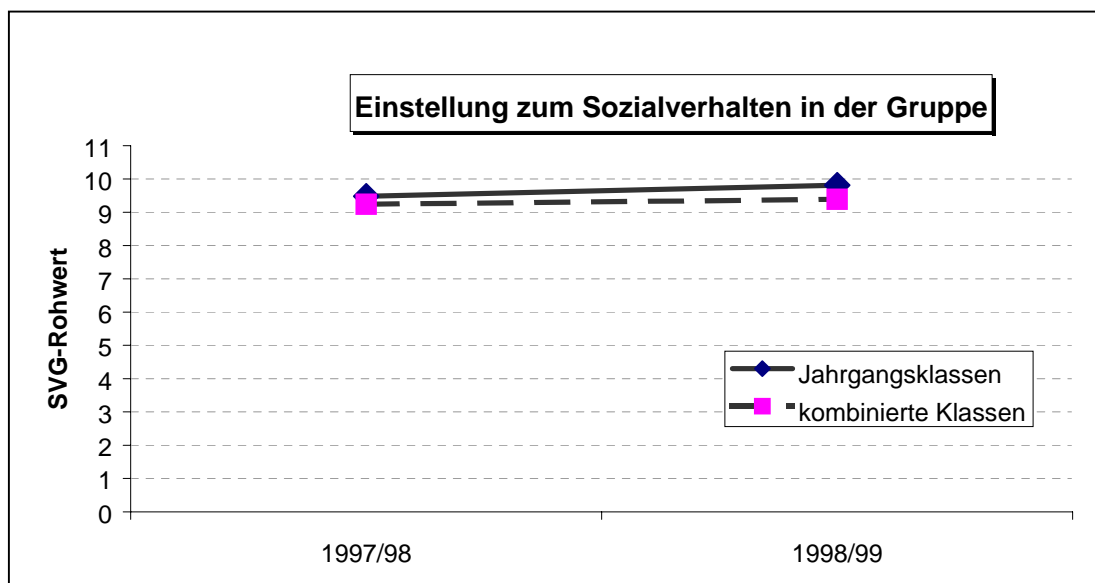
Der Unterschied zwischen den beiden Analysegruppen bleibt unverändert auch am Ende der zweiten Klasse weiterhin signifikant bestehen.

Unterschiede im Längsschnittvergleich (siehe dazu auch die Abbildung 2)

Es sind auch, wenn man die Gesamtstichprobe nach dem Kriterium II einteilt, keine signifikanten Veränderungen im Sozialverhalten im Laufe von etwa anderthalb Jahren zu sichern, obwohl der Mittelwert bei allen Analysegruppen sogar noch leicht zunimmt.

Die Wechselwirkung ist nicht signifikant, also sind auch hier keine systematischen „modellspezifischen Treatment-Einflüsse“ auf die Einstellung der Schüler zu erkennen.

Insgesamt kann auch bei diesem Ansatz kein Einfluss des „modellspezifischen Treatments“ auf die Einstellung der Schüler zum Sozialverhalten aufgezeigt werden.



**Abbildung 2: Längsschnitt (97/98 bis 98/99) der Gesamtstichprobe nach dem Kriterium II**

#### *Zusammenschau der Ergebnisse der Vergleichsuntersuchungen (Kohorte 1)*

Zunächst ist festzuhalten, dass die Intensivschulen sich als besondere Schulen ausweisen. Sie starten mit sehr hohen Durchschnittswerten. Obwohl die Unterschiede zwischen Erprobungs- und Vergleichsklassen in ähnliche Richtung zeigen wie bei der Gesamtstichprobe, sind sie nicht signifikant. Die signifikante Abnahme des Mittelwerts zum Testzeitpunkt 2 – anders als dies bei den beiden Analysen mit der Gesamtstichprobe der Fall ist – dürfte sich am besten mit der „Tendenz zur Mitte“, einem häufig zu beobachtenden Effekt, erklären lassen.

Bei der Gesamtstichprobe starten die Erprobungsklassen, am deutlichsten ist dies bei den A-Modellen zu sehen, mit niedrigeren Mittelwerten. Dabei muss allerdings berücksichtigt werden, dass diese Erprobungsklassen zum Testzeitpunkt 1 noch nicht jahrgangsgemischt



waren. Ob sich hier erste Effekte des Modellansatzes (weniger Zurückstellungen) bemerkbar machen, kann ohne Zusatzinformationen nicht beurteilt werden.

Interessant ist, dass sich bei der Gesamtstichprobe keinerlei „modellspezifischen Treatmenteffekte“ aufzeigen lassen. Es lassen sich also vom Testzeitpunkt 1 bis zum Testzeitpunkt 2, die etwa eineinhalb Jahre auseinander liegen, keine signifikanten Veränderungen und keinerlei besondere Modelleffekte aufzeigen. Dieses Ergebnis ist unabhängig davon, nach welchem der drei Ansätze die Analysegruppen gebildet werden. Die kombinierten Klassen verhalten sich also im Längsschnittsverlauf wie die Jahrgangsklassen, was die Einstellung zum Sozialverhalten in der Gruppe betrifft.

### Ergebnisse bei den 41 Versuchsschulen der Kohorte 3

Anmerkung:

Im Gegensatz zu den differenzierten Analysen bei der Kohorte 1 wird hier nur von Analysen berichtet, bei denen die unabhängige Variable „Modell“ aus diesen drei Gruppierungen besteht: Gruppe 1 = A-Modelle, Gruppe 2 = B-Modell, Gruppe 3 = Vergleichsschulen

Unterschiede im Querschnittsvergleich

Schuljahr 1999/00 (Mitte)

Unterschiede bezüglich des Sozialverhaltens sind im ersten Schulhalbjahr der Klasse 1 erkennbar: Die Schüler der A-Modelle und des B-Modells haben signifikant höhere Werte als die der Vergleichsschulen.

Schuljahr 1999/00 (Ende)

Die Unterschiede bezüglich des Sozialverhaltens sind am Ende des ersten Schuljahres der Klasse 1 weiterhin erkennbar: Die Schüler der A-Modelle und des B-Modells haben signifikant höhere Werte als die der Vergleichsschulen.

Schuljahr 2000/01 (Ende)

Es sind zwar am Ende des zweiten Schuljahres numerische Unterschiede bezüglich des Sozialverhaltens von Schülern der A-Modelle und des B-Modells und denen der Vergleichsklassen zu erkennen, diese sind aber nicht signifikant.

Unterschiede im Längsschnittsvergleich

Für die Gesamtstichprobe sind Veränderungen im Sozialverhalten im Laufe von etwa anderthalb Jahren zu sehen. Zunächst fällt auf, dass der durchschnittliche Mittelwert signifikant zunimmt. Die Wechselwirkung ist nicht signifikant, also sind keine systematischen modellspezifischen „Treatment-Einflüsse“ (hier: Wirkungen der anderthalbjährigen Unter-richtung in verschiedenen modellspezifischen Strukturen) auf die Einstellung der Schüler zu erkennen.

Die Gruppenunterschiede fallen nicht signifikant aus.

Bei den 41 Versuchsschulen in Kohorte 3 kann demnach der Einfluss der „modellspezifischen Treatments“ auf die Ausprägung des Sozialverhaltens als vernachlässigbar klein eingeschätzt werden.

### *Zusammenfassung der Ergebnisse von Kohorte 1 und 3*

Bei statistischen Analysen, ob modellspezifische Einflüsse auf die Einstellung zum Sozialverhalten in Gruppen nachzuweisen sind, ergaben sich folgende Ergebnisse:

Sowohl bei den Analysen der Gesamtstichprobe (Kohorte 1) als auch bei der Stichprobe der Kohorte 3 sind zu beiden Testzeitpunkten signifikante Unterschiede zwischen den Analysegruppen festzustellen, die allerdings keine Modellsystematik aufweisen.

Für beide Kohorten ist augenfällig, dass die Mittelwerte in Kl. 2 Ende höher sind, als die in Kl.1 Mitte (bei Kohorte 1 ist dieses Ergebnis nur numerisch festzustellen, in Kohorte 3 ist es dann signifikant). D.h., zumindest als Einstellung verfestigt sich das schon in Kl. 1 beobachtete sehr positive Sozialverhalten.

Sowohl in der Gesamtstichprobe der Schülerinnen und Schüler der Kohorte 1 als auch in der Stichprobe der Kohorte 3 sind keine signifikanten Wechselwirkungen festzustellen, d.h., es sind keine systematischen modellspezifischen „Treatment-Einflüsse“ (hier Wirkungen der anderthalbjährigen Unterrichtung in verschiedenen modellspezifischen Strukturen) auf die Einstellung der Schülerinnen und Schüler zu erkennen.

*Insgesamt lässt sich zur Entwicklung des Sozialverhaltens festhalten, dass die Kinder das hohe durchschnittliche Einstellungsniveau zum Schulbeginn auch noch ein Jahr später zeigen; allerdings nicht ganz unabhängig vom Modell, das sie besuchen, wobei es in den Ergebnissen auch Unterschiede zwischen den beiden Kohorten (Schülerjahrgänge 1 und 3) gibt. Somit muss eher von einer modellunabhängigen als von einer modellabhängigen Förderung des Sozialverhaltens ausgegangen werden.*

## Einstellungstest: SVG-4.1 – Sozialverhalten, Soziale Kompetenz und Schulfriedenheit (Klasse 4)

### *Einführung*

Eines der Ziele der Begleituntersuchung des Modellprojekts „Schulanfang auf neuen Wegen“ bestand darin, die Entwicklung des Sozialverhaltens der Grundschüler von Klasse 1 bis Klasse 4 nachzuzeichnen, natürlich unter der Berücksichtigung des Modelleinflusses. Musste für die Klassenstufen 1 und 2 (die Eingangsstufe) ein völlig neuartiger Weg, um die Schülereinstellung zu erfassen, gewählt werden, so konnte in Klassenstufe 4 auf die bewährte Fragebogenmethode zurückgegriffen werden.

### *Konstruktion*

Der SVG-4.1 „Fragen an Schülerinnen und Schüler der Klasse 4“ besteht aus zwei Teilen mit insgesamt 43 Items<sup>21</sup>. Teil I trägt die Überschrift: „Die Kinder meiner Klasse, meine Lehrerinnen und Lehrer und meine Grundschule“. Es wird ein Satz angefangen und drei mögliche Alternativen zur Beantwortung angeboten, die jeweils angekreuzt werden können.

Teil II stellt die Leitfrage: „Wie versuchst du Probleme zu lösen, die es schon mal beim Zusammensein mit anderen Kindern und auch mit Lehrerinnen/Lehrern gibt?“ Hier wird eine alltägliche Schulsituation beschrieben, die ein Konfliktpotential enthält. Zur Lösung werden vier Alternativvorschläge vorgestellt, von denen einer ausgewählt werden soll.

### *Konzept des SVG 4.1*

Das Konzept lehnt sich zum einen Teil an den „Fragebogen zur sozialen Kompetenz“ im Rahmen des Projekts „Wissenschaftliche Begleituntersuchung von IMPULSE- Guter Start an der Hauptschule“<sup>22</sup> an und zum anderen an die Studie „Belastung in der Schule?“<sup>23</sup>

Nach einer Testerprobung an zwei Grundschulen wurde von 8 Skalen ausgegangen:

Schulfriedenheit

Einstellung zu Lehrerinnen und Lehrern

Gegenseitige Hilfsbereitschaft

Klassenzusammenhalt

Konfliktlösungsansatz: soziale Kompetenz

Konfliktlösungsansatz: soziale Aggressivität

Konfliktlösungsansatz: Lösung durch Autoritätspersonen

Konfliktlösungsansatz: soziale Gleichgültigkeit

---

<sup>21</sup> Fragebogen kann auf Wunsch beim Landesinstitut für Schulentwicklung, Ref. 11, angefordert werden.

<sup>22</sup> Thiel, R.-D., Sikorski P. B. & Dr. Peschl V. M. IMPULSE – Guter Start in der Hauptschule - Ergebnisse der wissenschaftlichen Begleituntersuchung: Abschlussbericht. Landesinstitut für Erziehung und Unterricht, Ref. 11 Ber. 11-04-05, 2004.

<sup>23</sup> Projektgruppe Belastung: Belastung in der Schule? Eine Untersuchung an Hauptschulen, Realschulen und Gymnasien Baden-Württembergs. Beltz – Deutscher Studien Verlag, 1998.

### *Testanalyse des SVG 4.1*

Der SVG 4-1 wurde im Schuljahr 2002/2003 zum Ende von Kl. 4, Kohorte 3 eingesetzt. In die Auswertung konnten N=680 Bögen einbezogen werden.

Erste Analysen zeigten, dass die acht postulierten Skalen nicht bestätigt werden konnten. Daher wurden die Items einer Faktorenanalyse unterzogen.

### Faktorenanalysen<sup>24</sup>

Durch die Faktorenanalysen konnten bis 57 % der Gesamtvarianz aufgeklärt werden.

Für eine schlüssige Interpretation schälten sich 5 Faktoren heraus:

Einschätzung von Lehrerinnen/Lehrern, ihrem Unterricht und der Schule insgesamt (Skala 1)<sup>25</sup>

Einschätzung der Mitschüler und der gegenseitigen Hilfsbereitschaft (Skala 2)<sup>26</sup>

Störtendenzen gegen ein gutes Klassenklima (Skala 3)<sup>27</sup>

Konfliktlösungsansatz: soziale Kompetenz (Skala 4)<sup>28</sup>

Konfliktlösungsansatz: Aggressivität (Skala 5)<sup>29</sup>

### Mittelwerte (MW) und Standardabweichungen (SW)

Skala 1 (9 Items): MW = 14,81, SW = 3,49

Skala 2 (6 Items): MW = 9,74, SW = 2,25

Skala 3 (6 Items): MW = 9,84, SW = 1,89

Skala 4 (15 Items): MW = 7,26, SW = 3,08

Skala 5 (10 Items): MW = 1,11, SW = 1,58

### Reliabilitäten<sup>30</sup>

Skala 1: **.84**

Skala 2: **.66**

Skala 3: **.60**

Skala 4: **.70**

Skala 5: **.72**

(Gesamt: **.85**)

---

<sup>24</sup> Extraktionsmethode: Hauptkomponentenanalyse; Rotationsmethode: Varimax mit Kaiser-Normalisierung

<sup>25</sup> Beispiel: (Frage 14) Ich gehe gern in diese Schule.

<sup>26</sup> Beispiel: (Frage 12) Die Kinder unserer Klasse helfen sich gegenseitig im Unterricht.

<sup>27</sup> Beispiel: (Frage 8) Wenn ein Kind etwas Dummes gesagt hat, dann lachen (alle / einige/ keiner) in der Klasse über es.

<sup>28</sup> Beispiel: (Frage 43, d) [Konflikt:] Zwei deiner liebsten Klassenkameraden hatten einen Riesenkrach. Sie sprechen seitdem nicht mehr miteinander. [Lösung d:] Ich versuche die beiden wieder miteinander zu versöhnen.

<sup>29</sup> Beispiel: (Frage 36, c) [Konflikt:] Eine Mitschülerin in eurer Klasse boxt, stößt oder zwickt die schwächeren Kinder um sie herum. Lösung c: Ich drehe mich um und boxe sie.

<sup>30</sup> Innere Konsistenz

## Diskussion

Die Höhe der Reliabilitätskoeffizienten leidet u. a. auch darunter, dass bestimmte Einstellungen (Skala 5, abgeschwächt auch Skala 3) eher in die Gruppe der seltenen Ereignisse einzustufen sind, was zwangsläufig schiefe Verteilungen mit sich bringt. D.h., die verwendeten Rechenprozeduren ergeben nicht ganz zuverlässige Werte, wobei erfahrungsgemäß Unterschätzungen relativ häufig sind.

Laut Lienert<sup>31</sup> reichen zur Errechnung von Gruppendifferenzen Werte von  $>.50$  gerade noch aus. D.h., zur Ermittlung von Modellunterschieden können alle 5 Skalen herangezogen werden.

## *Testgesamtwert*

Die in Klassenstufen 1 und 2 eingesetzten SVG-Testversionen erbrachten nur einen Gesamtwert. Zum besseren Vergleich der Ergebnisse wurde daher auch vom SVG 4-1 ein Gesamtwert berechnet; die Skala 4 wurde dabei invertiert<sup>32</sup>. Um die einzelnen Skalen gleichgewichtet zu behandeln, wurden die Rohwerte in z-Werte transformiert.

## *Ergebnisse*

### Modellvergleich

Es werden pro Skala vier Ansätze bzw. Gruppierungen miteinander verglichen. Diese vier Ansätze lassen sich davon ableiten, dass ab Klasse 2 (B-Modell) bzw. ab Klasse 3 (die A-Modelle) die Schülerinnen und Schüler nach denselben Ansätzen (Jahrgangsprinzip) unterrichtet werden, also quasi alle „Regelunterricht“ erhalten.

Die Grundschulzeit umfasst in Baden-Württemberg vier Klassenstufen. Geht man nun davon aus, dass Nachwirkungen durch das unterschiedliche „pädagogische Treatment“, also die „Modellbeeinflussung“ zu erwarten sind, so stellt sich die pädagogische Situation für die A-Modelle und die B-Modelle unterschiedlich dar: Die A-Modelle unterrichten zwei Jahre nach dem „Modellansatz“ und zwei Jahre nach dem „Regelschulansatz“; die B-Modelle entsprechend ein Jahr und drei Jahre. Im Unterschied zu den Vergleichsschulen ist das Gemeinsame bei den A- und den B-Modellen, dass sie während der Grundschulzeit auch nach einem Modellansatz unterrichtet wurden. Somit kann man vier Vergleichsgruppierungen bzw. -ansätze berechnen:

1. Alle vier Modellgruppen: A1-Modell, A2-Modell, B-Modell und Vergleichsschulen
2. Zwei Gruppen: Die beiden A-Modelle (Gruppe 1) vs. B-Modell und Vergleichsgruppen (Gruppe 2)
3. Drei Gruppen: Die beiden A-Modelle vs. das B-Modell und vs. die Vergleichsschulen

---

<sup>31</sup> Lienert & Raatz: Testaufbau und Testanalyse. Beltz Verlag, 1998, S. 269.

<sup>32</sup> Eine Faktorenanalyse der 5 Skalen erbrachte nach dem Eigenwertkriterium nur einen substantiellen Faktor. Danach ist die Bildung eines Gesamtwerts zu vertreten.

4. Zwei Gruppen: (Gruppe 1) die Klassen mit Modellansätzen (A- und B-Modelle) vs. (Gruppe 2) Vergleichsschulen .

Querschnittsvergleiche

**Skala 1:** (Positive) Einschätzung von Lehrerinnen/Lehrern, ihrem Unterricht und der Schule insgesamt [Lesehilfe: Je kleiner die Mittelwerte, desto positiver die Einschätzung]

**Tabelle 3: SVG-4.1-Skala1 (Schule, Lehrer, Unterricht) – vier Vergleichsansätze**

Faktor 1 (Skala 1)	<i>Vergleichsansatz 1: A1-Modell vs. A2-Modell vs. B-Modell vs. Vergleichsschulen</i>					
<b>Modell</b>	A1	A2	B	V	Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
<b>Anzahl</b>	180	134	122	220		
<b>Mittelwerte</b>	16,13	15,51	16,67	16,24	0,10 (ns.)	A2 < B *

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 1 (Skala 1)	<i>Vergleichsansatz 2: A1-Modell+ A2-Modell vs. B-Modell+ Vergleichsschulen</i>					
<b>Modell</b>	A1 + A2		B + V		Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
<b>Anzahl</b>	314		342			
<b>Mittelwerte</b>	15,87		16,39		0,08 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 1 (Skala 1)	<i>Vergleichsansatz 3: A1-Modell+ A2-Modell vs. B-Modell vs. Vergleichsschulen</i>					
<b>Modell</b>	A1 + A2		B	V	Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
<b>Anzahl</b>	314		122	220		
<b>Mittelwerte</b>	15,87		16,67	16,24	0,12 (ns.)	A1+A2 < B *

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 1 (Skala 1)	<i>Vergleichsansatz 4: A1-Modell+ A2-Modell+ B-Modell vs. Vergleichsschulen</i>					
<b>Modell</b>	A1 + A2 + B			V	Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
<b>Anzahl</b>	436			220		
<b>Mittelwerte</b>	16,09			16,24	0,65 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Beim generellen Vergleich aller vier Vergleichsansätze ergeben sich keine signifikanten Unterschiede.

Beim Post-hoc-Test (also im hierarchisch nachgeordneten Vergleich zwischen den einzelnen Modellen) unterscheidet sich beim Vergleichsansatz 1 nur die Einschätzung der Schülerinnen/Schüler aus den A2-Modellen gegenüber denen der B-Modelle (die ersteren geben die signifikant positivere Einschätzung an).

Auch im Vergleichsansatz 3 schneidet das B-Modell gegenüber den zusammengefassten A-Modellen signifikant weniger gut ab, was hauptsächlich auf das Abschneiden des A2-Modells zurückzuführen ist.

*Fazit: Die Einschätzung der Schülerinnen und Schüler der vier Modelle unterscheidet sich nicht wesentlich. Die B-Modellschüler sehen ihre Schulen und Lehrerinnen und Lehrer etwas distanzierter als die Schülerinnen und Schüler der anderen drei Modelle.*

**Skala 2:** (Positive) Einschätzung der Mitschüler und der gegenseitigen Hilfsbereitschaft  
[Lesehilfe: Je kleiner die Mittelwerte, desto positiver die Einschätzung]

**Tabelle 4: SVG-4.1-Skala2 (Mitschüler) – vier Vergleichsansätze**

Faktor 2 (Skala 2)	<i>Vergleichsansatz 1: A1-Modell vs. A2-Modell vs. B-Modell vs. Vergleichsschulen</i>					
Modell	A1	A2	B	V	Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	181	135	125	226		
Mittelwerte	9,94	9,70	9,42	9,79	0,24 (ns.)	[A1>B *] <sup>33</sup>

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 2 (Skala 2)	<i>Vergleichsansatz 2: A1-Modell+ A2-Modell vs. B-Modell+Vergleichsschulen</i>					
Modell	A1 + A2		B + V		Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	316		351			
Mittelwerte	9,84		9,66		0,30 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

<sup>33</sup> Ergebnis sollte aus methodischen Gründen nur als Tendenz interpretiert werden.

Faktor 2 (Skala 2)	<i>Vergleichsansatz 3: A1-Modell+ A2-Modell vs. B-Modell vs. Vergleichsschulen</i>				
<b>Modell</b>	A1 + A2	B	V	Signifikanz der Modellunter- schiede, alle Modelle	Signifikanz der Modellunter- schiede, einzel- ne Modelle
<b>Anzahl</b>	316	125	226		
<b>Mittel- werte</b>	9,84	9,42	9,79	0,19 (ns.)	---

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau,  
(ns.) = nicht signifikant, --- = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 2 (Skala 2)	<i>Vergleichsansatz 4: A1-Modell+ A2-Modell+ B-Modell vs. Vergleichsschulen</i>			
<b>Modell</b>	A1 + A2 + B	V	Signifikanz der Modellunter- schiede, alle Modelle	Signifikanz der Modellunter- schiede, einzel- ne Modelle
<b>Anzahl</b>	441	226		
<b>Mittel- werte</b>	9,72	9,79	0,69 (ns.)	---

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau,  
(ns.) = nicht signifikant, --- = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Beim generellen Modellvergleich sind bei allen vier Vergleichsansätzen keinerlei signifikante Unterschiede festzustellen.

Beim Post-hoc-Test schätzen die B-Modellschüler ihre Mitschüler signifikant positiver ein als die A1-Modellschüler.

*Fazit: Die Einschätzung der Schülerinnen und Schüler der vier Modelle unterscheidet sich nicht wesentlich. Die B-Modellschüler beurteilen ihre Mitschüler ein wenig positiver.*

### **Skala 3: Störtendenzen gegen ein gutes Klassenklima**

*[Lesehilfe: Je kleiner die Mittelwerte, desto positiver die Einschätzung]*

**Tabelle 5: SVG-4.1-Skala 3 (Störtendenzen) – vier Vergleichsansätze**

Faktor 3 (Skala 3)	<i>Vergleichsansatz 1: A1-Modell vs. A2-Modell vs. B-Modell vs. Vergleichsschulen</i>					
<b>Modell</b>	A1	A2	B	V	Signifikanz der Modellunter- schiede, alle Modelle	Signifikanz der Modellunter- schiede, einzel- ne Modelle
<b>Anzahl</b>	178	131	125	220		
<b>Mittel- werte</b>	10,17	9,80	9,70	9,68	0,048 *	A1 > B *, A1 > V *

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau,  
(ns.) = nicht signifikant, --- = keine signifikanten Unterschiede zwischen den einzelnen Modellen



Faktor 3 (Skala 3)	<i>Vergleichsansatz 2: A1-Modell+ A2-Modell vs. B-Modell+Vergleichsschulen</i>			
<b>Modell</b>	A1 + A2	B + V	Signifikanz der Modellunter- schiede, alle Modelle	Signifikanz der Modellunter- schiede, einzel- ne Modelle
<b>Anzahl</b>	309	345		
<b>Mittel- werte</b>	10,01	9,69	0,03 *	A1 + A2 > B+ V

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau,  
(ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 3 (Skala 3)	<i>Vergleichsansatz 3: A1-Modell+ A2-Modell vs. B-Modell vs. Vergleichsschulen</i>				
<b>Modell</b>	A1 + A2	B	V	Signifikanz der Modellunter- schiede, alle Modelle	Signifikanz der Modellunter- schiede, einzel- ne Modelle
<b>Anzahl</b>	309	125	220		
<b>Mittel- werte</b>	10,01	9,70	9,68	0,09 (ns.)	A1 + A2 > B *,

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau,  
(ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 3 (Skala 3)	<i>Vergleichsansatz 4: A1-Modell+ A2-Modell+ B-Modell vs. Vergleichsschulen</i>			
<b>Modell</b>	A1 + A2 + B	V	Signifikanz der Modellunter- schiede, alle Modelle	Signifikanz der Modellunter- schiede, einzel- ne Modelle
<b>Anzahl</b>	434	220		
<b>Mittel- werte</b>	9,92	9,68	0,12 (ns)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau,  
(ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Beim generellen Modellvergleich sind bei den Vergleichsansätzen 3 und 4 keine signifikanten Unterschiede festzustellen, wohl aber bei den Vergleichsansätzen 1 und 2.

Die Post-hoc-Tests geben darüber Aufschluss, wie es dazu kommt: Die A1-Modellschüler geben etwas mehr Störtendenzen an, bei den zusammengefassten Werten (A-Modelle) gehen die signifikanten Unterschiede ebenfalls auf den etwas höheren Wert des A1-Modells zurück.

Fazit:

Die Einschätzung der Schülerinnen und Schüler der vier Modelle unterscheidet sich bei drei Modellen nur unwesentlich. Zieht man zusätzlich noch die Post-hoc-Tests zu Rate, so geben die A1-Modellschüler eine etwas höhere Störrate an.

*Skala 4: Konfliktlösungsansatz: Soziale Kompetenz [Lesehilfe: Je größer die Mittelwerte, desto positiver die Einschätzung]*

**Tabelle 6: SVG-4.1-Skala 4 (Soziale Kompetenz) – vier Vergleichsansätze**

Faktor 4 (Skala 4)	<i>Vergleichsansatz 1: A1-Modell vs. A2-Modell vs. B-Modell vs. Vergleichsschulen</i>					
Modell	A1	A2	B	V	Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	188	136	125	227		
Mittelwerte	7,07	7,53	7,16	7,31		

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 4 (Skala 4)	<i>Vergleichsansatz 2: A1-Modell+ A2-Modell vs. B-Modell+Vergleichsschulen</i>					
Modell	A1 + A2		B + V		Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	320		352			
Mittelwerte	7,23		7,26			

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 4 (Skala 4)	<i>Vergleichsansatz 3: A1-Modell+ A2-Modell vs. B-Modell vs. Vergleichsschulen</i>					
Modell	A1 + A2		B	V	Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	320		125	227		
Mittelwerte	7,23		7,16	7,31		

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 4 (Skala 4)	<i>Vergleichsansatz 4: A1-Modell+ A2-Modell+ B-Modell vs. Vergleichsschulen</i>					
Modell	A1 + A2 + B		V		Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	445		227			
Mittelwerte	7,21		7,31			

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Beim generellen Modellvergleich sind bei allen vier Vergleichsansätzen keine signifikanten Unterschiede festzustellen. Auch die Post-hoc-Tests zeigen keine signifikanten Unterschiede.

Fazit:

Der Modellansatz hat keinen erkennbaren Einfluss auf das Bevorzugen von Konfliktlösungen seitens der Schülerinnen/Schüler, die man dem Bereich „Soziale Kompetenz“ zuordnen kann.

*Skala 5: Konfliktlösungsansatz: Aggressivität [Lesehilfe: Je kleiner die Mittelwerte, desto positiver die Einschätzung]*

**Tabelle 7: SVG-4.1-Skala 5 (Aggressivität) – vier Vergleichsansätze**

Faktor 5 (Skala 5)	<i>Vergleichsansatz 1: A1-Modell vs. A2-Modell vs. B-Modell vs. Vergleichsschulen</i>					
Modell	A1	A2	B	V	Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	188	136	125	227		
Mittelwerte	1,10	,93	1,26	1,14		

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 5 (Skala 5)	<i>Vergleichsansatz 2: A1-Modell+ A2-Modell vs. B-Modell+Vergleichsschulen</i>					
Modell	A1 + A2		B + V		Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	320		352			
Mittelwerte	1,04		1,18			

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 5 (Skala 5)	<i>Vergleichsansatz 3: A1-Modell+ A2-Modell vs. B-Modell vs. Vergleichsschulen</i>				
Modell	A1 + A2	B	V	Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	320	125	227		
Mittelwerte	1,04	1,26	1,14		

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Faktor 5 (Skala 5)	<i>Vergleichsansatz 4: A1-Modell+ A2-Modell+ B-Modell vs. Vergleichsschulen</i>			
<b>Modell</b>	A1 + A2 + B	V	Signifikanz der Modellunter- schiede, alle Modelle	Signifikanz der Modellunter- schiede, einzel- ne Modelle
<b>Anzahl</b>	445	227		
<b>Mittel- werte</b>	1,10	1,14	0,80 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau,  
(ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Beim generellen Modellvergleich sind bei allen vier Vergleichsansätzen keine signifikanten Unterschiede festzustellen.

Auch bei den Post-hoc-Tests sind keine signifikanten Unterschiede festzustellen.

Fazit:

Der Modellansatz hat keinen erkennbaren Einfluss auf das Bevorzugen von Konfliktlösungen seitens der Schülerinnen/Schüler, die man dem Bereich „(Soziale) Aggressivität“ zuordnen kann.

*Testgesamtwert (z-Werte): Schul-Erleben und Sozialkompetenz*

**Tabelle 8: SVG-4.1-Testgesamtwert – vier Querschnitts-Vergleichsansätze**

*[Lesehilfe: Je kleiner die gemittelten z-Werte, desto positiver die Einschätzung]*

Testgesamtwert (z-Werte)						
<b>Modell</b>	A1	A2	B	V	Signifikanz der Modellunter- schiede, alle Modelle	Signifikanz der Modellunter- schiede, einzel- ne Modelle
<b>Anzahl</b>	167	128	122	214		
<b>Mittel- werte</b>	,0559004	-,1033553	-,0160842	-,0138347	,327 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau,  
(ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Testgesamtwert (z-Werte)						
<b>Modell</b>	A1 + A2		B + V		Signifikanz der Modellunter- schiede, alle Modelle	Signifikanz der Modellunter- schiede, einzel- ne Modelle
<b>Anzahl</b>	295		336			
<b>Mittel- werte</b>	-,0120316		-,0062734		,916 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau,  
(ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Testgesamtwert (z-Werte)					
Modell	A1 + A2	B	V	Signifikanz der Modellunter-schiede, alle Modelle	Signifikanz der Modellunter-schiede, einzel-ne Modelle
Anzahl	295	122	214		
Mittel-werte	-,0120316	,0173651	-,0197495	,887 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Testgesamtwert (z-Werte)				
Modell	A1 + A2 + B	V	Signifikanz der Modellunter-schiede, alle Modelle	Signifikanz der Modellunter-schiede, einzel-ne Modelle
Anzahl	417	214		
Mittel-werte	-,0034311	-,0197495	0,777 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Beim generellen Modellvergleich mittels des Testgesamtwerts sind bei allen vier Vergleichsansätzen keine signifikanten Unterschiede festzustellen.

Auch bei den Post-hoc-Tests konnten keine signifikanten Einzelvergleiche aufgezeigt werden.

Fazit:

Der Modellansatz hat keinen erkennbaren Einfluss auf die Ausprägung eines Testgesamtwerts, der als „Schul-Erleben und Sozialkompetenz“ benannt werden kann.

### Längsschnittvergleiche

In die Längsschnittvergleiche werden als Kovariaten die drei SVG Untersuchungen in Klasse 1 (Anfang und Ende) und Klasse 2 (Ende) eingebracht<sup>34</sup>.

**Tabelle 9: SVG-4.1-Testgesamtwert – vier Längsschnitt-Vergleichsansätze**

*[Lesehilfe: Je kleiner die gemittelten z-Werte, desto positiver die Einschätzung]*

Testgesamtwert (z-Werte)						
Modell	A1	A2	B	V	Signifikanz der Modellunterschiede, alle Modelle <sup>35</sup>	Signifikanz der Modellunterschiede, einzelne Modelle <sup>36</sup>
Anzahl	134	40	90	171		
Mittelwerte	,028	-,054	,023	-,049	,695 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Testgesamtwert (z-Werte)						
Modell	A1 + A2		B + V		Signifikanz der Modellunterschiede, alle Modelle <sup>37</sup>	Signifikanz der Modellunterschiede, einzelne Modelle <sup>38</sup>
Anzahl	174		261			
Mittelwerte	,010		-,024		,599 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Testgesamtwert (z-Werte)						
Modell	A1 + A2		B	V	Signifikanz der Modellunterschiede, alle Modelle <sup>39</sup>	Signifikanz der Modellunterschiede, einzelne Modelle <sup>40</sup>
Anzahl	174		90	171		
Mittelwerte	,010		,024	-,050	,610 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

<sup>34</sup> Wie fast immer bei Längsschnitten reduzieren sich die verwertbaren Fallzahlen. Da aber keine erkennbaren systematischen Ausfallgründe zu erkennen sind, kann das Ergebnis dennoch zu Interpretationszwecken verwendet werden.

<sup>35</sup> Die Kovariaten im Modell werden anhand der folgenden Werte berechnet: SVGRW1 SVG-11Items-Rohwert (Jan 00) = 9,46, SVGRW2 SVG-11Items-Rohwert (Jul 00) = 9,46, SVGRW3 SVG-11Items-Rohwert (Jun-Jul 01) = 9,99.

<sup>36</sup> Basiert auf geschätzten Randmitteln

<sup>37</sup> siehe Fußnote 35

<sup>38</sup> siehe Fußnote 36

<sup>39</sup> siehe Fußnote 35

<sup>40</sup> siehe Fußnote 36

Testgesamtwert (z-Werte)				
Modell	A1 + A2 + B	V	Signifikanz der Modellunter-schiede, alle Modelle <sup>41</sup>	Signifikanz der Modellunter-schiede, einzel-ne Modelle <sup>42</sup>
Anzahl	264	171		
Mittel-werte	,014	-,050	,326 (ns.)	- - -

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede zwischen den einzelnen Modellen

Beim generellen Modellvergleich sind auch im Rahmen des kovarianzanalytischen Ansatzes bei allen vier Vergleichsansätzen keine signifikanten Unterschiede zu erkennen.

Bei den Post-hoc-Tests wurden keine signifikanten Unterschiede festgestellt.

Fazit:

Die unterschiedlichen Modellansätze haben im Längsschnitt keinen erkennbaren Einfluss auf die Entwicklung des Sozialverhaltens und der positiven Einschätzung der Schule, der Lehrerinnen und Lehrer und der Mitschüler.

#### Zusammenhang zwischen einigen Merkmalen und dem SVG- 4.1

In diesem Teilkapitel sollen einige Ergebnisse dargestellt werden, die durch Aufbrechen der Gesamtgruppe der Schülerinnen und Schüler nach bestimmten Merkmalen erzielt wurden.

Damit ist eine doppelte Absicht verbunden: Zum einen soll ein kleiner Überblick über die Einstellungen der Schülerinnen und Schüler am Ende ihrer Grundschule zu ihrem schulischen Umfeld und zu Lösungsansätzen bei sozialen Konflikten gegeben werden, zum anderen können so erste Hinweise zur Validität des SVG- 4.1 gegeben werden.

---

<sup>41</sup> siehe Fußnote 35

<sup>42</sup> siehe Fußnote 36

**Merkmale „Notendurchschnitt Kl. 4“ und „Bildungsempfehlung“**

Merkmale „Notendurchschnitt“

**Tabelle 10: Zusammenhang von SVG-4.1-Skalen und der gemittelten Note aus Deutsch und Mathematik, Ende Kl. 4**

SVG4.1-Skalen	Standardisierte Koeffizienten (Beta)	Signifikanz	
SVG4.1-Skala 1 (Schule, Lehrer, Unterricht)	,038	,446 (ns.)	
SVG4.1-Skala 2 (Mitschüler)	,053	294 (ns.)	
SVG4.1-Skala 3 (Störtendenzen)	-,045	,332 (ns.)	
SVG4.1-Skala 4 (soziale Kompetenz) <sup>43</sup>	,264	,000**	
SVG4.1-Skala 5 (Aggressivität)	-,130	,013*	
<b>Ergebnis der Regressionsberechnung</b> (abhängige Variable: gemittelte Note aus D und M, Ende Kl. 4)			
Koeffizienten	R	Korrigiertes R-Quadrat	Signifikanz
<b>Werte</b>	,244	,051	,000**

Legende: \* = Koeffizient signifikant 5%-Niveau, \*\* = Koeffizient signifikant 1%-Niveau, (ns.) = nicht signifikant

Der Zusammenhang ist nicht sehr eng, aber gegen den Zufall solide abgesichert. Ein Blick auf die standardisierten Beta-Gewichte zeigt, dass die Skala 4 (soziale Kompetenz) und die Skala 5 (Aggressivität als Konfliktlösungsansatz) die „gewichtigen“ Beiträge stellen. Allerdings mit verschiedener Polung: je besser (also numerisch kleiner) die Endnote, desto wahrscheinlicher ist eine stärkere Bevorzugung von Konfliktlösungen, die als „sozial kompetent“ bezeichnet werden können. Bei der „Aggressivität“ ist der Zusammenhang umgekehrt.

Bemerkenswert ist, dass die Einstellungen zur Schule, zu Lehrern und Mitschülern keinen wesentlichen Zusammenhang mit der Schulleistung Ende Kl. 4 zeigen.

<sup>43</sup> invertierte Skala: je kleiner der Wert, desto höher die soziale Kompetenz



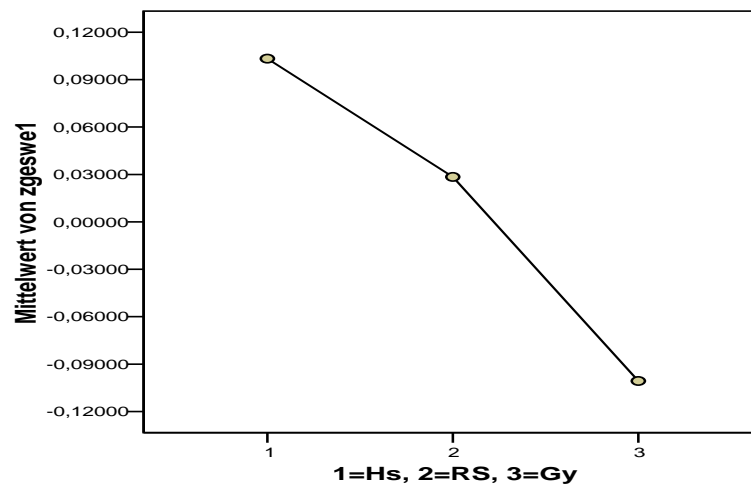
Merkmal „Bildungsempfehlung“

**Tabelle 11: SVG-4.1-Testgesamtwert, aufgebrochen nach Bildungsempfehlung**

[Lesehilfe: Je kleiner die gemittelten z-Werte, desto positiver die Einschätzung]

Testgesamtwert (z-Werte)					
Bildungsempfehlung	Hauptschule	Realschule	Gymnasium	Signifikanz der Modellunterschiede, alle Modelle	Signifikanz der Modellunterschiede, einzelne Modelle
Anzahl	295	122	214		
Mittelwerte	,1032489	,0284053	-,1006799	,014 *	HS>Gy**, RS>Gy#

Legende: \* = Unterschiede signifikant 5%-Niveau, \*\* = Unterschiede signifikant 1%-Niveau, # = Unterschiede signifikant 10%-Niveau, (ns.) = nicht signifikant, - - - = keine signifikanten Unterschiede bei den einzelnen Modellen



**Abbildung 3: SVG-4.1-Testgesamtwert (zgeswe1), aufgebrochen nach der Bildungsempfehlung**

[Lesehilfe: Je kleiner die gemittelten z-Werte, desto positiver die Einschätzung]

Es besteht ein signifikanter Zusammenhang zwischen der Bildungsempfehlung am Ende der Grundschulzeit und dem Testgesamtwert des SVG-4.1.

Aus Abbildung 4 ist ersichtlich, dass es einen fast linearen Zusammenhang zwischen der von der Grundschule empfohlenen weiterführenden Schulart und der Höhe des durchschnittlichen Testgesamtwerts innerhalb der Schülergruppierungen nach der Bildungsempfehlung gibt: Die HS-Empfohlenen haben den höchsten, sprich ungünstigsten, die gymnasial Empfohlenen den niedrigsten, sprich günstigsten Testgesamtwert.

Die auf der nächsten Seite gezeigte multiple Regression zeigt die Gewichtung, mit der die einzelnen Skalen in die Berechnung eingehen.

**Tabelle 12: Zusammenhang von SVG4.1-Skalen und der Bildungsempfehlung**

[Lesehilfe: Die Kodierung der Bildungsempfehlung ist HS=1, RS=2, Gy=3; negative Beta-Gewichte weisen auf einen Zusammenhang hin, der zu Gunsten der Schülergruppe mit gymnasialen Bildungsempfehlung ausfällt.]

SVG-4.1-Skalen	Standardisierte Koeffizienten (Beta)	Signifikanz	
SVG4.1-Skala 1 (Schule, Lehrer, Unterricht)	-,051	,301 (ns.)	
SVG4.1-Skala 2 (Mitschüler)	-,021	,668 (ns.)	
SVG4.1-Skala 3 (Störtendenzen)	,046	,312 (ns.)	
SVG4.1-Skala 4 (soziale Kompetenz) <sup>44</sup>	-,235	,000**	
SVG4.1-Skala 5 (Aggressivität)	,096	,063 (ns.)	
<b>Ergebnis der Regressionsberechnung</b> (abhängige Variable: gemittelte Note aus D und M, Ende Kl. 4)			
Koeffizienten	R	Korrigiertes R-Quadrat	Signifikanz
Werte	,216	,038	,000**

Legende: \* = Koeffizient signifikant 5%-Niveau, \*\* = Koeffizient signifikant 1%-Niveau, (ns.) = nicht signifikant

Der Zusammenhang ist nicht sehr eng, aber ebenfalls gegen den Zufall solide abgesichert. Ein Blick auf die standardisierten Beta-Gewichte zeigt, dass die Skala 4 (soziale Kompetenz) und erheblich weniger deutlich die Skala 5 (Aggressivität als Konfliktlösungsansatz) die „gewichtigen“ Beiträge stellen. Allerdings mit verschiedener Polung: Je „höher“ die Bildungsempfehlung für die einzelnen Schülerinnen und Schüler ausfällt, desto wahrscheinlicher ist eine stärkere Bevorzugung von Konfliktlösungen, die als „sozial kompetent“ bezeichnet werden können. Bei der „Aggressivität“ ist der Zusammenhang umgekehrt.

Zu vermerken ist auch hier, dass die Einstellungen zur Schule, zu Lehrern und Mitschülern keinen wesentlichen Zusammenhang mit der Schulleistung bzw. Bildungsempfehlung Ende Kl. 4 zeigen.

Fazit:

Nach Untersuchungen, die sich mit Formen der Gewalt und Aggression an Schulen beschäftigen<sup>45</sup>, muss erwartet werden, dass sich die Konfliktlösungen der Schülerinnen und Schüler der drei weiterführenden Schularten im Schnitt graduell unterscheiden. Das hier dargestellte Ergebnis ist demnach erwartungsgemäß und ein erster Hinweis auf die Validität des SVG-4.1.

Im Übrigen sollte nach Meinung des Autors der Zusammenhang von Schulleistung bzw. Bildungsempfehlung und Einstellungen in Bezug auf den sozialen Interaktionsraum Schule nicht kausal gesehen werden („gute“ Einstellungsausprägung bedingen „gute“ Noten

<sup>44</sup> invertierte Skala: je kleiner der Wert, desto höher die soziale Kompetenz

<sup>45</sup> z.B. Sikorski & Thiel: Gewalt an Schulen: Ergebnisse einer Befragung von Schulleitern in Baden-Württemberg zur Gewaltentwicklung und zu möglichen Interventionsmaßnahmen. Zusammenfassender Bericht. Landesinstitut für Erziehung und Unterricht, LEU-Bericht 95/3, 1995.

und vice versa), sondern als parallel verlaufende Äußerungen einer dahinter liegenden individuellen Persönlichkeits- und Sozialisationsstruktur von Schülerinnen und Schülern.

**Merkmale „Geschlecht“ und „Nationalität“**

**Tabelle 13: Merkmale „Geschlecht“ und „Nationalität“ und die SVG 4 - Skalen**  
*[Lesehilfe: Je kleiner die gemittelten z-Werte, desto positiver die Einschätzung]*

Variable (z-Werte)	Merkmal	Signifikanz	Richtung
Testgesamtwert (Schul-Erleben und Sozialkompetenz)	<i>Geschlecht</i>	,005**	♀ > ♂
	Nationalität	,833 (ns.)	
SVG4.1-Skala1 (Schule, Lehrer, Unterricht)	<i>Geschlecht</i>	,495 (ns.)	
	Nationalität	,067 (ns.)	
SVG4.1-Skala2 (Mitschüler)	<i>Geschlecht</i>	,197 (ns.)	
	Nationalität	,675 (ns.)	
SVG4.1-Skala 3 (Störtendenzen)	<i>Geschlecht</i>	,162 (ns.)	
	Nationalität	,642 (ns.)	
SVG4.1-Skala 4 (soziale Kompetenz) <sup>46</sup>	<i>Geschlecht</i>	,000**	♀ > ♂
	Nationalität	,094 (ns.)	
SVG4.1-Skala 5 (Aggressivität)	<i>Geschlecht</i>	,000**	♀ > ♂
	Nationalität	,841 (ns.)	

Legende: \* = Koeffizient signifikant 5%-Niveau, \*\* = Koeffizient signifikant 1%-Niveau,  
 (ns.) = nicht signifikant

Beim Merkmal „Geschlecht“ konnten auf zwei Skalen und beim Gesamtwert signifikante Unterschiede aufgezeigt werden. Jungen zeigen demnach im Schnitt ein etwas negatives Schul-Erleben und eine etwas negativere Sozialkompetenz. Letztere wird noch durch die signifikanten Unterschiede bei den Konfliktlösungen: „Soziale Kompetenz“ und „Aggressivität“ präzisiert: Jungen tendieren weniger als Mädchen zu sozial kompetenten Konfliktlösungen und mehr als Mädchen zu solchen, denen aggressive Momente zugemischt sind.

Beim Merkmal „Nationalität“ haben sich keinerlei signifikante Unterschiede zwischen deutschen und nicht-deutschen Schülerinnen und Schülern ergeben.

**Fazit:**

Geschlechtsunterschiede bei den bevorzugten Konfliktlösungsstrategien waren zu erwarten und weisen ebenfalls auf die Validität des SVG-4.1 hin.

Nicht ganz erwartungsgemäß fällt das Ergebnis aus, dass sich zwischen den deutschen und nicht-deutschen Schülerinnen und Schülern keinerlei signifikanten Unterschiede ergeben haben. Immerhin deutet es darauf hin, dass bei den untersuchten Schulen die Integrationsbemühungen, zumindest was den Einstellungsbereich der Schülerinnen und Schüler betrifft, anscheinend zum Erfolg geführt haben.

<sup>46</sup> invertierte Skala: je kleiner der Wert, desto höher die soziale Kompetenz

### *Zusammenfassende Diskussion über die Ergebnisse der Modellvergleiche*

Bei den Querschnittsvergleichen, unter Einbeziehung der fünf SVG-4.1-Skalen und des Testgesamtwerts, ergeben sich sowohl bei den generellen Gruppenvergleichen (Mehrfachvergleichen) als auch bei den Einzelvergleichen bei Anwendung der Post-hoc-Tests nur wenige signifikante Modellunterschiede<sup>47</sup>.

Bei den Längsschnittvergleichen, die nur mit dem Testgesamtwert durchgeführt wurden, sind sowohl alle Mehrfachvergleiche als auch alle Einzelvergleiche nicht signifikant.

Man muss daher von einer im hohen Maße modellunabhängigen Förderung des Sozialverhaltens bzw. der Entwicklung der Einstellung zur Schule ausgehen.

Dennoch sollte nicht unerwähnt bleiben, dass die Lehrerinnen und Lehrer der A-Modelle mit einer etwas schwierigeren Ausgangslage konfrontiert wurden: In ihren Eingangsklassen befanden sich deutlich mehr Schülerinnen und Schüler<sup>48</sup>, die in den B-Modellen und in den Vergleichsschulen vom Schulbesuch zurückgestellt worden wären; statistisch gesehen hatten ihre Eingangsklassen auch einen signifikant höheren Ausländeranteil und einen deutlich höheren Anteil an vorzeitig eingeschulerten Kindern<sup>49</sup>.

Die durch die Begleituntersuchung vorgelegten Ergebnisse bescheinigen den Lehrerinnen und Lehrern der A-Modelle, dass sie trotz dieser unterschiedlichen Ausgangssituation im Bereich der Formung von Einstellungen und Konfliktlösungsstrategien vergleichbare pädagogische Resultate erzielt haben wie ihre Kolleginnen und Kollegen aus den anderen beiden Modellen. Dieser Befund reiht sich in die Folge von Erkenntnissen ein, die sich bei der Untersuchung scheinbar noch extremerer pädagogischer Herausforderungen, wie der Integration von Behinderten, gewinnen ließen. Deren Fazit lässt sich einfach so beschreiben: Die Möglichkeiten, pädagogisch mit scheinbar erschwerten sozialen Situationskonstellationen umzugehen, sind häufig wesentlich günstiger als viele Eltern oder Lehrerinnen und Lehrer annehmen.

---

<sup>47</sup> Bei den generellen Vergleichen: ca. 8% und bei den Einzelvergleichen bei Anwendung der Post-hoc-Tests ca.9%.

<sup>48</sup> Die Schülerzusammensetzung der Eingangsklassen der A-Modelle auf der einen und der B-Modelle + Vergleichsschulen auf der anderen Seite nach dem Merkmal „Einschulungsmodalität“ ist signifikant unterschiedlich.

<sup>49</sup> dito

## II.6 Schulische Leistungen

### Schulische Leistungen gemessen mit standardisierten Verfahren

Die Vermittlung der drei Kulturtechniken Lesen, Schreiben und Rechnen wird immer wieder als das zentrale Ziel des Grundschulunterrichts dargestellt. Daher sollten auch die direkt in diesen Bereichen erworbenen Fertigkeiten bei der Evaluation des neu gestalteten Anfangsunterrichts betrachtet werden. Neben den Schulnoten in den Fächern Deutsch und Mathematik am Ende der zweiten Klasse kamen für diesen Aspekt der Evaluation pädagogisch-psychologische Testverfahren in Frage. Bei der Auswahl der Verfahren zur Erfassung der basalen Fertigkeiten in diesen Bereichen fanden drei Gesichtspunkte besondere Berücksichtigung: Erstens sollten möglichst normierte Verfahren herangezogen werden, die auch längsschnittlich zumindest in den ersten Grundschuljahren eingesetzt werden können; zweitens sollten die berücksichtigten Fertigkeitstests eine differenzierte Erfassung besonders leistungsstarker Schüler ermöglichen; und drittens sollte auch eine Differenzierung im unteren Leistungsbereich möglich sein.

#### *Eingesetzte Verfahren und Ergebnisse in der Kohorte 1 (8 Intensivschulen)*

In der Intensiverhebung wurde im Verlauf des zweiten Schuljahres (1998/99) für die Kinder zunächst lediglich das Leistungsniveau im Bereich des Lesens und Schreibens erfasst. Während die Rechtschreibleistung am Ende des ersten Schulhalbjahres der zweiten Klasse über den „Weingartener Grundwortschatz Rechtschreib-Test für 1. und 2. Klassen“ (WRT 1+ von Birkel, 1995) erfasst wurde, wurde die Leseleistung mit Hilfe des differenzierten Gruppenlesetests KNUSPEL-L (Marx, 1998) am Ende der zweiten Klassenstufe erhoben. Dieser Test enthält vier Subtests, von denen der erste das Hörverstehen und die drei folgenden spezifische Lesefertigkeiten (Rekodieren, Dekodieren und Leseverstehen) messen.

Das *Hörverstehen* (Subtest 1) ist operationalisiert als „Verstehen mündlich gestellter Fragen und Aufforderungen“, das *Leseverstehen* (Subtest 4) als „Verstehen schriftlich gestellter Fragen und Anweisungen“. Der Inhalt und die Aspekte der jeweils 14 Fragen beider Subtests korrespondieren weitgehend miteinander. Die Fragen beziehen sich entweder auf das Fabelwesen „Knuspel“ (Haben die Knuspel glatte oder lockige Haare?; Wie viele Finger haben die Knuspel an einer Hand? etc.) oder auf das Kind selber (Geburtsdatum, Anfangsbuchstaben des Namens etc.), wobei im vorderen Teil der Frage immer der Wissensaspekt und im hinteren Teil der Ausführungsaspekt angesprochen und mit einem oder zwei Rohpunkten bewertet wird (gesamt je 17 Rohpunkte für Wissens- und Ausführungsaspekt).

Das *Rekodieren* (Subtest 2) wird über die Vorgabe von 40 orthographisch korrekt geschriebenen, sinnvollen Wortpaaren (z.B. „Meer-mehr“, „Kiepe-Kippe“), die auf Homophonie zu testen sind, überprüft (gesamt 40 Rohpunkte).

Das *Dekodieren* (Subtest 3) wird über die Präsentation von 40 Pseudowörtern (echte Nachnamenwörter, z.B. „KNAPE“, „DOHSE“), die nach erfolgter Rekodierung auf ihren semantischen Gehalt hin zu prüfen sind, festgestellt (gesamt 40 Rohpunkte).

In einem ersten Auswertungsschritt wurde zunächst geprüft, ob es bedeutsame Leistungsunterschiede zwischen Erprobungs- und Vergleichsklassen im Lesen und Rechtschreiben gibt. Die für diese Überprüfung relevanten Mittelwerte und T-Werte sind aus Tabelle 2 auf der nächsten Seite zu entnehmen. Obwohl bei allen vier Subskalen des Lesetests und auch bei den Zeugnisnoten in den Fächern Deutsch und Mathematik die Kinder der Erprobungsklassen numerisch etwas bessere Werte als die Kinder der Vergleichsklassen erzielten, unterscheiden sich beide Gruppen nicht statistisch bedeutsam am Ende der

zweiten Klassenstufe. Etwas anders verhält es sich bei der über den WRT 1+ erfassten Rechtschreibleistung. Hier lagen die mittleren Rohwerte in den Erprobungsklassen nicht nur tendenziell, sondern sogar statistisch bedeutsam über denen der Vergleichsklassen.

**Tabelle 14: Mittelwerte (MW) und Standardabweichungen (SW) der vier Subskalen des Gruppenlesetests KNUSPEL-L (T-Werte), der Deutsch- und Mathematiknote am Ende der 2. Klasse und des Rechtschreibtests WRT (Rohwerte), getrennt für die Kinder der Erprobungs- und der Vergleichsklassen sowie die T-Werte zur inferenzstatistischen Prüfung auf Leistungsunterschiede zwischen den beiden Gruppen.**

Variable	Bedingung	N	MW	SW	t-Wert	df	sign.
Hörverstehen (T-Wert)	Erprobungsklasse	74	53.62	10.96	1.22	182	nein
	Vergleichsklasse	110	51.65	10.51			
Rekodierfähigkeit (T-Wert)	Erprobungsklasse	74	46.12	6.58	1.24	181	nein
	Vergleichsklasse	109	44.55	9.49			
Dekodierfähigkeit (T-Wert)	Erprobungsklasse	74	47.46	8.54	.79	182	nein
	Vergleichsklasse	110	46.25	11.05			
Leseverstehen (T-Wert)	Erprobungsklasse	74	52.62	10.30	.12	181	nein
	Vergleichsklasse	109	52.45	9.27			
Zeugnisnote Deutsch	Erprobungsklasse	64	2.34	0.89	.46	175	nein
	Vergleichsklasse	113	2.42	1.08			
Zeugnisnote Mathematik	Erprobungsklasse	64	2.27	0.93	.91	175	nein
	Vergleichsklasse	113	2.41	1.03			
WRT-Rohwert	Erprobungsklasse	74	18.85	5.24	2.15	184	ja
	Vergleichsklasse	112	17.15	5.31			

Auf der Ebene der schulischen Leistungen in den Bereichen des Lesens und des Rechtschreibens sowie hinsichtlich der Zeugnisnoten in den Fächern Deutsch und Mathematik ließ sich in keiner Weise ein Nachteil der Erprobungsklassen feststellen. Im Gegenteil: In allen Leistungsbereichen zeigte sich tendenziell eher eine Überlegenheit der Kinder, die eine Erprobungsklasse im A-Modell besucht hatte. Hinsichtlich der Leistungen des etwa Mitte der zweiten Klasse erhobenen Rechtschreibtests fiel diese Überlegenheit sogar nicht nur tendenziell aus, sondern statistisch bedeutsam.

Somit konnte bei den 8 Intensivschulen der Kohorte 1 das Fazit lauten:

Bezüglich der schulischen Leistungen hat sich somit zumindest die von uns hier evaluierte A-Modell-Variante des „Schulanfangs auf neuen Wegen“ voll bewährt<sup>50</sup>.

### *Eingesetzte Verfahren und Ergebnisse in der Kohorte 3 (41 Versuchsschulen)*

#### Verfahren zur Messung der Lese- und Rechtschreibleistung

Im Verlauf des zweiten Schuljahres (2000/01) und am Ende des vierten Schuljahres (Juni 2003) wurde für die Kinder zunächst lediglich das Leistungsniveau im Bereich des Lesens und Schreibens erfasst. Während die Rechtschreibleistung zu Beginn des zweiten Schulhalbjahres der zweiten Klasse über den „Weingartener Grundwortschatz Rechtschreib-Test für 1. und 2. Klassen“ (WRT 1+, Birkel, 1995<sup>51</sup>) und am Ende der vierten Klassenstufe mit dem entsprechenden Verfahren für 3. und 4. Klassen (WRT 3+, Birkel, 1994) erfasst wurde, wurde die Leseleistung mit Hilfe der „Würzburger Leise Leseprobe“ (WLLP, Küspert & Schneider, 1998) am Ende der zweiten Klassenstufe und am Ende der vierten Klassenstufe erhoben. Dieser Test ermöglicht die ökonomische Erfassung der Leseleistung im gesamten Grundschulbereich im Rahmen einer Gruppensitzung. Er ist als Speed-Test konstruiert und misst die Dekodier(=Lese-)geschwindigkeit, indem geschriebenen Wörtern jeweils vier Bildalternativen gegenübergestellt sind und das korrespondierende Bild anzustreichen ist. Mit den Testheften der Form A und B liegen Pseudo-Parallelformen mit jeweils 140 Items vor. Am Ende der vierten Klassenstufe wurde zusätzlich das Leseverständnis über den entsprechenden Subtest des Hamburger Lesetests für 3. und 4. Klassen (HAMLET 3-4, Lehmann, Peek & Poerschke, 1997) erhoben. Bei diesem Leseverständnistest wird das sinnerschließende, stille Lesen geprüft, wobei ein breites Spektrum von Textformen (Erzähltexte, Sachtexte, Gebrauchstexte) berücksichtigt wird.

Im Vergleich zu den Berechnungen bei den 8 Intensivschulen der Kohorte 1 wurde ein verfeinertes methodisches Vorgehen gewählt. Die Mittelwerte der Modelle werden mittels kovarianzanalytischer Techniken um wichtige Eingangsvoraussetzungen der Schüler, die beiden Untertests des Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC; Jansen, Mannhaupt, Marx & Skowronek, 1996 bzw. 1999) „Reimen“ und „Silben segmentieren“ korrigiert (geschätzt). Somit können Mittelwertsunterschiede der Modelle zum großen Teil auf modellspezifische Treatmenteffekte bezogen werden.

---

<sup>50</sup> Weitere Auswertungen zur Frage, inwiefern sich die interindividuellen Leistungsunterschiede im Lesen und Schreiben sowie in den Fächern Deutsch und Mathematik durch die zu Beginn der ersten und zu Beginn der zweiten Klasse in Einzeluntersuchungen (bzw. durch Lehrer-Urteil) erfassten Entwicklungsmerkmale bzw. durch das Modell (Erprobungs- vs. Vergleichsklasse) erklären lassen, sind im Tabellenanhang des Zwischenberichts dokumentiert.

<sup>51</sup> Kurzbeschreibung siehe beim Unterkapitel II.3 (8 Intensivschulen)

Ergebnisse zum Rechtschreiben, gemessen mit dem WRT 1+ und WRT 3+ (Kohorte 3, 41 Versuchsschulen)

**Tabelle 15: Mittelwerte<sup>52</sup> (in Klammern Standardabweichungen) der Leistungen im Rechtschreiben (WRT 1+, März 01; WRT 3+, Juni 03)**

Gruppierung	WRT 1+	WRT 3+	Anzahl der Fälle
A-Modelle	18.8 (4.65)	44.9 (9.69)	470 / 320
B-Modelle	18.9 (4.55)	43.9 (7.72)	172 / 114
Vergleichsschulen	19.5 (3.97)	43.4 (8.11)	229 / 218

Signifikante Mittelwertunterschiede: nein

Trotz gewisser numerischer Unterschiede der Mittelwerte bei den Modellen sind weder im Leistungsquerschnitt am Ende von Kl. 2 noch im Leistungslängsschnitt von Beginn Kl. 1 bis Ende Kl. 2 die Unterschiede signifikant. D.h., die Schüler der verschiedenen Modelle unterscheiden sich am Ende von Kl. 2 und auch am Ende von Kl. 4 nicht in ihrer durchschnittlichen Rechtschreibleistung und auch nicht beim kovarianzanalytisch geschätzten Leistungszuwachs von Beginn Kl. 1 bis Ende Kl. 2 bzw. Ende Kl. 4.

Ergebnisse zum Lesen, gemessen mit der WLLP (Kohorte 3, 41 Versuchsschulen)

**Tabelle 16: Mittelwerte (in Klammern Standardabweichungen) der Leseleistungen am Ende von Klasse 2 und zu Beginn von Klasse 4 (gemessen über WLLP bzw. HAMLET)**

Gruppierung	Lesetempo Ende Klasse 2	Lesetempo Ende Klasse 4	Leseverständnis Ende Kl. 4
A-Modelle	77.40 (25.50)	111.21 (22.29)	11.94 (2.36)
B-Modelle	79.82 (30.16)	111.47 (21.38)	11.85 (2.57)
Vergleichsschulen	76.03 (23.95)	107.41 (19.16)	11.93 (2.52)

Signifikante Mittelwertunterschiede: kein Unterschied im Lesetempo, Ende Klasse 2 und im Leseverständnis Ende Klasse 4; tendenzieller Vorteil von A und B gegenüber V beim Lesetempo Ende Klasse 4 ( $p < .08$ )

Trotz gewisser numerischer Unterschiede der Mittelwerte bei den Modellen sind weder im Leistungsquerschnitt am Ende von Kl. 2 noch im Leistungslängsschnitt von Beginn Kl. 1

<sup>52</sup> Die aufgeführten Mittelwerte geben die Zahl der durchschnittlich richtig erkannten Wörter an und nicht die Fehlerzahl.



bis Ende Kl. 2 die Unterschiede signifikant. D.h., die Schüler der verschiedenen Modelle unterscheiden sich am Ende von Kl. 2 nicht in ihrer durchschnittlichen Lesegeschwindigkeit und auch nicht beim kovarianzanalytisch geschätzten Leistungszuwachs von Beginn Kl.1 bis Ende Kl.2. Dies ändert sich tendenziell erst zur 4. Klassenstufe. Hier erweist sich die Leseleistung der Kinder aus den A- und den B-Modellen verglichen mit denen aus den Vergleichsschulen als tendenziell überlegen ( $p < .08$ ). Der Effekt ist allerdings sehr klein. Im Hinblick auf das am Ende von Klasse 4 erfasste Leseverständnis zeigten sich überhaupt keine nennenswerten Unterschiede zwischen den drei verglichenen Gruppen.

Zur Leistung der Schüler im Lesen und Rechtschreiben kann folgendes Fazit gezogen werden:

In diesem Bereich sind kaum Unterschiede zwischen den drei Gruppen auszumachen. Eine leichte Überlegenheit im Bereich des Lesetempos ist für die Kinder der A-Modelle und B-Modelle am Ende der Grundschulzeit feststellbar. Weitergehende Analysen zeigen, dass die schriftsprachlichen Schuleingangsvoraussetzungen für die Schülerschaft in den A-Modellen die ungünstigsten Werte aufweisen. Angesichts dessen ist es umso bemerkenswerter, dass für diese Kinder am Ende der vierten Klasse bessere Leseleistungen als in den Vergleichsschulen resultieren.

#### Verfahren zur Messung der Rechenleistung

Zusätzlich zur Lese- und Rechtschreibleistung wurden die Schüler der 41 Versuchsschulen in der Kohorte 3 auf ihre Rechenleistung hin überprüft. Mit der Reihe der Deutschen Mathematiktests (DEMAT) stehen neuerdings curricular valide Tests zur Verfügung, mit deren Hilfe der erreichte Mathematikleistungsstand nach der ersten (DEMAT 1+, Krajewski, Küspert & Schneider, 2002), zweiten (DEMAT 2+, Krajewski, Liehm & Schneider, 2004), dritten (DEMAT 3+, Roick, Gölitz & Hasselhorn, 2004) sowie am Ende der vierten Klassenstufe (DEMAT 4, Gölitz, Roick & Hasselhorn, in Druck) diagnostiziert werden kann.

Die DEMATs sind als Gruppentest mit zwei Parallelformen A und B konzipiert und damit zur ökonomischen Erfassung der Rechenleistung einer gesamten Schulklasse geeignet. Dem Aufbau des Tests liegen die Lehrpläne aller deutschen Bundesländer zu Grunde. Das Anlegen von Schablonen erlaubt eine rasche Auswertung der Testhefte. Durch Punktvergabe auf Subtestebene kann ein Leistungsprofil erstellt werden, das Hinweise auf mögliche Problembereiche gibt.

Das methodische Vorgehen lehnt sich an das bei den sprachlichen Leistungstests (WRT, WLLP) an. Die Mittelwerte der Modelle werden mittels kovarianzanalytischer Techniken korrigiert (geschätzt). Somit können Mittelwertsunterschiede der Modelle zum großen Teil auf modellspezifische Treatmenteffekte bezogen werden.

Bei der folgenden Ergebnisdarstellung wurden die Anfang der vierten Klasse erfassten Werte des DEMAT 3+ unberücksichtigt gelassen, da hier die Normwerte (T-Werte) wegen der etwas verspäteten Durchführung nicht berechnet werden konnten. Eine Analyse der DEMAT 3+ Daten auf der Ebene der Rohwerte ergab darüber hinaus exakt das gleiche Befundmuster, wie es am Ende der vierten Klasse für den DEMAT 4 gefunden wurde und im folgenden Abschnitt auch berichtet wird.

Ergebnisse zum Rechnen, gemessen mit den Verfahren der DEMAT-Reihe (Beginn Kl. 2, Beginn Klasse 3 und Ende von Klasse 4, Kohorte 3, 41 Versuchsschulen)

**Tabelle 17: Mittlere T-Werte (in Klammern Standardabweichungen) der Leistungen in den bundesweit normierten Deutschen Mathematiktests (DEMAT-Reihe)**

<b>Gruppierung</b>	<b>Anfang Klasse 2 (DEMAT 1+)</b>	<b>Anfang Klasse 3 (DEMAT 2+)</b>	<b>Ende Klasse 4 (DEMAT 4)</b>
A-Modelle	49.2 (10.03)	48.6 (9.97)	54.0 (8.60)
B-Modelle	52.9 ( 8.88)	50.6 (9.47)	57.1 (8.45)
Vergleichsschulen	50.0 (10.77)	50.0 (9.19)	53.5 (9.94)

Signifikante Mittelwertunterschiede:

B-Modelle besser als A-Modelle und V-Schulen  
zu Beginn der 2. und am Ende der 4. Klasse

Die in Tabelle 5 angegebenen Werte basieren auf einer bundesweiten T-Normierung. Zu diesem Zwecke wurden die Verfahren von einer großen Anzahl an Schülern in allen 16 Bundesländern bearbeitet. Die Durchschnittsleistung dieser Normierungsstichprobe beträgt 50 (bei einer mittleren Streuung bzw. Standardabweichung von 10).

Die Schüler der A-Modelle und der Vergleichsschulen erreichen nach der 1. Klasse einen signifikant niedrigeren Mittelwert als die der B-Modelle. Das Effektstärkemaß  $\eta^2$  erreicht allerdings mit 0,019 einen Wert, den man nach einer Einteilung von Gediga (1998) nur geringe praktische Bedeutsamkeit zumessen kann.

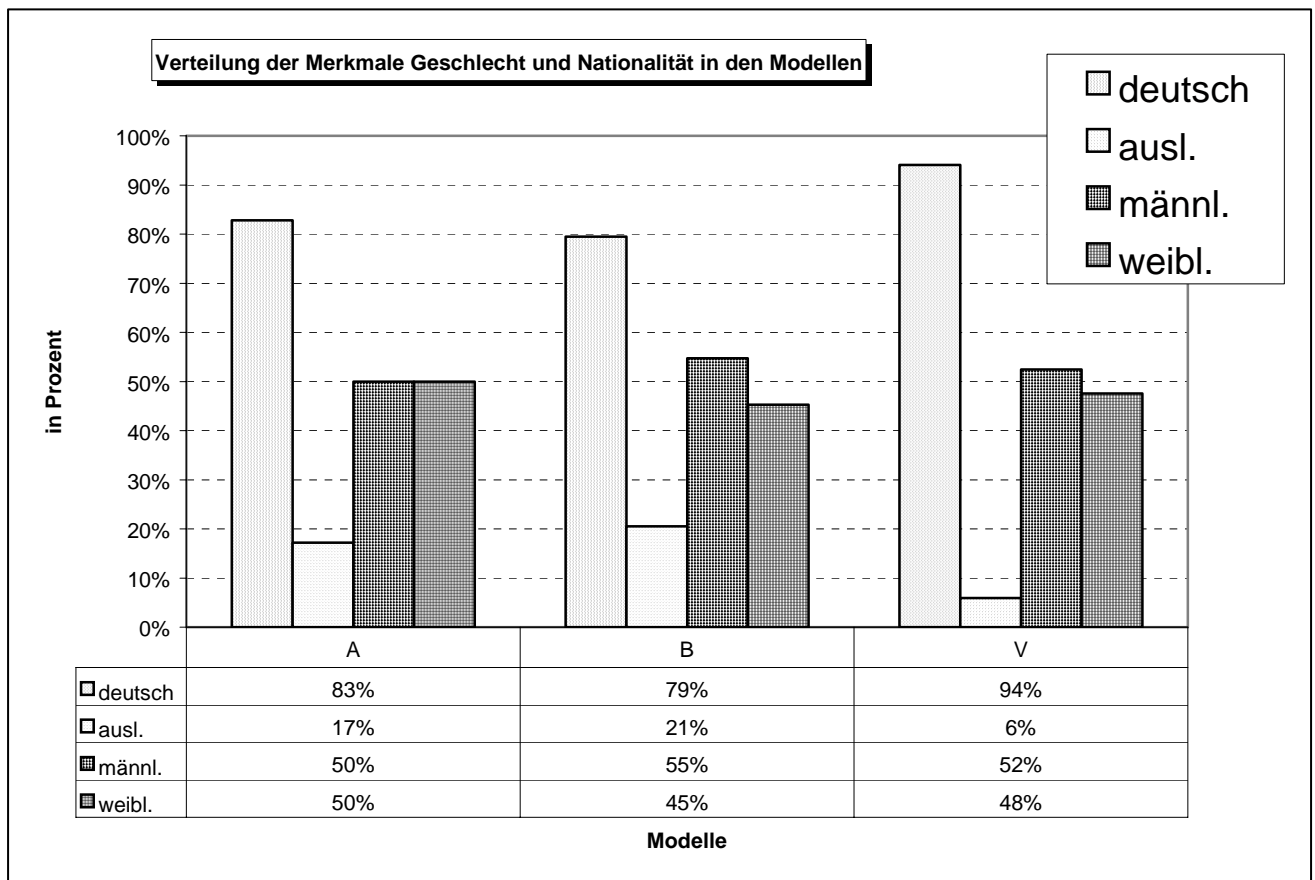
Ein ähnliches Bild zeigt sich auch ein Jahr später beim DEMAT 2+. Die modellspezifischen Leistungsunterschiede haben sich allerdings nicht weiter verschärft. Bringt man das DEMAT 1+ Ergebnis als Kovariate in die Berechnungen mit ein, so stellt sich heraus, dass nach der zweiten Klasse keine modellspezifischen Leistungsunterschiede mehr bestehen, d.h. das pädagogische Treatment scheint in diesem Zeitraum in allen Modellen gleich wirksam gewesen zu sein.

Die Leistungsunterschiede am Ende der 4. Klasse zugunsten der Kinder aus den B-Modellen nehmen jedoch wieder zu. Hier erweist sich der Unterschied auch bei kovarianz-analytischer Auspartialisierung der Werte des DEMAT 1+ als statistisch bedeutsam ( $p < 003$ ). Für die A-Modelle finden sich keine Unterschiede zu den Vergleichsschulen.

Die Tatsache, dass in allen drei Schülergruppen am Ende der vierten Klassenstufe im Mittel ein T-Wert deutlich über 50 (Bundesdurchschnitt, s.o.) erreicht wurde (auch bei den unter 50 gestarteten A-Modell-Schülern) ist ein deutlicher Hinweis auf die gelungene (d.h. bundesweit überdurchschnittliche) Vermittlung mathematischer Grundfertigkeiten und Kenntnisse in allen beteiligten Schulen.

Sucht man nach Erklärungen für das Abschneiden der A-Modelle, so weisen weitergehende Analysen daraufhin, dass es Sinn macht, sich mit der Schülerzusammensetzung der Modelle zu beschäftigen. D.h., die Frage stellt sich, ob es Modellstrukturen sind oder Faktoren, die in der Schülerzusammensetzung liegen, die als mögliche Erklärung für die Rechen- bzw. Mathematikleistung der Schüler herangezogen werden können.

Teilt man die Schülerschaft der Modelle in Gruppierungen, die nach den Merkmalen Geschlecht und Nationalität gebildet werden, so fällt auf, dass der relative Anteil der Gruppierungen je nach Modell schwankt. Dies zeigt die folgende Abbildung.



**Abbildung 4: Relativer Anteil der Schülerschaft der einzelnen Modelle, nach den Merkmalen Geschlecht und Nationalität getrennt ausgewiesen**

Unterschiede zwischen den drei Modellen mittels Chi-Quadrat:

Merkmal Geschlecht:  $\Phi = 0,037$ , nicht signifikant

Merkmal Nationalität:  $\Phi = 0,156$ , sehr signifikant

Aus der Abbildung 4 wird ersichtlich, dass die A-Modelle den höchsten relativen Anteil an weiblichen Schülern und den zweithöchsten relativen Anteil an ausländischen Schülern haben.

Berechnungen zeigen, dass die numerischen Unterschiede beim Merkmal Geschlecht nicht zu einem signifikanten Modellunterschied ausreichen, wohl aber die des Merkmals Nationalität.

Zur Leistung der Schüler im Rechnen (Mathematik) kann folgendes Fazit gezogen werden:

Zwar zeigen die Schüler des A-Modells numerisch eine signifikant niedrigere Durchschnittsleistung im Rechnen gegenüber den B-Modellen, nicht aber gegenüber den Vergleichsschulen; weitergehende Analysen zeigen, dass die Ergebnisse teilweise auf die unterschiedliche Zusammensetzung der Schülerschaft in den Modellen zurückzuführen

ist. Dennoch zeigt sich ein Mathematikvorteil der B-Modelle am Ende der Grundschulzeit. Am Ende von Klasse 4 erreichen alle drei untersuchten Gruppen (A-Modelle, B-Modelle, Vergleichsschulen) im Bundesvergleich überdurchschnittlich gute Leistungen.

### Abschneiden in den landesweiten Vergleichsarbeiten in den Fächern Deutsch und Mathematik am Ende von Kl. 3

Im Juli 2002 wurden in Baden-Württemberg in allen dritten Klassen der Grundschule landesweite Diagnosearbeiten in den Fächern Deutsch und Mathematik durchgeführt. Auch die Kinder unserer wissenschaftlichen Evaluation des „Schulanfangs auf neuen Wegen“ nahmen an diesen Vergleichsarbeiten teil. Über einen Vergleich der Leistungen in diesen Vergleichsarbeiten haben wir die Möglichkeit zu prüfen, ob sich die Teilnahme der Kinder an den A-Modellen oder den B-Modellen in irgendeiner Weise auf das erzielte Leistungsniveau in Deutsch und Mathematik auswirkt.

**Tabelle 18: Mittelwerte (in Klammern Standardabweichungen) der Leistungen der landesweiten Diagnosearbeiten in den Bereichen Deutsch und Mathematik (Ende Klasse 3)**

<b>Gruppierung</b>	<b>Deutsch</b>	<b>Mathematik</b>	<b>Anzahl der Fälle</b>
A-Modelle	26.8 (5.9)	16.4 (6.0)	281
B-Modelle	26.5 (5.9)	18.4 (6.1)	109
Vergleichsschulen	25.7 (5.8)	16.2 (6.2)	216

Signifikante Mittelwertunterschiede:

Sprache: nein

Mathematik: B-Modelle besser als A-Modelle u. Vergleichsschulen

Während für die Leistungen im Bereich Deutsch keinerlei statistisch bedeutsame Unterschiede zwischen den drei Vergleichsgruppen feststellbar waren, zeigte sich ein statistisch bedeutsamer Vorteil der Kinder aus den B-Modellen im Vergleich zu den beiden übrigen Gruppen hinsichtlich der Mathematikleistung am Ende von Klassenstufe 3 ( $p < .006$ ). Dieser Vorteil der B-Modelle blieb auch nach kovarianzanalytischer Kontrolle der Lernvoraussetzungsunterschiede zum Einschulungszeitpunkt im Bereich des mathematischen Denkens (Zahlabstraktion, Zählen von Objekten) bestehen.

Im Bereich der sprachlichen Leistungen lässt sich der Gesamtwert noch weiter ausdifferenzieren, da dieser sich zusammensetzt aus einem Teilwert für Lesen, einem Teilwert für Rechtschreibung (Diktat) und einem Teilwert für Aufsatz. In Tabelle 7 sind daher noch einmal die Mittelwerte, getrennt nach diesen drei Teilbereichen aufgeführt.

Die varianzanalytische Überprüfung auf mögliche Gruppenunterschiede in den Bereichen Lesen, Rechtschreibung und Aufsatz ergab jedoch wiederum keinerlei statistisch bedeutsamen Unterschied.

**Tabelle 19: Mittelwerte (in Klammern Standardabweichungen) der Leistungen in Lesen, Rechtschreibung (Diktat) und Aufsatz in den landesweiten Diagnosearbeiten im Juli 2002 (Ende Klasse 3)**

<b>Gruppierung</b>	<b>Lesen</b>	<b>Rechtschreibung</b>	<b>Aufsatz</b>
A-Modelle	12.3 (3.4)	6.5 (5.7)	8.8 (2.7)
B-Modelle	11.9 (3.2)	6.7 (4.6)	8.8 (3.2)
Vergleichsschulen	12.4 (3.4)	5.8 (4.9)	9.1 (2.8)

Signifikante Mittelwertunterschiede: nein

Vergleicht man die Ergebnisse der landesweiten Diagnosearbeiten mit den in Kapitel II.6 dokumentierten Schulleistungsentwicklungen, so bestätigt sich auch hier der Mathematikvorteil der B-Modelle. Ein Sprachvorteil der A-Modelle kann zwar ebenfalls numerisch bestätigt werden, lässt sich jedoch nicht statistisch absichern.

### Leistungsentwicklung von potentiell leistungsstarken Schülern und von Risikoschülern in Abhängigkeit vom Modell

In den beiden voran stehenden Teilkapiteln wurde die Leistungsentwicklung der in der Hauptstudie (Kohorte 3) untersuchten Schülerinnen und Schüler in Abhängigkeit von der Art und Weise der pädagogischen Gestaltung ihres Erstunterrichts (A-Modelle, B-Modelle, Vergleichsschulen) analysiert. Die dort berichteten Unterschiede erlauben Aussagen zur Frage, ob die Beschulung im Rahmen von A-Modellen oder B-Modellen im Allgemeinen zu Vorteilen oder Nachteilen der schulischen Entwicklung von Grundschulkindern führen. In diesem Abschnitt werden weitergehende Analysen dargestellt mit dem Ziel zu klären, ob Risikokinder von den pädagogischen Innovationen der A-Modelle und B-Modelle unterschiedlich profitieren bzw. ob potentiell leistungsstarke Schülerinnen und Schüler sich modellabhängig unterschiedlich entwickeln.

Um entsprechende statistische Analysen durchführen zu können wurde in einem ersten Schritt festgelegt, was im Rahmen dieses Berichtes unter einem potentiell leistungsstarken Kind bzw. unter einem Risikokind zu verstehen ist. Zum Einschulungszeitpunkt der Kinder der Hauptstudie (Kohorte 3) war der kognitive Entwicklungsstand der Kinder in vier Bereichen erfasst worden, die als zentrale individuelle kognitive Voraussetzungen schulischer Leistungsentwicklung in den Bereichen Lesen, Rechtschreiben und Mathematik gelten. Im Einzelnen handelt es sich dabei um die phonologische Bewusstheit, das phonologische Arbeitsgedächtnis, das räumliche Gedächtnis sowie das mathematische Denken. In Abschnitt II.6 ist dargestellt, wie diese vier Bereiche empirisch erfasst wurden.

In jedem dieser vier Bereiche können nun weit überdurchschnittliche Kompetenzen oder auch Störungen auftreten, die einen potentiellen Vorteil bzw. ein Risiko für die angemessene schulische Entwicklung darstellen. In Anlehnung an internationale Definitionen schulischer Teilleistungsstörungen könnte man z.B. davon ausgehen, dass eine risikoverdächtige Störung vorliegt, wenn die individuelle Leistung eines Kindes zu den 15% schlechtesten Werten der Altersgruppe gehört. Wendet man diese Logik auf die Absicht an, Risikokinder zu identifizieren, dann könnte man alle die Kinder als Risikokinder bezeichnen, die

in einem der vier aufgeführten Funktionsbereiche zu den 15% schlechtesten ihrer Altersgruppe gehören. Die Auswertungen ergaben nun, dass fast ein Drittel der untersuchten Stichprobe in diese Kategorie gehört, d.h. dass etwa 33 Prozent der im Rahmen der Hauptstudie untersuchten Kinder in wenigstens einem der vier Bereiche (phonologische Bewusstheit, phonologisches Arbeitsgedächtnis, räumliches Gedächtnis, mathematisches Denken) zu den 15 Prozent schlechtesten der Altersgruppe gehören. Daher schien eine solche Festlegung von „Risiko“ als zu weich. Macht man das Kriterium etwas härter, in dem man fordert, dass bei Vorliegen von schlechten Leistungen in wenigstens zwei der vier Bereiche von einem ernst zu nehmenden Risiko für die schulische Leistungsentwicklung auszugehen ist, dann verbleiben etwa 10 Prozent der untersuchten Kinder (103 von 1040) als Risikokinder.

Analog zu dieser Festsetzung wurden all die Kinder als potentiell leistungsstark klassifiziert, die in wenigstens zwei der vier genannten Bereichen zu den 15% leistungsstärksten gehörten und in keinem der Bereiche zu den 15% leistungsschwächsten. Auch die Gruppe der so definierten potentiell leistungsstarken Schüler umfasste ca. 10 % der kurz nach der Einschulung untersuchten Kinder (106 von 1040).

Auf der Grundlage dieser operationalen Definitionen von potentiell leistungsstarken Kindern und von Risikokindern wurden zwei Auswertungsstrategien verfolgt, um die Frage zu beantworten, inwiefern die pädagogischen Innovationen der A- und B-Modelle im Rahmen des „Schulanfangs auf neuen Wegen“ für Kinder mit ungünstigen Entwicklungsvoraussetzungen (Risikokinder) oder mit günstigen Entwicklungsbedingungen (potentiell leistungsstarke Kinder) Vorteile bezüglich der schulischen Leistungsentwicklung im Vergleich zur herkömmlichen Art des Schulanfangsunterrichts mit sich bringen. Die erste Strategie bestand darin, für alle vorliegenden Schulleistungen (als abhängige Variablen) zweifaktorielle Varianzanalysen (3 Gruppierungen x 3 Entwicklungsvoraussetzungen) zu berechnen. Die Erwartung dabei war, dass sich regelmäßig Haupteffekte für den Faktor Entwicklungsvoraussetzungen ergeben würden (d.h. generell sollte die Leistung bei Risikokindern schlechter und die der potentiell leistungsstarken Kinder besser ausfallen als bei unauffälligen Kindern). Von besonderem Interesse waren in diesem ersten Auswertungsschritt signifikante statistische Interaktionseffekte zwischen den beiden Faktoren Gruppierung und Entwicklungsvoraussetzung. Solche Interaktionen können nämlich auf systematische Unterschiede der Leistungen der Risikokinder bzw. der potentiell leistungsstarken Schüler in Abhängigkeit vom Modell ihres Schulanfangsunterrichts hinweisen.

In einigen Fällen wurde zusätzlich eine zweite Auswertungsstrategie angewandt. Diese unterschied sich von der beschriebenen zweifaktoriellen Varianzanalyse darin, dass zusätzlich eine frühere Leistung des gleichen Bereichs als Kovariante in die Analyse eingeführt wurde. Auf diese Art und Weise werden mögliche frühe Leistungsunterschiede zwischen den drei verglichenen Gruppen statistisch kontrolliert, so dass geprüft werden kann, ob im Lernzuwachs vom früheren zum späteren Zeitpunkt die Modelle in Abhängigkeit von den Entwicklungsvoraussetzungen der Kinder unterschiedlich Wirkungen haben. In den Tabellen 20 bis 23 sind die Mittelwerte (und Standardabweichungen) der analysierten schulischen Leistungswerte der Schülerinnen und Schüler der Hauptuntersuchung (Kohorte 3), getrennt nach Gruppierung (A-Modell vs. B-Modell vs. Vergleichsschule) und Entwicklungsvoraussetzung (potentiell leistungsstarkes Kind vs. unauffälliges Kind vs. Risikokind) dargestellt.

**Tabelle 20: Mittelwerte (in Klammern Standardabweichungen) der Leistungen im Rechtschreiben (WRT 1+, März 01; WRT 3+, Juni 03), getrennt nach Modell und Entwicklungsvoraussetzungen der Kinder**

<b>Gruppierung</b>	<b>WRT 1+</b>	<b>WRT 3+</b>	<b>Anzahl der Fälle</b>
<b>A-Modelle</b>			
pot. leistungstark	21.3 (2.7)	48.4 (6.6)	44
unauffällig	19.2 (4.2)	44.6 (10.1)	244
Risiko	18.7 (4.7)	41.2 (9.9)	21
<b>B-Modelle</b>			
pot. leistungstark	20.5 (4.1)	47.2 (4.9)	18
unauffällig	19.7 (3.5)	44.0 (7.7)	84
Risiko	17.1 (3.6)	38.7 (9.0)	11
<b>Vergleichsschulen</b>			
pot. leistungstark	20.5 (3.9)	47.5 (5.7)	15
unauffällig	19.8 (3.8)	43.5 (7.9)	170
Risiko	18.0 (4.3)	36.5 (13.5)	8

Signifikante Mittelwertunterschiede: Risiko-Kinder zeigen schlechtere und potentiell leistungsstarke Kinder zeigen bessere Leistungen als die unauffälligen Kinder; und dieser Leistungsunterschied vergrößert sich vom Anfang der 2. bis zum Ende der 4. Klasse; modellspezifische Effekte sind nicht nachweisbar.

Tabelle 20 ist zu entnehmen, dass im Bereich des Rechtschreibens keine modellspezifischen Wirkungen (auch nicht in Abhängigkeit von den individuellen Entwicklungsvoraussetzungen) auszumachen sind. Ebenfalls zu entnehmen ist der Tabelle allerdings auch, dass bis Ende der vierten Klassenstufe (Juni 2003) eine rapide Abnahme der Anzahl der als Risikokinder klassifizierten Schülerinnen und Schüler (von 103 auf 40) stattfindet. Diese Abnahme ist nur teilweise durch Mobilität der Eltern zu erklären. Ein Großteil dieser Kinder ist schlichtweg aus der Untersuchungsstichprobe heraus gefallen, weil sie im Laufe der Grundschulzeit eine Klasse wiederholen mussten. Interessanterweise verringert sich der Anteil der ursprünglich als Risikokinder identifizierten Schülerinnen und Schüler dadurch an der Gesamtstichprobe von ca. 10 % auf ca. 7 %, während der relative Anteil der potentiell Leistungsstarken von 10% auf 12 % ansteigt.

**Tabelle 21: Mittlere Leistungen (in Klammern Standardabweichungen) der Leseleistungen (Lesetempo WLLP, Leseverständnis HAMLET), getrennt nach Modell und Entwicklungsvoraussetzungen der Kinder**

Gruppierung	Lesetempo (Ende Klasse 2)	Lesetempo (Ende Klasse 4)	Leseverständnis (Ende Klasse 4)
A-Modelle			
pot. leistungsstark	88.3 (20.1)	118.9 (18.8)	13.1 (1.6)
unauffällig	81.9 (25.9)	111.2 (21.9)	11.8 (2.4)
Risiko	71.1 (25.4)	101.4 (22.4)	10.8 (2.5)
B-Modelle			
pot. leistungsstark	87.1 (23.5)	121.7 (20.9)	13.3 (2.5)
unauffällig	83.2 (30.6)	110.9 (19.7)	11.9 (2.5)
Risiko	62.6 (42.5)	102.7 (27.4)	9.0 (2.3)
Vergleichsschulen			
pot. leistungsstark	93.2 (23.0)	118.2 (14.5)	12.4 (2.2)
unauffällig	76.4 (24.1)	107.0 (19.1)	12.0 (2.7)
Risiko	63.5 (18.4)	91.7 (20.5)	9.9 (4.1)

Signifikante Mittelwertunterschiede:

Potentiell leistungsstarke Kinder zeigen bessere, Risiko-Kinder zeigen schlechtere Leistung als unauffällige Kinder; beim Leseverständnis zeigt sich ein tendenzieller Interaktionseffekt ( $p < .08$ ) der darauf hindeutet, dass die Risikokinder in den A-Modellen am Ende der 4. Klasse ein besseres Leseverständnis zeigen als die Risikokinder der B-Modelle. Signifikante modellspezifische Effekte sind nicht nachweisbar.

**Tabelle 22: Mittlere T-Werte (in Klammern Standardabweichungen) der Mathematikleistungen in der DEMAT-Reihe, getrennt nach Modell und Entwicklungsvoraussetzungen der Kinder**

Gruppierung	DEMAT 1+ (Anfang Klasse 2)	DEMAT 2+ (Anfang Klasse 3)	DEMAT 4 (Ende Klasse 4)
A-Modelle			
pot. leistungsstark	54.0 (7.0)	52.5 (8.3)	58.5 (9.1)
unauffällig	51.2 (8.0)	50.1 (8.9)	53.8 (8.3)
Risiko	45.8 (6.9)	43.2 (9.0)	49.5 (7.6)
B-Modelle			
pot. leistungsstark	57.9 (7.5)	55.9 (7.0)	64.3 (8.3)
unauffällig	54.3 (7.2)	51.2 (8.4)	56.8 (7.7)
Risiko	44.9 (10.4)	43.7 (11.2)	46.9 (7.3)
Vergleichsschulen			
pot. leistungsstark			



unauffällig Risiko	56.7 (4.8)	58.0 (5.0)	60.7 (6.8)
	51.2 (9.9)	49.6 (9.4)	52.5 (9.5)
	50.5 (8.0)	44.8 (9.6)	44.0 (14.4)

Signifikante Mittelwertunterschiede:

Für den Faktor Entwicklungsvoraussetzungen zeigt sich der erwartete Effekt; ein genereller Modelleffekt ist nicht nachweisbar; allerdings ist der Leistungszuwachs vom Anfang der 2. Klassenstufe bis zum Ende der 4. Klassenstufe für die Risikokinder bei den A-Modellen größer als bei den Vergleichsschulen ( $p < .02$ ) und tendenziell sogar besser als bei den B-Modellen ( $p < .08$ ).

**Tabelle 23: Mittelwerte (in Klammern Standardabweichungen) der in den landesweiten Diagnosearbeiten am Ende der dritten Klassenstufe, getrennt nach Modell und Entwicklungsvoraussetzungen der Kinder**

Gruppierung	Lesen	Diktatfehler	Aufsatz	Mathematik
A-Modelle pot. leistungstark unauffällig Risiko	14.4 (2.2)	4.7 (5.2)	9.6 (2.8)	20.5 (4.4)
	12.2 (3.3)	6.5 (5.6)	8.8 (2.7)	16.1 (6.0)
	10.1 (4.0)	8.9 (6.3)	8.2 (3.0)	12.6 (5.0)
B-Modelle pot. leistungstark unauffällig Risiko	13.1 (3.4)	4.5 (4.2)	9.0 (3.4)	23.2 (5.7)
	12.1 (2.9)	6.9 (4.5)	9.0 (3.1)	18.1 (5.7)
	8.5 (2.8)	9.5 (5.5)	6.2 (3.1)	13.2 (4.7)
Vergleichsschulen pot. leistungstark unauffällig Risiko	14.9 (2.0)	2.6 (2.2)	10.3 (2.1)	20.2 (4.8)
	12.2 (3.4)	5.9 (4.8)	9.1 (2.6)	16.2 (6.1)
	10.7 (4.2)	8.7 (6.2)	7.7 (2.7)	11.8 (7.2)

Signifikante Mittelwertunterschiede:

Bei den drei sprachlichen Leistungen zeigen sich lediglich signifikante Haupteffekte für den Faktor Entwicklungsvoraussetzungen, jedoch keine differentiellen Modelleffekte. Bei den Mathematikleistungen zeigt sich ein tendenzieller Haupteffekt zugunsten von den B-Modellen ( $p < .09$ ).

Insgesamt zeigen sich kaum differentielle Auswirkungen der drei miteinander verglichenen Gruppen. Lediglich in zwei Fällen ergaben die kovarianzanalytischen Auswertungen Leistungsunterschiede zwischen den Gruppen für die Risikokinder. In beiden Fällen zeigten sich Vorteile der A-Modelle. Im Bereich der mit den DEMAT-Verfahren erfassten Mathematikleistungen erwiesen sich die A-Modelle gegenüber den Vergleichsschulen für die Risikokinder als günstiger. Im Bereich des Leseverständnisses am Ende der Grundschulzeit zeigte sich für die Risikokinder ein tendenzieller Vorteil der A-Modelle im Vergleich zu den B-Modellen.

## II.7 Noten in den Fächern Deutsch und Mathematik in den Klassen 2 bis 4 und Bildungsempfehlung Ende Kl. 4

Zeugnisnoten Kl. 2 bis Kl. 4

**Tabelle 24: Zeugnisnoten Klasse 2, Juli 01**

Gruppierung	Fach	Mittelwert	Standard-abweichung	Anzahl der Fälle
A1-Modell	<i>Deutsch</i>	2,33	,895	280
A2-Modell		2,45	,894	92
B-Modell		2,34	,932	144
Vergleichsschulen		2,20	,919	253
A1-Modell	<i>Mathematik</i>	2,16	,860	280
A2-Modell		2,20	,829	92
B-Modell		2,15	,968	144
Vergleichsschulen		2,04	,808	253

Signifikante Mittelwertunterschiede: Note Deutsch: nein; Note Mathematik: nein

**Tabelle 25: Zeugnisnoten Klasse 3, Juni 02**

Gruppierung	Fach	Mittelwert	Standard-abweichung	Anzahl der Fälle
A1-Modell	<i>Deutsch</i>	2,45	,800	210
A2-Modell		2,41	,692	129
B-Modell		2,61	,986	143
Vergleichsschulen		2,57	,905	228
A1-Modell	<i>Mathematik</i>	2,30	,843	210
A2-Modell		2,49	,708	129
B-Modell		2,48	1,018	141
Vergleichsschulen		2,43	,862	214

Signifikante Mittelwertunterschiede: Note Deutsch: nein; Note Mathematik: nein

**Tabelle 26: Zeugnissenoten Klasse 4, Juli 03**

Gruppierung	Fach	Mittelwert	Standard-abweichung	Anzahl der Fälle
A1-Modell	<i>Deutsch</i>	2,48	,825	243
A2-Modell		2,57	,795	131
B-Modell		2,61	,826	132
Vergleichsschulen		2,65	,862	226
A1-Modell	<i>Mathematik</i>	2,48	,794	243
A2-Modell		2,51	,768	131
B-Modell		2,40	,879	133
Vergleichsschulen		2,63	,888	221

*Trotz numerischer Unterschiede der Noten bewegen sich dennoch alle modellspezifischen Notendurchschnitte derselben Klassenstufe im selben Bereich, d.h. es gibt keine signifikanten Unterschiede zwischen den Modellen, was die Durchschnittsnoten in Deutsch und Mathematik am Ende der Kl. 2, Kl.3 und Kl. 4 betrifft.*

Alter und Bildungsempfehlung, aufgebrochen nach Modell (Stand Juli 03)

**Tabelle 27: Bildungsempfehlung und Modell**

Gruppierung	Bildungsempfehlung (in Prozent)			Anzahl
	HS	RS	GY	
A1-Modell	19,1%	25,3%	55,6%	254
A2-Modell	14,5%	29,0%	56,5%	131
B-Modell	16,0%	29,0%	55,0%	130
Vergleichsschulen	27,9%	26,5%	45,6%	208

Signifikante Mittelwertunterschiede: ja, und zwar ( $p=0,3$ )  
V-Schulen > A1-, A2- und B- Modell

*Der signifikante Mittelwertsunterschied zwischen den Modellen ist auf die unterschiedliche Struktur der Bildungsempfehlung der Vergleichsschulen zurückzuführen. Diese sprechen erkennbar mehr Hauptschulempfehlungen und weniger Gymnasialempfehlungen aus als die Schulen der drei Schuleingangsmodelle.*

**Tabelle 28: Alter und Modell**

<b>Gruppierung</b>	<b>Mittelwert</b>	<b>Standard- abweichung</b>	<b>Anzahl der Fälle</b>
A1-Modell	124,4	4,5	254
A2-Modell	124,5	4,7	131
B-Modell	125,7	4,0	130
Vergleichsschulen	126,9	4,2	208

Signifikante Mittelwertunterschiede: ja, und zwar ( $\eta^2=0,036$ )  
V-Schulen > A1-, A2- und B- Modell

**A1-Modelle:**

Jüngstes Kind = 9 Jahre, 3 Monate

Ältestes Kind = 11 Jahre, 9 Monate

**A2-Modelle:**

Jüngstes Kind = 9 Jahre, 4 Monate

Ältestes Kind = 11 Jahre, 11 Monate

**B-Modelle:**

Jüngstes Kind = 9 Jahre, 10 Monate

Ältestes Kind = 12 Jahre, 0 Monate

**Vergleichsschulen:**

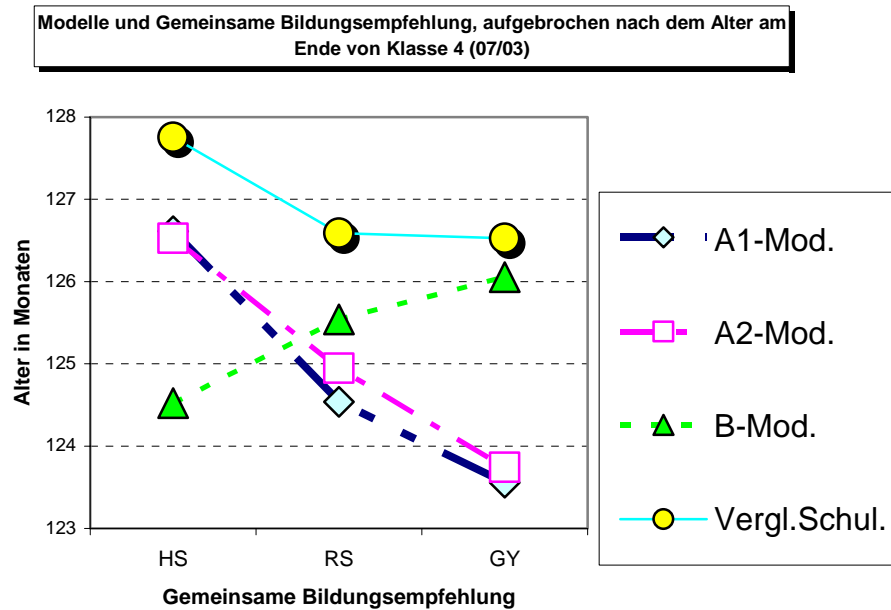
Jüngstes Kind = 9 Jahre, 9 Monate

Ältestes Kind = 12 Jahre, 9 Monate

*Die Schülerinnen und Schüler der Schulen der drei Schuleingangsmodelle sind im Schnitt signifikant jünger als die der Vergleichsschulen.*

*(Es ist allerdings darauf zu achten, dass diese Aussage nur für die Schülerinnen und Schüler gilt, die in unserer Analysestichprobe ohne Umwege die Klassenstufen 1 bis 4 durchlaufen haben).*

**Abbildung 5: Modelle und gemeinsame Bildungsempfehlung, aufgebrochen nach dem Alter am Ende von Klasse 4 (07/03)**



Signifikante Mittelwertunterschiede:

- Modell:

ja, und zwar ( $p=0,000$ )  
V-Schulen > A1-, A2- und B- Modell

- Grundschulempfehlung:

ja, und zwar ( $p=0,011$ )  
V-Schulen > A1-, A2- und B- Modell

- Wechselwirkung Modell und Bildungsempfehlung:

ja, und zwar ( $p=0,022$ )  
B- Modell hat umgekehrte Altersstruktur, bezogen auf die Bildungsempfehlung: HS-empfohlene Schüler jünger, Gy-empfohlene älter als bei A1-, A2-, und VS

*Die signifikante Wechselwirkung ist darauf zurückzuführen, dass die Schülerinnen und Schüler der Schulen des B-Modells eine umgekehrte Altersstruktur, bezogen auf die Bildungsempfehlung aufweisen:*

*HS-empfohlene Schüler sind im Schnitt jünger, Gy-empfohlene im Schnitt älter; bei den A1-, A2-, und VS-Modellen ist diese Relation genau umgekehrt.*

## **II.8 Unterrichtsgestaltung**

### Unterrichtsbezogene Ziele der wissenschaftlichen Begleitung

Zu den Zielen des Modellversuchs „Schulanfang auf neuen Wegen“ gehört es, neben schulorganisatorischen Veränderungen didaktisch-methodische Innovationen vornehmlich im Bereich des Anfangsunterrichts der Grundschule zu initiieren.

Aufgabe der wissenschaftlichen Begleitung ist es daher, bei den am Schulversuch teilnehmenden Klassen die didaktisch-methodische Gestaltung des Anfangsunterrichts zu untersuchen. Dabei soll insbesondere geprüft werden, wie sich die veränderte Schuleingangsstufe auf die konzeptionelle Weiterentwicklung des Unterrichts angesichts der Heterogenität der Schulanfänger in kognitiver, sprachlicher, emotionaler und sozialer Hinsicht auswirkt. Darüber hinaus soll die Einstellung der am Schulversuch teilnehmenden Lehrerinnen und Lehrer zu den Modellversuchen erhoben werden.

Den pädagogischen Orientierungsrahmen für die Anlage und Durchführung der auf die Unterrichtsgestaltung konzentrierten Untersuchung bildet das dem Modellversuch zugrunde liegende Schulverständnis. Dieses berücksichtigt die von der einschlägigen Forschung bestätigten veränderten soziokulturellen Bedingungen des Aufwachsens der heutigen Kindergeneration einschließlich der daraus sich ergebenden pädagogischen und didaktischen Ansprüche gegenüber der Grundschule (vgl. Fölling-Albers, 2003). Dabei geht es, um nur einige Beispiele zu nennen, konzeptionell um die Verfassung der Grundschule als Lern- und Lebensraum, didaktisch-methodisch um individualisierende und differenzierende Lernzugänge und Lernangebote, pädagogisch um die Rhythmisierung von Schul-, Unterrichts- und Lernzeiten und schulorganisatorisch um Jahrgangsmischung.

### Anlage und Durchführung der Hauptuntersuchung

Angesichts der Zielsetzungen des Modellversuchs konzentrierte sich das Forschungsinteresse auf die Frage, welche Veränderungen didaktisch-methodischer Art die neu konzipierte Schuleingangsstufe im Vergleich zur traditionellen Einschulungspraxis nach sich zieht. Im Vergleich zu der derzeit dominierenden effektivitätsorientierten empirischen Lehr- und Unterrichtsforschung (vgl. Weinert, 2001) konzentriert sich die wissenschaftliche Begleitung des Schulversuchs auf die Erhebung der praktizierten Unterrichtsgestaltung der Grundschullehrerinnen selbst.

#### **Erhebungsinstrumentarien**

Die wissenschaftliche Begleitung des Schulversuchs konnte zur Gewinnung empirischer Daten über die praktizierte Unterrichtsgestaltung nicht auf normierte und standardisierte Erhebungsverfahren zurückgreifen. Daher war es notwendig, ein eigenes Erhebungsinstrumentarium zu entwickeln.

Eine in regelmäßigen Abständen durchgeführte standardisierte Beobachtung durch geschultes Personal wäre wegen der zu erwartenden größeren Objektivität der Daten wünschenswert gewesen, konnte jedoch aus finanziellen, personellen und forschungspraktischen Gründen nicht durchgeführt werden.

Die wissenschaftliche Begleitung entschloss sich daher, einen speziell auf die Untersuchungsziele ausgerichteten Lehrerfragebogen zu erstellen, um auf diesem Wege die notwendigen Daten zur Unterrichtsgestaltung zu gewinnen. Demnach beruhen die nachfol-

gend darzustellenden Untersuchungsergebnisse auf Selbstauskünften der am Schulversuch teilnehmenden Lehrerinnen.

Der Fragebogen wurde in mehreren Varianten erstellt. Er besteht aus einem Katalog von Fragen, die für alle an der Untersuchung teilnehmenden Schulklassen einschließlich der Vergleichsklassen gelten, und einem variierenden modellspezifischen Frageteil. Die gestellten Fragen konnten zum geringsten Anteil offen beantwortet werden. Ihre überwiegende Anzahl musste mit Auswahlantworten quantifizierender und qualifizierender Art bearbeitet werden, wobei neben Einfachnennungen teilweise auch Mehrfachnennungen möglich waren.

Inhaltlich erstrecken sich die gestellten Fragen mit unterschiedlicher Gewichtung auf drei Bereiche. Der erste kleinere Teil bezieht sich auf berufsbiographische Angaben zu den Lehrerinnen sowie auf schulbiographische Angaben der am Modellversuch teilnehmenden Kinder. Der zweite umfangreichste Teil umfasst speziell unterrichtsbezogene Fragen, wie u. a. die Gewichtung von fachlichem und nichtfachlichem Unterricht, Umfang und Art der praktizierten Unterrichtsformen (z. B. Gruppenarbeit, Lernzirkel) jeweils differenziert nach Einzelfächern, weiterhin den unterrichtlichen Umgang mit leistungsschwachen und leistungsstarken Grundschulern. Mit dem dritten Fragebogenteil soll die Akzeptanz des Schulversuchs bei den teilnehmenden Lehrerinnen erfasst werden.

### **Erhebungszeitpunkte und Stichprobe**

Der neu entwickelte Lehrerfragebogen wurde zu insgesamt vier Erhebungszeitpunkten eingesetzt und zwar in folgenden Einschulungsjahrgängen:

1997/98  
1998/99  
1999/00  
2000/01.

Die Lehrerfragebogen wurden jeweils gegen Ende des Schuljahres im Frühsommer versandt. Da beim Ersteinsatz des Fragebogens beim Einschulungsjahrgang 1997/98 noch keine Jahrgangsmischung vorlag, bleiben die dazu erhobenen Daten bei den nachfolgend darzustellenden Untersuchungsergebnissen unberücksichtigt. Auswertungsbefunde zur Ersterhebung finden sich im Zwischenbericht (vgl. Arbeitskreis, 2000).

Für das Schuljahr 1998/99 lagen 121 ausgefüllte Fragebogen vor, die sich wie folgt auf die einzelnen Modelle verteilen:

33 A1-Erprobungsklassen, 19 A2 – Erprobungsklassen, 48 B-Erprobungsklassen, 21 Vergleichsklassen.

Für das Schuljahr 1999/00 wurden 70 Fragebogen ausgefüllt.

Davon entfielen 25 auf die A1-Erprobungsklassen, 18 auf die A2- Erprobungsklassen, 11 auf B- Erprobungsklassen und 16 auf die Vergleichsschulen.

Für das Schuljahr 2000/01 standen insgesamt 68 ausgefüllte Fragebogen zur Verfügung.

Davon stammten 24 von A1-Erprobungsklassen, 19 von A2- Erprobungsklassen, 10 von B-Erprobungsklassen und 15 von Vergleichsklassen.

Die befragten Lehrkräfte waren in den Modell- und Vergleichsklassen sowie zu allen drei Erhebungszeitpunkten annähernd zu 100 Prozent weiblichen Geschlechts. Das Lebensalter bewegte sich zwischen 26 und 62 Jahren. Die Dauer der beruflichen Tätigkeit in der Grundschule erstreckte sich von 1 bis 40 Jahren.

Bei den nachfolgend darzustellenden Auswertungsbefunden ist zu beachten, dass zwischen der Anzahl der jeweils beantworteten Fragen und der Zahl der zurückgegebenen Fragebogen leichte Abweichungen auftreten können, da nicht jeder Fragebogen vollständig ausgefüllt wurde und darüber hinaus bei Einzelfragen Mehrfachantworten möglich waren.

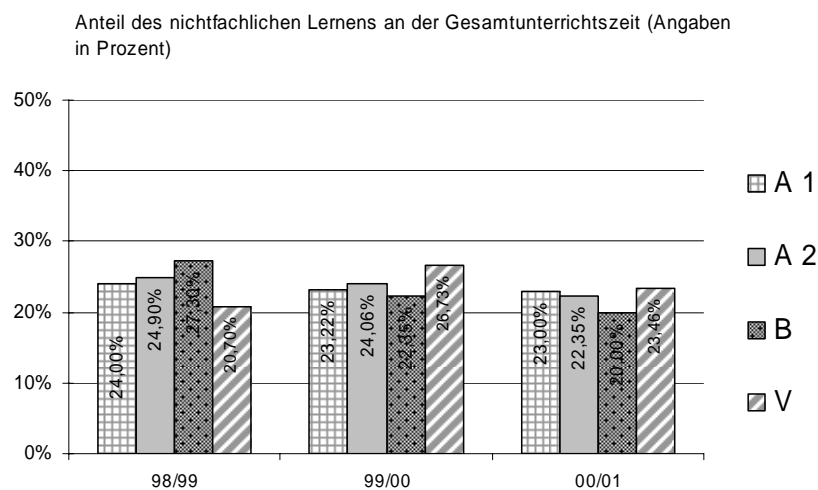
## Ergebnisse der Hauptuntersuchung

Die nachfolgend dargestellten Untersuchungsergebnisse geben Auskunft über modellspezifische unterrichtsbezogene Veränderungen im Vergleich zur regulären Einschulungspraxis. Dabei erfolgt die Auswertung für die vorstehend genannten Stichproben als Systemvergleich d.h. als Vergleich der Erprobungsklassen mit den Vergleichsklassen. Für den Zeitraum der Untersuchungsdauer soll darüber hinaus geprüft werden, ob sich auf der Basis der stichprobenspezifischen Einzelbefunde Stabilitäten in den Ergebnissen abzeichnen.

### Fachliches und nichtfachliches Lernen (freies Spiel, Stilleübungen etc.)

Da bedingt durch die Übergangssituation der Anfangsunterricht kindergartenähnliche Spiel- und Betätigungsformen berücksichtigt, erstreckte sich die erste Frage auf das Verhältnis von fachlichem und nichtfachlichem Lernen. Welchen prozentualen Anteil das nichtfachliche Lernen an der gesamten Unterrichtszeit einnimmt, stellt in modellspezifischer Differenzierung für alle drei Erhebungszeitpunkte die nachfolgende Graphik dar.

**Abbildung 6: Anteil des nichtfachlichen Lernens an der Gesamtunterrichtszeit (in %)**





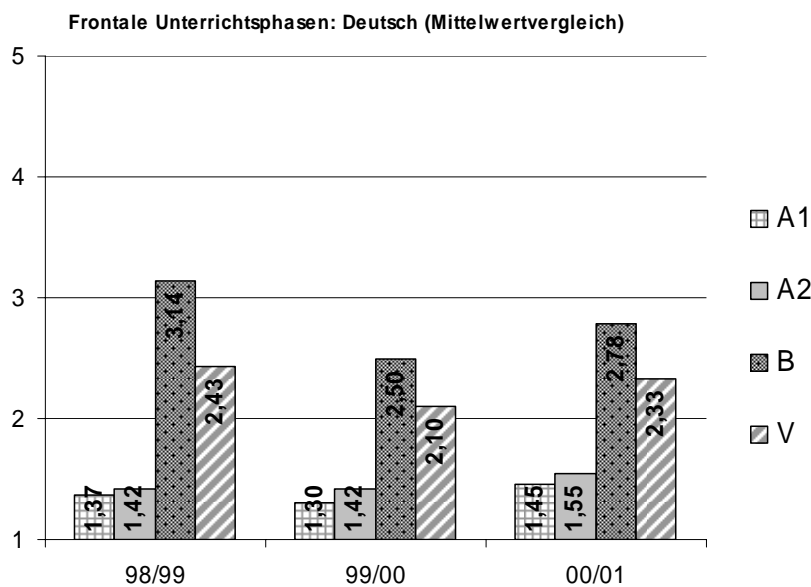
Über die drei Erhebungszeitpunkte hinweg lassen sich, was den prozentualen Anteil des nichtfachlichen Lernens an der Gesamtunterrichtszeit betrifft, keine signifikanten Unterschiede zwischen den Erprobungs- und Vergleichsklassen feststellen, allerdings mit einer Ausnahme. Zum Erhebungszeitpunkt 1998/99 wird in den B-Modellklassen signifikant häufiger nichtfachliches Lernen praktiziert als in den Vergleichsklassen.

Bei den A1-Modellen bleibt der Anteil nichtfachlichen Lernens an der Gesamtunterrichtszeit über die dreijährige Untersuchungsdauer hinweg relativ stabil, während er bei den B-Modellklassen tendenziell von 27,3 auf 20 Prozent sinkt und bei den Vergleichsklassen tendenziell zunimmt.

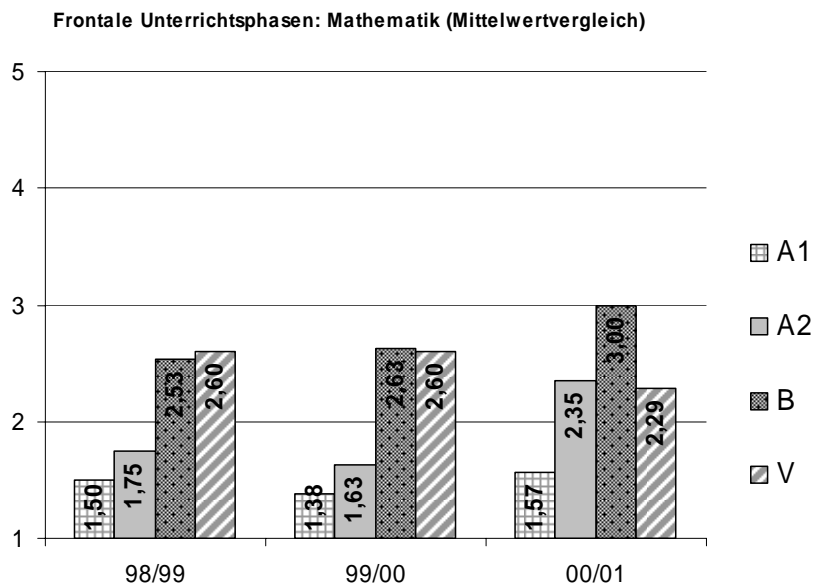
### Frontale Unterrichtsphasen

Die Frage nach dem Anteil frontaler Unterrichtsphasen an der wöchentlichen Unterrichtszeit in den Kernfächern Deutsch, Mathematik, sowie Heimat- und Sachunterricht wurde von den befragten Lehrkräften mittels einer Antwortskala mit gestuften prozentualen Häufigkeiten erhoben. Daraus wurde ein Mittelwert errechnet (1 = sehr selten; 2 = selten; 3 = oft; 4 = sehr oft). Die Auswertung der Antworten über Mittelwertvergleiche veranschaulichen folgende Grafiken.

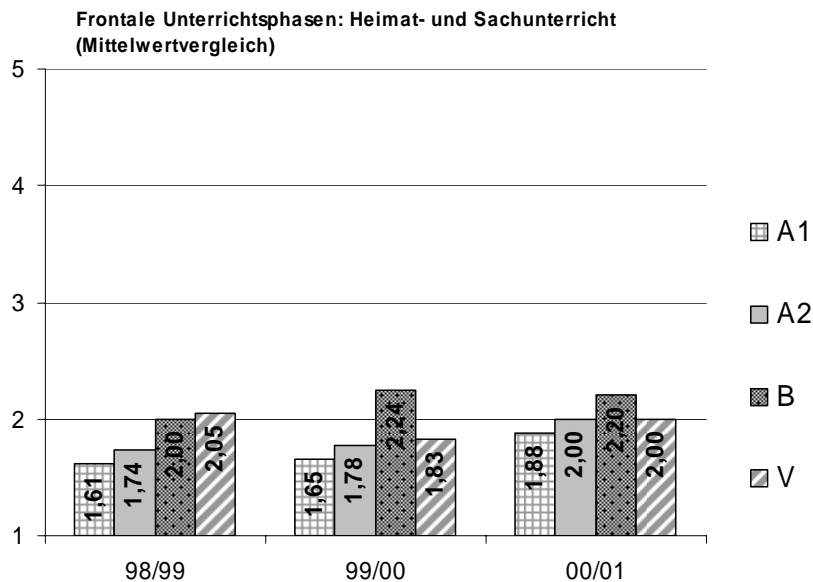
**Abbildung 7: Frontale Unterrichtsphasen: Deutsch (Mittelwertvergleich)**



**Abbildung 8: Frontale Unterrichtsphasen: Mathematik (Mittelwertvergleich)**



**Abbildung 9: Frontale Unterrichtsphasen: Heimat- und Sachunterricht (Mittelwertvergleich)**



Die Ergebnisse liefern für die Fächer Deutsch und Mathematik einen übereinstimmenden und für alle drei Erhebungszeitpunkte stabilen Befund, was die modellspezifischen Unterschiede angeht. Danach wird bei den A-Modellklassen signifikant weniger häufig Frontalunterricht erteilt als bei den Vergleichsklassen.<sup>53</sup> Dieses Resultat tritt beim Heimat- und Sachunterricht nur zum Erhebungszeit 1998/99 auf.

<sup>53</sup> Diese, wie auch alle weiteren im Text genannten Signifikanzen, sind berechnet für  $\alpha < 0,05$ .

## Praktizierte Unterrichtsformen in ausgewählten Grundschulfächern

Um die Auftretenshäufigkeit der praktizierten Unterrichtsformen (z.B. Kreisgespräch, Gruppenarbeit, Partnerarbeit, Lernzirkel) in den Erprobungs- und Vergleichsklassen festzustellen, wurde den Lehrerinnen eine vierstufige Antwortskala vorgegeben:

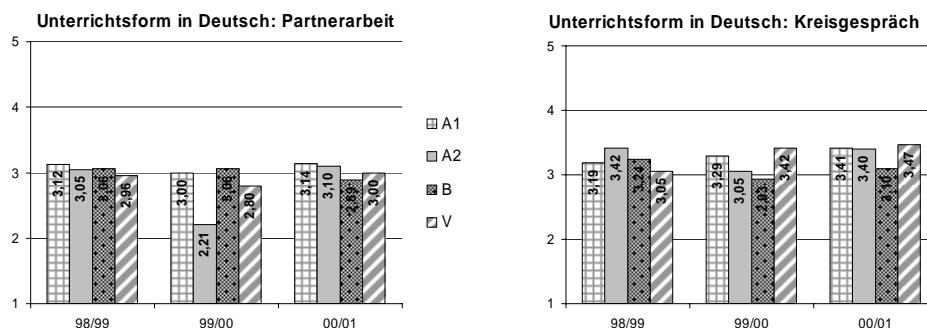
- 1= nie
- 2= selten
- 3 = häufig
- 4= sehr häufig

In den nachfolgenden Abbildungen wird für die einzelnen Unterrichtsformen über einen Mittelwertvergleich ausgewiesen, wie häufig diese jeweils in den Erprobungs- und Vergleichsklassen in den Fächern Deutsch, Mathematik und Heimat- und Sachunterricht praktiziert wurden. Je höher in den Säulendiagrammen der angegebene Mittelwert ist, umso häufiger wird die entsprechende Unterrichtsform angewendet.

### Deutsch

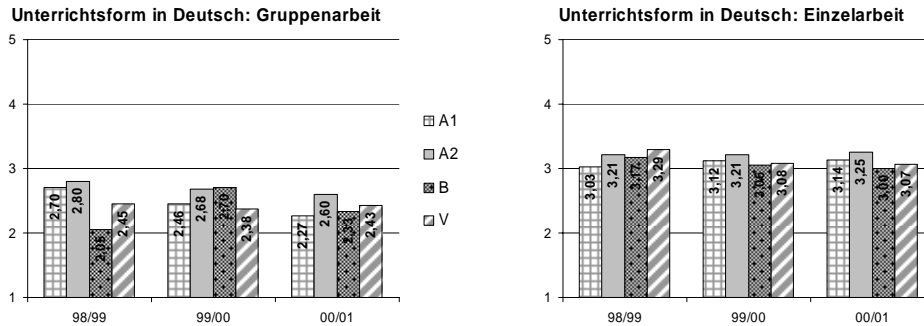
Für die im Grundschulfach Deutsch praktizierten Unterrichtsformen verdeutlichen die nachfolgenden Abbildungen die Auswertungsergebnisse (Mittelwertvergleich).

**Abbildung 10: Deutsch: Partnerarbeit, Kreisgespräch**



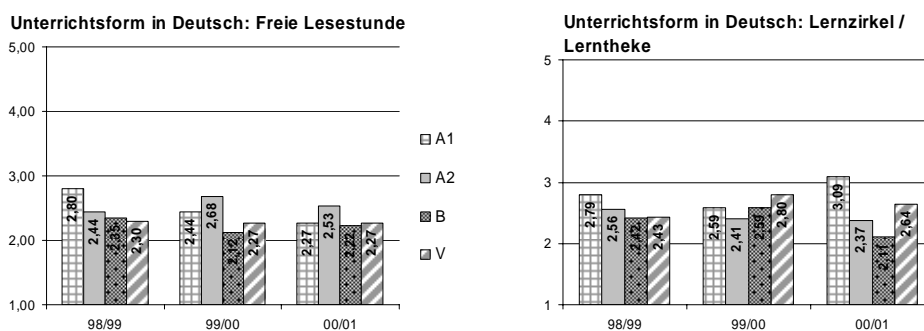
Wie aus den Darstellungen ersichtlich ist, gibt es im Fach Deutsch in der Auftretenshäufigkeit von Kreisgespräch und Partnerarbeit keine nennenswerten und über alle drei Erhebungszeitpunkte hinweg dauerhaften Unterschiede innerhalb der Modellvarianten, auch nicht zwischen diesen und den Vergleichsklassen. Lediglich für den Erhebungszeitraum 1999/00 zeigt sich, dass das Kreisgespräch in den Vergleichsklassen signifikant häufiger praktiziert wird. Für denselben Erhebungszeitraum ergibt die Auswertung der Daten, dass die Partnerarbeit in den A 2-Erprobungsklassen signifikant weniger häufig auftritt als in den Vergleichsklassen.

Abbildung 11: Deutsch: Gruppen- und Einzelarbeit



Während in der Auftretenshäufigkeit der Einzelarbeit weder innerhalb der Modellvarianten noch zwischen diesen und den Vergleichsklassen signifikante Unterschiede festzustellen sind, lassen sich diese bei der Gruppenarbeit nur für einen einzigen Erhebungszeitraum nachweisen. Diese Unterrichtsform wird im Untersuchungsjahrgang 1998/99 in den A 2-Modellen signifikant häufiger angewendet als in den B-Modellen und den Vergleichsklassen. Bei den Vergleichsklassen bleibt die Auftretenshäufigkeit über den Untersuchungszeitraum hinweg relativ unverändert, während sie bei den A-Modellen tendenziell abnimmt. Dagegen zeigt sich bei der Einzelarbeit über den Erhebungszeitraum hinweg bei den A-Modellen in der Auftretenshäufigkeit eine aufsteigende Tendenz, bei den B-Modellklassen und den Vergleichsklassen eine absteigende.

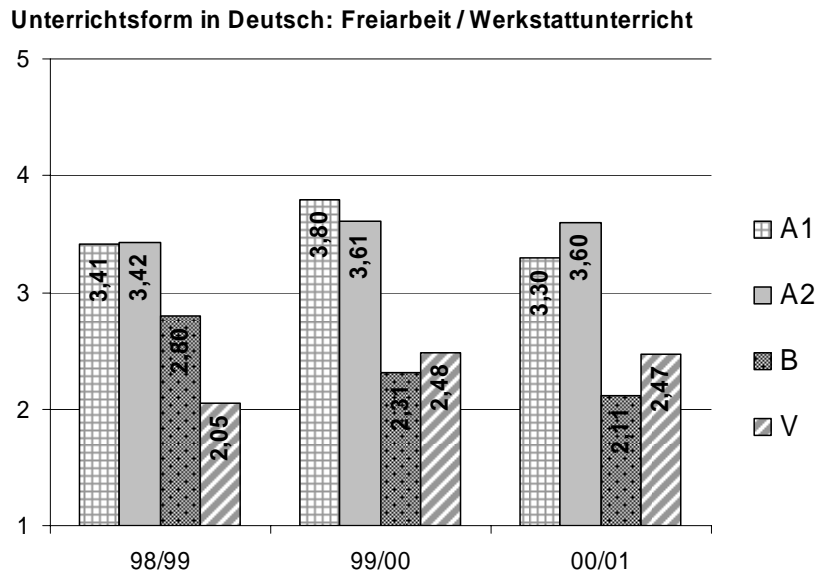
Abbildung 12: Deutsch: Freie Lesestunde und Lernzirkel/Lerntheke



Die freie Lesestunde wird nach den für 1998/99 und 1999/00 erhobenen Daten in den jahrgangsgemischten A-Modellen signifikant häufiger durchgeführt als in den jahrgangsgetreuten B-Modellklassen und Vergleichsklassen. Dabei zeigt sich über die drei Erhebungszeitpunkte hinweg unabhängig von der Modellzugehörigkeit eine abnehmende Tendenz in der Auftretenshäufigkeit dieser Unterrichtsform.

Bei Lernzirkel und Lerntheke ergeben sich nur modellintern signifikante Unterschiede zugunsten des A1-Modells gegenüber dem A2- und B-Modell. Ein tendenzieller Entwicklungstrend während der Untersuchungsdauer ist in den Daten nicht erkennbar.

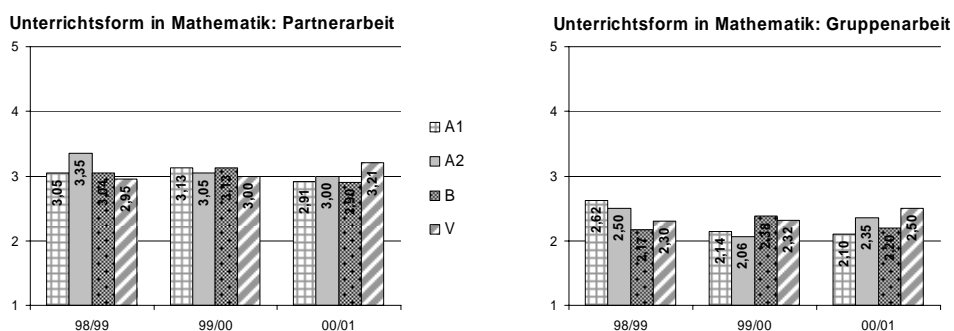
**Abbildung 13: Deutsch: Freiarbeit/Werkstattunterricht**



Auffallend im Vergleich zu den bisher berichteten Ergebnissen sind die Befunde zu Freiarbeit/Werkstattunterricht. Diese Unterrichtsform wird in den jahrgangsgemischten A-Modellklassen signifikant häufiger praktiziert als in den B-Modellklassen und den Vergleichsklassen, ein Ergebnis, das zu allen drei Erhebungszeitpunkten zu verzeichnen ist.

Mathematik

**Abbildung 14: Mathematik: Partnerarbeit/Gruppenarbeit**

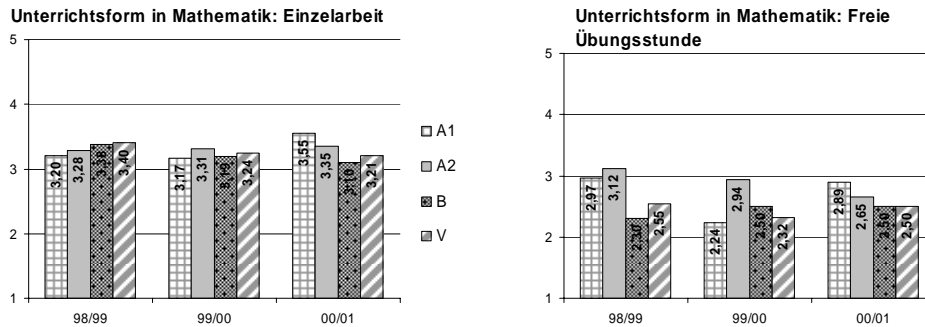


Für die Partnerarbeit wie für die Gruppenarbeit in Mathematik liefern, vergleichbar dem Deutschunterricht, die ausgewerteten Daten punktuelle Ergebnisse. Sie bescheinigen für einen einzigen Erhebungszeitraum(1998/99) dem A2 - Modell eine signifikant häufigere Anwendung der Partnerarbeit als den Vergleichsklassen. Letztere praktizieren signifikant häufiger Gruppenarbeit als die jahrgangsgemischten A1-Modellklassen, ein Ergebnis, das eher erwartungswidrig ausfällt.

Über die Untersuchungsdauer hinweg betrachtet zeigt sich bei den A2-Modellklassen eine tendenzielle Abnahme in der Auftretenshäufigkeit der Partnerarbeit, während diese bei den

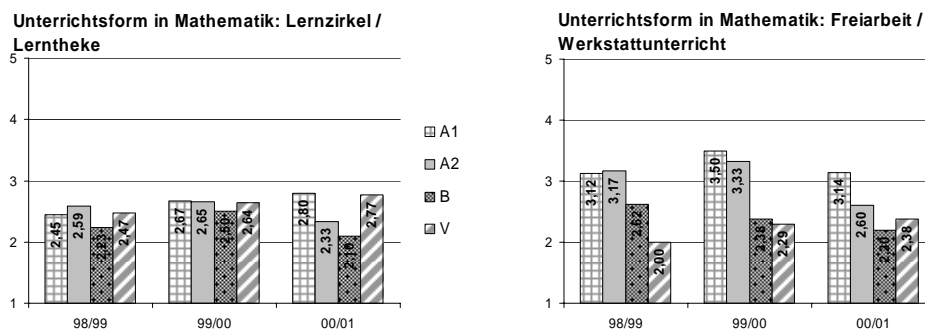
Vergleichsklassen tendenziell ansteigt. Vergleichbare Trends zeichnen sich im Falle der Gruppenarbeit nicht ab.

**Abbildung 15: Mathematik: Einzelarbeit/ Freie Übungsstunde**



Das vom Fach Deutsch her bekannte Ergebnis bezüglich der Einzelarbeit wiederholt sich in der Mathematik. Auch hier sind zu keinem Erhebungszeitpunkt signifikante modellspezifische Unterschiede nachweisbar. Demgegenüber wird die Freie Übungsstunde zu zwei Erhebungszeitpunkten in den jahrgangsgemischten A- Modellklassen signifikant häufiger durchgeführt als in den B- Modellen und Vergleichsklassen, ein Befund, der sich weitgehend mit den zur freien Lesestunde ermittelten Ergebnissen deckt (vgl. Abb.12).

**Abbildung 16: Mathematik: Lernzirkel/Lerntheke und Freiarbeit/Werkstattunterricht**

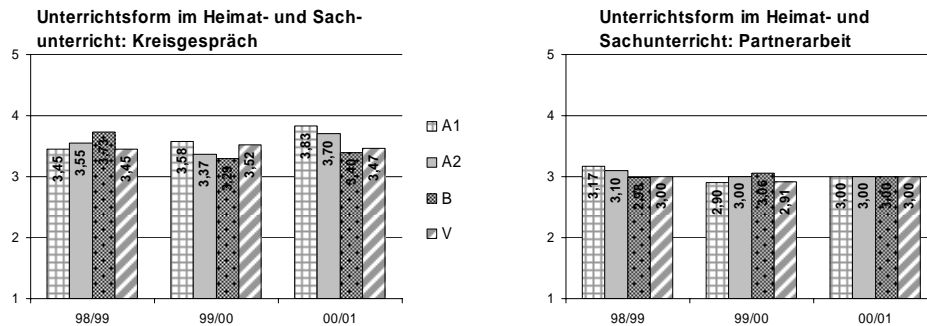


In Mathematik werden Lernzirkel/Lerntheke in den jahrgangsgetrenten B-Modellklassen am wenigsten häufig praktiziert, ein Ergebnis, das jedoch nur für das Schuljahr 2000/01 signifikant ist. Sowohl bei den jahrgangsgemischten A1 -Modellklassen wie bei den jahrgangsgetrenten Vergleichsklassen nimmt während der Untersuchungsdauer die Auftrenshäufigkeit von Lernzirkel/Lerntheke tendenziell zu.

Das im Deutschunterricht ermittelte Ergebnismuster bezüglich Freiarbeit/ Werkstattunterricht wird auch durch die für Mathematik erhobenen Daten bestätigt. Danach wird die genannte Unterrichtsform in den jahrgangsgemischten A- Modellklassen zu allen drei Erhebungspunkten signifikant häufiger angewendet als in den B- Modellen und den Vergleichsklassen.

## Heimat- und Sachunterricht

**Abbildung 17: Heimat- und Sachunterricht: Kreisgespräch/Partnerarbeit**

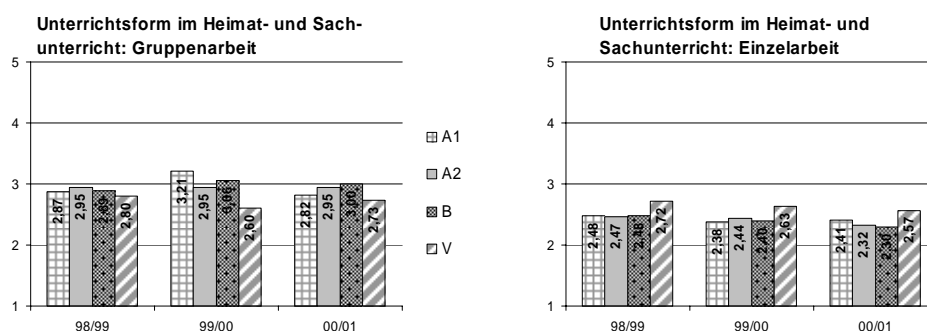


Im Heimat- und Sachunterricht wird das Kreisgespräch für den Untersuchungsjahrgang 2000/01 in den A1- Klassen signifikant häufiger praktiziert als in B-Modellklassen. Für den Untersuchungsjahrgang 1998/99 verhält es sich umgekehrt, da das Kreisgespräch nunmehr signifikant häufiger in den B-Modellklassen als in den A1-Modellklassen zur Anwendung kommt.

Über die dreijährige Untersuchungsdauer hinweg nimmt bei den A-Modellklassen die Auftretenshäufigkeit dieser Unterrichtsform tendenziell stetig zu, während sie bei den Vergleichsklassen in etwa konstant bleibt.

Bei der Partnerarbeit treten überhaupt keine modellspezifischen Signifikanzen auf und über den Untersuchungszeitraum hinweg allenfalls geringfügige tendenzielle Veränderungen.

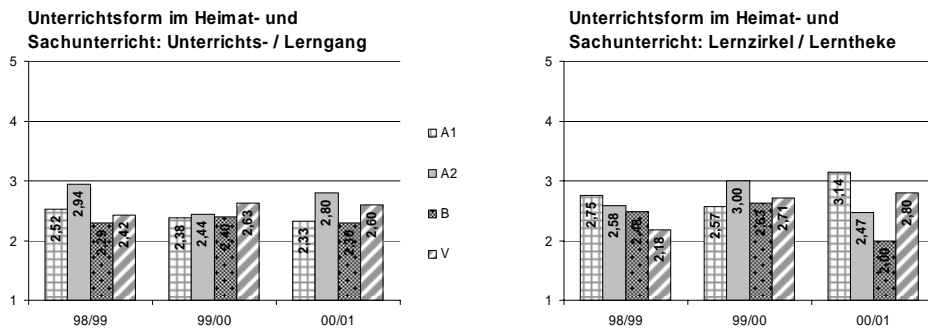
**Abbildung 18: Heimat- und Sachunterricht: Gruppenarbeit/Einzelarbeit**



Die Gruppenarbeit wird tendenziell in den Erprobungsklassen aller Modellvarianten über alle drei Erhebungszeiträume hinweg häufiger praktiziert als in den Vergleichsklassen.

Die erhobenen Daten zur Einzelarbeit bestätigen das Ergebnis, das für diese Unterrichtsform im Deutsch- und Mathematikunterricht schon nachweisbar war. Es sind keine signifikanten Unterschiede feststellbar. Sowohl bei den A2 - Modellen wie bei den Vergleichsklassen zeichnet sich im Laufe des Untersuchungszeitraumes eine abnehmende Tendenz in der Auftretenshäufigkeit der Einzelarbeit ab.

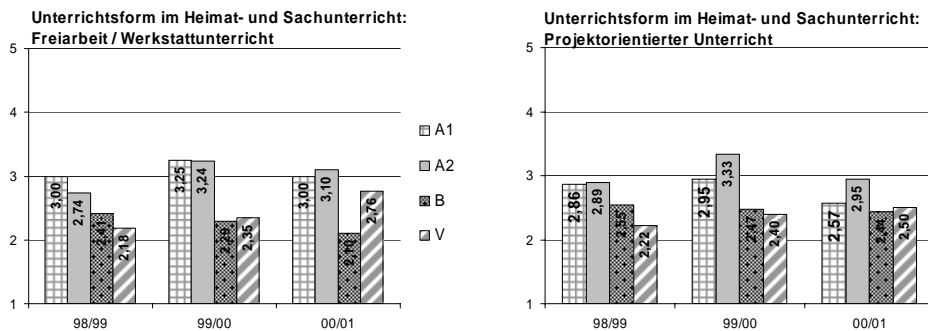
**Abbildung 19: Heimat- und Sachunterricht: Unterrichts-/Lerngang und Lerntheke**



Für den Unterrichtsgang ist die Auftretenshäufigkeit bei allen drei Erhebungszeitpunkten in den A2 -Modellklassen signifikant höher als in den A1- Modellklassen und zeitlich bedingt auch in den B- Modellklassen. Für das Schuljahr 99/00 ergibt sich als Detailergebnis, dass in jahrgangstrennten Vergleichsklassen signifikant häufiger der Lerngang durchgeführt wird als in jahrgangsgemischten A1- Modellen.

Was Lernzirkel und Lerntheke angeht, so zeigen die Daten mit Ausnahme des zweiten Erhebungszeitpunktes, dass diese Unterrichtsform in den A1-Modellen signifikant häufiger praktiziert wird als in den A2- Modell-, B- Modell- und Vergleichsklassen.

**Abbildung 20: Heimat- und Sachunterricht: Freiarbeit/ Werkstattunterricht und projektorientierter Unterricht**



Bei Freiarbeit/Werkstattunterricht ergibt sich als ein eindeutiges Ergebnis, dass diese Unterrichtsform in den jahrgangsgemischten A-Modellklassen über die gesamte Untersuchungsdauer hinweg signifikant häufiger durchgeführt wird als in den jahrgangstrennten B-Modell- und Vergleichsklassen, ein Befund, der auf der Linie der für den Deutsch- und Mathematikunterricht ermittelten Untersuchungsergebnisse liegt (vgl. Abb. 13 und 16).

Analog zu Freiarbeit/Werkstattunterricht fällt die Auftretenshäufigkeit von Projektunterricht in den A- Modellklassen signifikant höher aus als in den B-Modell- und Vergleichsklassen, allerdings begrenzt auf zwei Erhebungszeitpunkte.



## **Veränderungen der Unterrichtsgestaltung**

### Zeitorganisatorische Veränderungen

Im Gegensatz zu den bisher berichteten Ergebnissen beziehen sich die nachfolgenden Ausführungen ausschließlich auf die A- und B- Modellklassen, also nicht auf die Vergleichsklassen.

Die erhobenen Antworten zur Frage, ob durch die Teilnahme am Modellversuch die Unterrichtsgestaltung in der Zeitorganisation verändert wurde, liefern über die gesamte Untersuchungsdauer statistisch signifikante Ergebnisse. Danach wird von den Lehrerinnen und Lehrern der A- Modellklassen diese Frage weitaus häufiger bejaht als bei den Lehrerinnen der B-Modellklassen.

Die konkreten Veränderungen betreffen zum einen den Tagesablauf, zum anderen den Ablauf der Schulwoche. Als häufigste Veränderungen im Tagesablauf werden von den Lehrerinnen der jahrgangsgemischten A-Modellklassen zu allen drei Erhebungszeitpunkten die „Lockerung des 45-Minuten-Taktes“ und der „Wechsel zwischen lehrergesteuertem Unterricht und Freiarbeit“ genannt. Weniger Nennungen entfallen in der Regel auf die „offene Schlussphase“ und die „flexible tägliche Pause“.

An der Spitze der Nennungen zu den zeitorganisatorischen Veränderungen im Wochenablauf steht in allen drei Untersuchungsjahrgängen die „Arbeit mit dem Wochenplan“ und die „Auflockerung des Wochenrhythmus durch musisch bestimmte Phasen“. Vereinzelt wird auch die Durchführung von Wochenanfangs- und -abschlussritualen als zeitorganisatorische Neuerung genannt.

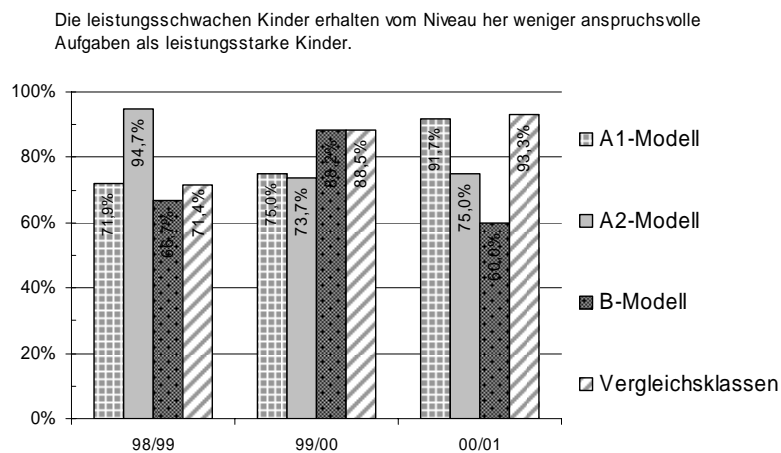
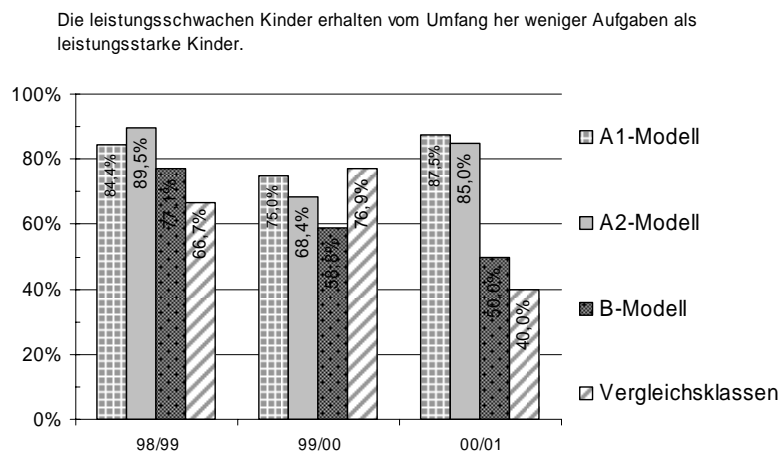
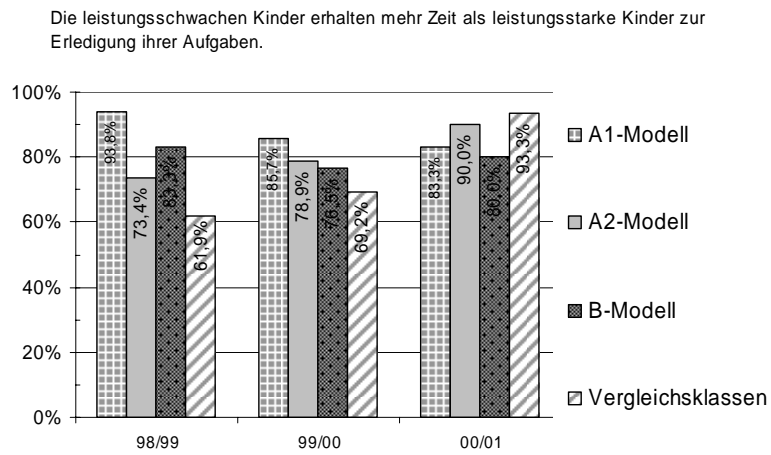
### Unterrichtsmethodische Veränderungen

Zu allen drei Erhebungszeitpunkten geben die Lehrerinnen der A-Modelle häufiger als die der B-Modelle an, dass sie durch die Teilnahme am Modellversuch ihren Unterricht in methodischer Hinsicht verändert haben, ein Ergebnis, das statistisch signifikant ausfällt. Auf die „Intensivierung individualisierender/differenzierender Unterrichtsmaßnahmen“ entfallen mit einer Spitze bei den A-Modellen für die gesamte Untersuchungsdauer die häufigsten Nennungen, gefolgt von „Werkstattunterricht und Freiarbeit“.

## **Unterrichtlicher Umgang mit leistungsstarken und leistungsschwachen Schülerinnen und Schülern.**

Die Befragung zur Unterrichtsgestaltung schloss neben den bisher berichteten Ergebnissen auch eine Erhebung zum Umgang der Grundschullehrer und -lehrerinnen mit den leistungsstarken und -schwachen Schülern und Schülerinnen ein. Die ermittelten Ergebnisse für die drei Erhebungszeitpunkte sind in den nachfolgenden Grafiken dargestellt.

Abbildung 21 a. – c.: Unterrichtlicher Umgang mit leistungsschwachen Schülern

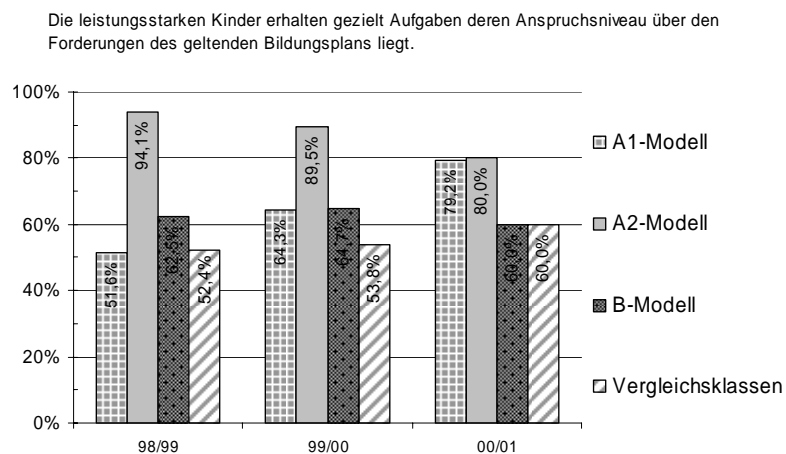
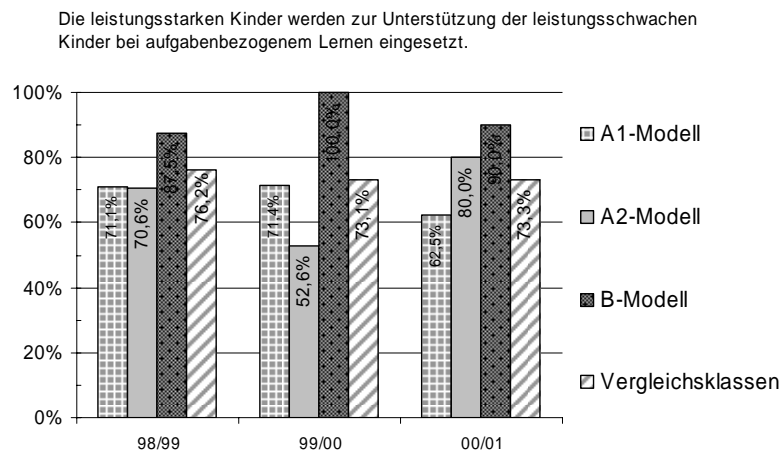


Im Untersuchungsjahr 1998/99 erhalten nach den prozentualen Anteilen in den A-Modellklassen die Leistungsschwachen sowohl mehr Zeit zur Aufgabenerledigung als auch in Umfang und Anspruchsniveau reduzierte Aufgaben als in den Vergleichsklassen.

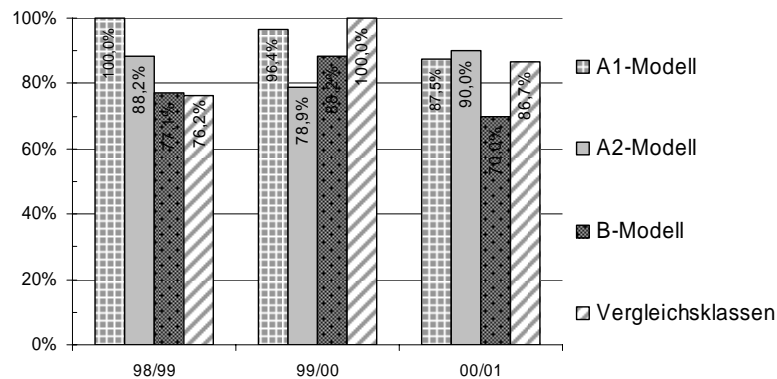
Dabei wird innerhalb der Modellvarianten speziell in den Erprobungsklassen des A1- und des B- Modells bei der unterrichtlichen Förderung Leistungsschwacher der Lernzeitverlängerung der Vorzug gegeben, während eine Senkung des Anforderungsniveaus beim unterrichtlichen Umgang mit dieser Gruppe in den A2- Modellklassen dominiert. Dieser Befund findet in den Ergebnissen der nachfolgenden Erhebungszeitpunkte keine eindeutige Bestätigung. Es lassen sich beim Umgang mit leistungsschwachen Schülern keine dauerhaften modellspezifischen Präferenzen feststellen.

Die ermittelten Ergebnisse zum Umgang mit Leistungsstarken werden für alle drei Erhebungszeitpunkte in Abbildung 22a. - c. dargestellt.

**Abbildung 22 a. – c.: Unterrichtlicher Umgang mit leistungsstarken Schülern**



Die leistungsstarken Kinder können sich entsprechend ihrer Interessen mit selbst gewählten Aufgaben beschäftigen.



Beim unterrichtlichen Umgang mit Leistungsstarken zeigt sich für alle drei Erhebungszeitpunkte, dass speziell bei den A- 2 Modellklassen die gezielte, über dem curricularen Anforderungsniveau liegende Aufgabenstellung die prozentual höchsten Werte erhält. Auf das Helferprinzip entfallen in den B-Modell-Klassen die prozentual höchsten Anteile ebenfalls über den gesamten Untersuchungszeitraum hinweg. Wie an voran stehenden Abbildungen erkennbar, variieren die übrigen erfragten Ergebnisse je nach Erhebungszeitpunkt und Modellzugehörigkeit.

### **Vor- und Nachteile des Modellversuchs aus der Sicht der teilnehmenden Lehrerinnen**

Die Lehrerinnen wurden nach den Vor- und Nachteilen des Modellversuchs befragt und zwar jeweils bezogen auf die Kinder, den Unterricht und die Lehrerin.

#### Vorteile und Nachteile für die Kinder

Die Lehrerinnen der A- Modelle geben in den drei Untersuchungsjahrgängen als Vorteile für die Kinder folgende, in der Häufigkeit ihrer Nennung aufgeführte positive Effekte an:

- (1) „Die Sozialbeziehungen zwischen den Kindern werden positiv beeinflusst“.
- (2) „Bei den Kindern wird selbstständiges und selbstverantwortliches Lernen gefördert“.
- (3) „Die Kinder können nach ihrem eigenen Entwicklungsstand und Lerntempo voranschreiten“.

Die wenigsten Nennungen entfallen durchgängig bei allen Erhebungen auf die Antwortvorgabe „Bei den Kinder wird die Kreativität gefördert“.

Als Nachteile für die Kinder wird von den Lehrerinnen der A-Modelle zu jedem Erhebungszeitpunkt am häufigsten das Ausweichen der Kinder vor Leistungsanforderungen genannt, am zweithäufigsten die Vergrößerung der Leistungsunterscheide zwischen den Kindern.

#### Vorteile und Nachteile für den Unterricht

Als Vorteile für den Unterricht geben die Lehrerinnen der A-Modelle durchgängig für alle Untersuchungszeiträume am häufigsten an, dass durch den Modellversuch die Handlungsdominanz der Lehrerin im Unterricht zugunsten der Eigenaktivität der Kinder abgeschwächt wird. Als der am zweithäufigsten genannte Vorteil wird im Falle der A-Modelle

die Intensivierung von Differenzierungs- und Individualisierungsmaßnahmen von den Lehrerinnen gewertet.

Die am meisten genannten Nachteile für den Unterricht sind aus der Sicht der Lehrerinnen der jahrgangsgemischten A-Modellklassen die Erschwerung von Unterrichtsphasen mit der Gesamtklasse, sowie die größere Störanfälligkeit des Unterrichtsverlaufes, mit Einschränkungen auch ein verlangsamtes Erreichen der Anforderungen des gültigen Bildungsplanes.

#### Vorteile und Nachteile für die Lehrerin

In den A-Modellklassen wird als der am häufigsten genannte Vorteil für die Lehrerin bei zwei Erhebungszeitpunkten die durch den Modellversuch initiierte Weiterentwicklung eigener didaktisch-methodischer Kompetenzen angesehen, bei einem Erhebungszeitpunkt die Schärfung der eigenen Wahrnehmung für die Lernstärken und Lernschwächen der Schüler. Keinen Vorteil sehen die befragten Lehrerinnen in der Intensivierung der Elternarbeit, denn auf diese Antwortvorgabe entfallen zu allen drei Erhebungszeitpunkten die wenigsten Nennungen.

Von den Lehrerinnen der A-Modellklassen wird die Zunahme des Arbeits- und Zeitaufwandes für die Vorbereitung des Unterrichts in eindeutiger Klarheit über den gesamten Erhebungszeitraum hinweg am häufigsten als Nachteil des Modellversuchs wahrgenommen.

#### Auswirkungen der Jahrgangsmischung auf das Sozialverhalten und die Lernentwicklung der Kinder

Zu allen drei Erhebungszeitpunkten wurden die Lehrerinnen zu den von ihnen wahrgenommenen Auswirkungen der in den A-Modellklassen vorhandenen Jahrgangsmischung befragt. Zu diesem Zwecke wurden Auswahlantworten mit der Möglichkeit von Mehrfachnennungen vorgegeben, zum einen zum Sozialverhalten, zum anderen zur Lernentwicklung der am Schulversuch teilnehmenden Kinder.

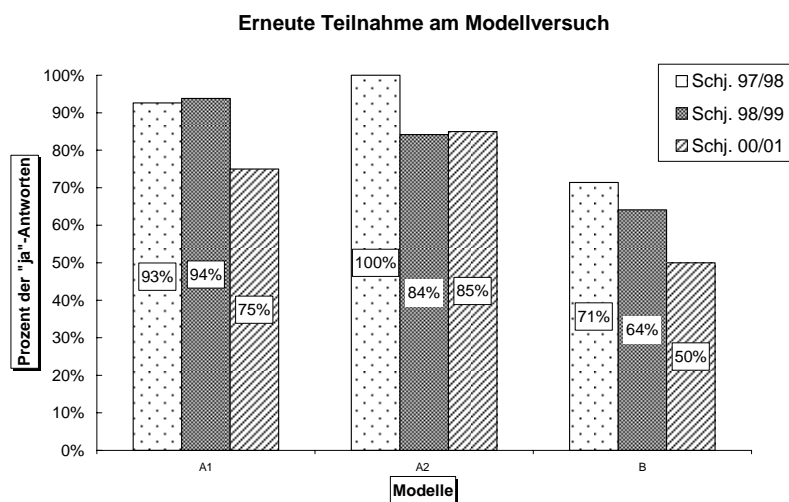
Durchgängig bei allen drei Erhebungen erhält im Falle des Sozialverhaltens die Antwortvorgabe „Die Jahrgangsmischung erleichtert durch Helfereffekte die Eingewöhnung der Schulanfänger in die neue Umgebung“ die höchste Zustimmung, die zweithöchste entfällt auf „Die Jahrgangsmischung bereichert die Sozialbeziehungen der Kinder untereinander“. Die geringste Zustimmung findet bei den Lehrerinnen der A-Modelle die Antwortmöglichkeit: „Durch die Jahrgangsmischung wird das Sozialleben in der Lerngruppe konfliktträchtiger“.

Was die Auswirkung der Jahrgangsmischung auf die Lernentwicklung der Kinder anbelangt, so wird die Verbreiterung des Spektrums der Lernmöglichkeiten und Lernwege (z.B. durch Mithören, Zuschauen) zu allen drei Erhebungszeitpunkten als ein Effekt wahrgenommen, der bei den Lehrerinnen mit Abstand die höchste Zustimmung erfährt, gefolgt von folgender Antwortmöglichkeit: „Die Jahrgangsmischung steigert infolge eines größeren Anregungsmilieus die Lernmotivation insbesondere der leistungsstarken Kinder“. Dass durch die Jahrgangsmischung der systematische Aufbau von fachbezogenem Lernen vernachlässigt werden könnte, findet die geringste Zustimmung bei den Lehrerinnen.

Eine positive Rückwirkung auf die Lernmotivation insbesondere der leistungsstarken Kinder infolge eines größeren Anregungsmilieus registrieren die Lehrerinnen als einen Effekt der Jahrgangsmischung, auf den die zweithöchste Zustimmung entfällt.

#### **Erneute Teilnahme am Schulversuch**

Da nach den erfragten Lehrerinnen- und Lehrerurteilen die Vorteile des Modellversuchs die Nachteile bei weitem übertreffen, ist es erklärlich, dass über alle drei Erhebungszeitpunkte hinweg durchschnittlich knapp 80 Prozent der Lehrerinnen und Lehrer an den A-Modellklassen eine erneute Teilnahme an dem Schulversuch bejahen (Abbildung 23). Bei der Erstbefragung fiel die Bereitschaft für eine erneute Teilnahme am Versuch bei den Lehrerinnen in den A-Modellklassen höher aus als beim dritten Erhebungszeitpunkt. Zur Begründung ihrer Entscheidung wurden den Lehrerinnen Auswahlantworten mit der Möglichkeit von Mehrfachnennungen vorgegeben.



**Abbildung 23: Erneute Teilnahme am Modellversuch**

Als häufigster Grund für die erneute Teilnahme wird genannt, dass der Schulversuch eine den individuellen Lern- und Leistungsvoraussetzungen der Kinder angepasste Förderung ermögliche. Als zweithäufigster Grund wird die Bereicherung des sozialen Lernens durch den Schulversuch von den Lehrerinnen der jahrgangsgemischten A-Modellklassen angegeben. Im Vergleich dazu entfielen auf die Antwortvorgabe „Durch die Teilnahme am Schulversuch wird die eigene Berufsarbeit interessanter“ nur wenige Nennungen.

Sofern eine erneute Teilnahme am Modellversuch abgelehnt wurde, wurde diese Entscheidung am häufigsten mit unzureichenden zeitlichen, räumlichen, personellen und materiellen Rahmenbedingungen begründet.

## Zusammenfassung und Diskussion der Ergebnisse der Hauptuntersuchung

Betrachtet man zusammenfassend die vorstehend berichteten Einzelergebnisse, so zeichnen sich zur Unterrichtsgestaltung Befunde ab, die die Aussagen des Zwischenberichts bekräftigen, aber auch davon abweichen.

Im Hinblick auf den Anteil des Frontalunterrichts am gesamten Unterricht bestätigen sich die Ergebnisse des Zwischenberichts insofern als in den A-Modellen in den Fächern Deutsch und Mathematik signifikant weniger Frontalunterricht erteilt wird als in den Vergleichsklassen. Das spricht dafür, dass durch die bei den A-Modellklassen vorhandene Jahrgangsmischung binnendifferenziertes Lernen begünstigt wird.

Dennoch lassen sich für die Mehrzahl der praktizierten Unterrichtsformen mit Differenzierungsanspruch (z.B. Partner- und Gruppenarbeit, Lernzirkel) für deren Auftretenshäufigkeit zwischen jahrgangsgemischten und jahrgangsgetrenten Klassen keine signifikanten Unterschiede feststellen, die während des Erhebungszeitraums stabil bleiben. Sofern diese nachweisbar sind, variieren sie zwischen den Erhebungszeitpunkten und in Abhängigkeit von den untersuchten Fächern und damit von curricularen Vorgaben.

Anders verhält es sich nur bei der Freiarbeit, deren Auftretenshäufigkeit nicht vom Fach, sondern von der Modellzugehörigkeit beeinflusst wird. Als einzige Unterrichtsform wird die Freiarbeit durchgängig in den drei Fächern Deutsch, Mathematik sowie Heimat- und Sachunterricht in den jahrgangsgemischten A-Modellklassen signifikant häufiger durchgeführt als in den Vergleichsklassen. Für die auffallende einseitige Bevorzugung der Freiarbeit gegenüber anderen Differenzierungsformen sind mehrere Erklärungsgründe denkbar. Zum einen lässt sich die Freiarbeit in einem breiten Konkretisierungsspektrum unter Einschluss von Partner-, Einzel- oder Gruppenarbeit realisieren, so dass letztere nicht mehr als eine von der Freiarbeit losgelöste und gesonderte Möglichkeit der Differenzierung von den befragten Lehrerinnen gesehen werden. Zum anderen könnte die Favorisierung der Freiarbeit in der aktuellen grundschulpädagogischen Diskussion die dominante Bevorzugung dieser Unterrichtsform erklären. Ungeachtet möglicher Begründungszusammenhänge geben die erhobenen Daten weder Auskunft über die didaktisch-methodische Qualität der Freiarbeit, noch über deren Wirkung auf die Leistungs- und Sozialentwicklung der Schüler und Schülerinnen.

Beim unterrichtlichen Umgang mit leistungsschwachen Schülern und Schülerinnen liefern die Daten im Hinblick auf die unterschiedlichen Fördermaßnahmen keinen klaren modellspezifischen Befund. Etwas anders präsentiert sich die Datenlage beim Umgang mit den Leistungsstarken. Bei den erfragten Differenzierungsvarianten fällt auf, dass in den A2-Modellklassen eine didaktisch anspruchsvollere Förderung durch Steigerung der Aufgabenschwierigkeit als in den Vergleichsklassen bevorzugt wird. Bei letzteren wie bei den B-Modellklassen überwiegt das Helferprinzip und damit wird eine Spielart der Differenzierung favorisiert, bei der sozialerzieherische Intentionen dominieren und eine gezielte aufgabenbezogene Förderung insbesondere der leistungsstarken Grundschüler vernachlässigt wird.

Bei der Bewertung des Modellversuchs durch die teilnehmenden Lehrerinnen selbst äußern diese eine auffallend hohe Akzeptanz gegenüber der veränderten Schuleingangsstufe, der sie als ausgesprochen positive Effekte die Förderung der Sozialerziehung, der kindlichen Eigenaktivität und die Erweiterung ihres eigenen didaktisch-methodischen Handlungsrepertoires zuschreiben. Eine Vernachlässigung des fachbezogenen systematischen Lernens durch die Jahrgangsmischung werten die Lehrerinnen nur zum ganz geringen Teil als Nachteil der neuen Schuleingangsstufe.

In der Sichtweise der Lehrerinnen wird die Jahrgangsmischung vornehmlich in einem Argumentationskontext positiv eingeschätzt, der mit einem ausgesprochen kindzentrierten Unterrichts- und Lernverständnis korrespondiert. Unter curricularem wie fachsystematischem Anspruch wird sie bei den befragten Lehrern und Lehrerinnen nicht wahrgenommen, weder in positiver noch in negativer Bewertung. Dieser Befund entspricht den in der aktuellen grundschulpädagogischen Literatur dominierenden Denkmustern, die die für jeden Unterricht notwendige Vermittlungsleistung zwischen Kind und Sache in ungleichgewichtiger Weise stärker zugunsten des Kindes thematisiert.

Für die breite Zustimmung wie die positive Einschätzung der Lehrerinnen zum Modellversuch „Schulanfang auf neuen Wegen“ lassen sich mehrere Gründe anführen. Da die Teilnahme am Modellversuch auf freiwilliger Basis stattfand, kann mit einiger Sicherheit angenommen werden, dass die Lehrerinnen der A-Modellklassen den Reforminitiativen von

vornherein positiv gegenüber standen und sich den aus der Jahrgangsmischung resultierenden pädagogischen und didaktischen Herausforderungen gewachsen fühlten. Zudem spricht eine Reihe der ausgewerteten Lehrerantworten dafür, dass der Modellversuch ein Schulverständnis repräsentiert, das den eigenen pädagogischen Überzeugungen und bildungspolitischen Vorstellungen der teilnehmenden Lehrkräfte entspricht. Was auch immer zutreffen mag, in jedem Fall geben die erhobenen Daten nur Auskunft über Effekte, wie sie von den befragten Grundschullehrerinnen wahrgenommen und eingeschätzt wurden, nicht über Effekte, die aus Daten direkt auf Schülerebene erhoben wurden.

Vergegenwärtigt man sich, dass es zu den Zielsetzungen des Modellversuchs „Schulanfang auf neuen Wegen“ gehörte, pädagogisch-didaktische Innovationen vornehmlich im Bereich des Anfangsunterrichts zu initiieren, so sprechen die Befunde für eine Realisierung des Ziels. Das zeigen die berichteten Ergebnisse zum Anteil des Frontalunterrichts am gesamten Unterricht sowie zur Freiarbeit, die den A-Modellklassen entsprechend der Heterogenität der Lernausgangslagen in weit größerem Maße als den Vergleichsklassen einen durch Differenzierungsmaßnahmen gekennzeichneten Unterricht bestätigen. Da nach den Selbstaussagen der Lehrerinnen die Teilnahme am Schulversuch eine Weiterentwicklung eigener didaktisch-methodischer Kompetenzen bewirkt hat, war mit dem Modellversuch ein Gewinn an Professionalität verbunden.

Ein weiteres Ziel des Schulversuchs war es, unter Verzicht auf Zurückstellungen und Schulfähigkeitsüberprüfungen alle dem Alter nach schulpflichtigen Kinder einzuschulen. Die dadurch sich vergrößernde Heterogenität der Schulanfänger wurde von am Schulversuch teilnehmenden Lehrerinnen an keiner Stelle als Erschwernis oder Belastung ihrer Berufsarbeit wahrgenommen. Das spricht dafür, dass die am Schulversuch teilnehmenden Lehrerinnen die mit der Unterschiedlichkeit der Entwicklungs- und Lernstände der Kinder einhergehenden Problemlagen didaktisch und methodisch bewältigen konnten.

Ein Vergleich der Befunde aus dem baden-württembergischen Modellversuch „Schulanfang auf neuen Wegen“ mit den in anderen Bundesländern im Laufe der 1990er Jahre initiierten Maßnahmen zur Reform der Schuleingangsstufe wird aus mehreren Gründen erschwert. Eine ganze Reihe von Modellversuchen zur Neukonzipierung der Schulanfangsphase wird ohne wissenschaftliche Begleitung durchgeführt, so dass über deren Effekte keine forschungsbasierten Daten zur Verfügung stehen (vgl. Faust-Siehl, 2001). <Die zu einigen Modellversuchen, wie etwa zu dem hessischen, vorliegenden Erfahrungsbericht von teilnehmenden Lehrerinnen veranschaulichen Reformanstrengungen zwar im konkreten Unterrichtsalltag, genügen jedoch nicht den Ansprüchen einer wissenschaftlichen Evaluation der Modellversuche.> Sofern diese überhaupt bei den länderspezifischen Schulversuchen durchgeführt werden, unterscheiden sie sich von der baden-württembergischen Begleituntersuchung nicht nur in den Erhebungsinstrumentarien, der Untersuchungsdauer und dem Stichprobenumfang, sondern auch durch ein Untersuchungsinteresse, das standardisierte Erhebungen zur Unterrichtsgestaltung unter den Bedingungen der veränderten Schuleingangsstufe weitestgehend ausschließt.

Dasselbe gilt für die empirischen Forschungsprojekte, die losgelöst von der Schuleingangsproblematik das altersgemischte Lernen in der Grundschule untersuchen. Da sie sich vorrangig auf die Effekte konzentrieren, die die Altersmischung auf die Leistungsentwicklung und das Sozialverhalten der Grundschülerinnen haben, bleiben Fragen des Unterrichtsarrangements bei altersgemischten Lerngruppen unberücksichtigt (vgl. Knörzer, 1985; Laging, 1995).



## Anschlussuntersuchung in der dritten und vierten Jahrgangsstufe

In Ergänzung zu den auf die Eingangsstufe konzentrierten Erhebungen der Hauptuntersuchung wurden die Lehrerinnen und Lehrer der jahrgangsgetrenten 3. und 4. Grundschulklassen zu ihrer Unterrichtsgestaltung befragt, um Auskunft über mögliche Auswirkungen der modellspezifischen Besonderheiten wie der Jahrgangsmischung auf die nachfolgenden Grundschulklassen zu erhalten.

Zu diesem Zwecke wurden die jeweils für die A-Modellklassen, die B-Modellklassen und die Vergleichsklassen konstruierten Lehrerfragebogen mit geringfügigen jahrgangsbedingten Abänderungen in den versuchsnachfolgenden dritten und vierten Jahrgangsstufen erneut eingesetzt.

Die Erhebung wurde in den dritten Grundschulklassen gegen Ende des Schuljahres 2001/02 und in den vierten Grundschulklassen gegen Ende des Schuljahres 2002/03 durchgeführt. Differenziert nach Jahrgangsstufen und nach der früheren Modellzugehörigkeit der Dritt- und Viertklässler wurden folgende Fragebogen ausgewertet:

### 3. Jahrgangsstufe - Schuljahr 2001/02

30 Fragebogen von Drittklasslehrerinnen mit Schülern aus jahrgangsgemischten A-Modellklassen, 10 Fragebogen von Drittklasslehrerinnen mit Schülern aus B-Modellklassen, 15 Drittklasslehrerinnen der versuchsnachfolgenden Vergleichsklassen

### 4. Jahrgangsstufe- Schuljahr 2002/03

27 Fragebogen von Viertklasslehrerinnen mit Schülern aus jahrgangsgemischten A-Modellklassen, 9 Fragebogen von Viertklasslehrerinnen mit Schülern aus B-Modellklassen, 14 Viertklasslehrerinnen der versuchsnachfolgenden Vergleichsklassen

Die Auswertung der Lehrerfragebogen zeigt, dass in den versuchsnachfolgenden Grundschulklassen eine Unterrichtsgestaltung praktiziert wird, die sich nur geringfügig von derjenigen in den Vergleichsklassen unterscheidet.

Der Anteil des nichtfachlichen Lernens an der wöchentlichen Unterrichtszeit ist in der dritten Klasse mit A-Modellkindern auf einem statistisch signifikanten Niveau größer als in Klassen mit B-Modellkindern. Dieser Unterschied tritt nur modellintern auf, jedoch nicht zwischen Modell- und Vergleichsklassen. Im vierten Grundschuljahr sind überhaupt keine signifikanten Unterschiede mehr feststellbar.

Was den Anteil frontaler Unterrichtsphasen an der wöchentlichen Arbeitszeit in den Kernfächern Deutsch, Mathematik und Heimat- und Sachkunde anbelangt, so liefert die Auswertung der Fragebogen für die dritte und vierte Klasse Ergebnisse, die von denen der am Modellversuch teilnehmenden Schuleingangsklassen abweichen (vgl. 87 ff.). Während dort je nach Unterrichtsfach signifikante Unterschiede zwischen den A-Modellklassen und den Vergleichsklassen zu verzeichnen sind, entfallen diese bei den versuchsnachfolgenden Klassen.

Fragt man jenseits der frontalen Unterrichtsphasen nach der Auftretenshäufigkeit verschiedener Unterrichtsformen in den genannten Grundschulfächern, so kann für deren Mehrzahl in der Regel kein signifikanter Unterschied zwischen Klassen mit Modellkindern und den Vergleichsklassen nachgewiesen werden, wohl aber modellinterne Unterschiede, die jedoch nicht durchgängig deckungsgleich mit den in der Hauptuntersuchung erhobenen Unterschieden sind (vgl. S. 89 f.).

Ein solcher tritt in Abweichung vom Befund der Hauptuntersuchung (vgl. S. 88) beim Kreisgespräch im Fach Deutsch auf, das in der dritten und vierten Jahrgangsstufe in Klassen mit A-Modellkindern signifikant häufiger praktiziert wird als in solchen mit B-Modellkindern, aber auch in Vergleichsklassen signifikant häufiger als in dritten und vierten Grundschulklassen mit B-Modellkindern.

Im Deutschunterricht wird weiterhin die freie Lesestunde im dritten und vierten Schuljahr mit A-Modellkindern signifikant häufiger realisiert als in Vergleichsklassen, während der Lernzirkel nur im dritten Schuljahr in den Klassen mit A-Modellkindern signifikant häufiger auftritt als in Klassen mit B-Modellkindern. Aber auch hier ist die Auftretenshäufigkeit in den Vergleichsklassen signifikant höher als in den B-Klassen.

Im Fach Mathematik wird die freie Übungsstunde in den beiden versuchsnachfolgenden Schuljahren in den Vergleichsklassen wie in den Klassen mit B-Modellkindern signifikant häufiger praktiziert als in den dritten und vierten Klassen mit A-Modellkindern.

Im Heimat- und Sachunterricht ist im dritten Schuljahr bei den Klassen mit A-Modellkindern das Kreisgespräch signifikant häufiger anzutreffen als in Klassen mit B-Modellkindern, ein Unterschied, der im vierten Schuljahr nicht mehr nachweisbar ist. Hier findet das Kreisgespräch signifikant häufiger bei Klassen mit A-Modellkindern statt als in Vergleichsklassen. In der vierten Grundschulklasse wird der Lernzirkel im Heimat- und Sachunterricht in Klassen mit A-Modellkindern häufiger durchgeführt als in Klassen mit B-Modellkindern. Beim Lerngang als einer sachunterrichtsspezifischen Unterrichtsform ist kein Unterschied in der Auftretenshäufigkeit in den versuchsnachfolgenden Klassen nachweisbar.

Die Frage, ob die Existenz von Kindern aus jahrgangsgemischten Eingangsstufen in der dritten und vierten Klasse zeitorganisatorische Veränderungen des Unterrichts nach sich zieht, liefert keine statistisch signifikanten Ergebnisse mehr, wie sie im Falle der jahrgangsgemischten Eingangsstufe selbst zu verzeichnen waren (vgl. S.94).

Sofern zeitorganisatorische Veränderungen im Tagesablauf vorgenommen wurden, konkretisieren sie sich im dritten und vierten Grundschuljahr mit A-Modellkindern wie B-Modellkindern bevorzugt im „Wechsel zwischen lehrergesteuerten Phasen und Freiarbeit“, zwischen „Arbeits- und Entspannungsphasen“ sowie in der „Lockerung des 45-Minuten-Taktes“. Im Wochenablauf gehören in denselben Jahrgangsstufen die „Arbeit nach dem Wochenplan“ sowie in der „Auflockerung des Wochenrhythmus durch musisch bestimmte Phasen“ zu den favorisierten zeitorganisatorischen Veränderungen.

In den Förderpräferenzen von leistungsschwachen und leistungsstarken Schülern lassen sich in den dritten und vierten Klassen mit A- wie B-Modellkindern keine statistisch bedeutsamen Unterschiede nachweisen, auch nicht zu den entsprechenden Vergleichsklassen. Bei der Förderung leistungsschwacher Kinder tritt in der dritten und vierten Jahrgangsstufe speziell mit A-Modellkindern eine Bevorzugung qualitativer Differenzierungsmaßnahmen (Variation des Aufgabenniveaus) gegenüber quantitativen (Variation von Zeit- und Aufgabenpensum) auf, was sich auch beim Umgang mit leistungsstarken Grundschulern wiederholt. Hier wird bei den versuchsnachfolgenden B- und Vergleichsklassen als die am häufigsten praktizierte Fördermaßnahme die freie Aufgabenwahl durch die Kinder angegeben.

Da es sich bei den befragten Dritt- und Viertklasslehrerinnen in der Regel um Lehrkräfte handelte, die erstmals Kinder aus jahrgangsgemischten Vorläuferklassen unterrichteten, war für die Untersuchung ein Erfahrungsvergleich von Interesse. Daher wurden über die Unterrichtsgestaltung hinaus die Dritt- und Viertklasslehrerinnen mit A-Modellkindern zusätzlich gebeten, auf der Basis eigener Erfahrungen in Klassen ohne und in Klassen mit Kindern aus der jahrgangsgemischten Eingangsstufe Veränderungen zu identifizieren und

zwar bezogen auf das Sozialverhalten, das Lern- und Leistungsverhalten der Kinder und bezogen auf die eigene Unterrichtsvorbereitung.

Mit Blick auf das Sozialverhalten der Kinder stellt die Hälfte der befragten Lehrerinnen im dritten und im vierten Schuljahr Veränderungen fest, die auf die Herkunft der Kinder aus der jahrgangsgemischten Eingangsstufe zurückgeführt werden. Dabei werden mehrheitlich sozialverträgliche Verhaltensweisen wie Hilfsbereitschaft, Toleranz und Rücksichtnahme von den Lehrerinnen als positive Effekte hervorgehoben. Nur vereinzelt wird eine egoistische Haltung bei den Kindern als ein negativer Folgeeffekt der jahrgangsgemischten Eingangsstufe registriert.

Für die übrige Hälfte der befragten Lehrerinnen hat das Durchlaufen der jahrgangsgemischten Eingangsstufe weder positive noch negative Veränderungen im Sozialverhalten der Dritt- und Viertklässler zur Folge.

Was das Lern- und Leistungsverhalten der Schüler anbelangt, so werden im dritten Schuljahr von mehr als zwei Dritteln und im vierten Schuljahr noch von etwa der Hälfte der Lehrerinnen Veränderungen wahrgenommen. Die größere Selbständigkeit und die realistischere Selbsteinschätzung der Schüler werden von einer ganzen Reihe von Lehrerinnen als positive Nachfolgeeffekte der Jahrgangsmischung genannt. Vereinzelt wird auch die Erfahrung mitgeteilt, dass intelligente Schüler mehr profitieren als leistungsschwache und manche Schüler lange brauchen, um sich an verbindliche Leistungsanforderungen zu gewöhnen.

Mehr als die Hälfte der befragten Dritt- und Viertklasslehrerinnen geben an, dass die Existenz von Kindern aus jahrgangsgemischten Vorläuferklassen zu Veränderungen in der eigenen Unterrichtsvorbereitung und -gestaltung geführt hat. Sie konkretisieren sich in einem erhöhten Zeitaufwand für die Vorbereitung des Unterrichts sowie in einer vermehrten Praktizierung offener Unterrichtsformen (Wochenplanarbeit, Freiarbeit, Projekte) wie auch in einer Intensivierung der Einzelförderung.

Die Frage, ob die Jahrgangsmischung in der dritten und vierten Jahrgangsstufe fortgeführt werden sollte, wird von knapp über 50 Prozent der Drittklasslehrerinnen und von 43 Prozent der Viertklasslehrerinnen bejaht.

Obwohl in der Wahrnehmung der Dritt- und Viertklasslehrerinnen positive Nachfolgeeffekte der Jahrgangsmischung auf das Sozial-, das Lern- und Leistungsverhalten der Kinder überwiegen, trifft eine Fortführung der Jahrgangsmischung im dritten und vierten Grundschuljahr auf eine deutlich geringere Akzeptanz als bei den Lehrkräften der Eingangsstufe selbst. Demnach nimmt mit aufsteigender Jahrgangsstufe und - darin inbegriffen - mit der Zunahme fächerteiligen Unterrichts und steigenden Leistungsanforderungen die Akzeptanz der Grundschullehrerinnen gegenüber der Jahrgangsmischung ab.

## II.9 Unterrichtsgestaltung Grundschulförderklassen

Die Grundschulförderklasse besuchen vom Alter her schulpflichtige Kinder, die vom Schulbesuch zurückgestellt sind (vgl. § 74 Abs. 2 und 3 SchG). Die Zeit der Zurückstellung, ein Schuljahr, wird auf die Dauer der Pflicht zum Besuch der Grundschule nicht angerechnet. Grundschulförderklassen haben die Aufgabe, die zurückgestellten Kinder auf den Besuch der Grundschule vorzubereiten. Sie werden an Grundschulen geführt. Der Leiter der Grundschule ist zugleich Leiter der Förderklasse (vgl. § 5a SchG). Der Besuch dieser Einrichtung wird zurückgestellten Kindern empfohlen.

Der Schwerpunkt des B-Modells im Rahmen des Schulversuchs zur Flexibilisierung des Schuleingangs ist die enge Verzahnung von Grundschulförderklassen mit den ersten Klassen. Es wird davon ausgegangen, dass es in der Grundschulförderklasse und in der 1. Klasse eine Gruppe von Kindern mit vergleichbarem Förderbedarf gibt, die in einer Gruppe zusammenarbeiten können. Diese Gruppe wird im Folgenden als „spezifische Lerngruppe“ (sLG) bezeichnet.

Kennzeichnend für die Arbeit in dieser integrativen Lerngruppe ist eine intensive Kooperation des Personals beider Einrichtungen, also zwischen Erzieherin/Sozialpädagogin und Lehrer/Lehrerin, diese soll sich niederschlagen

- im Erstellen gemeinsamer Förderpläne für Kinder aus beiden Einrichtungen,
- im Austausch von Kindern zwischen erster Klasse und Grundschulförderklasse zu gemeinsamen Angeboten,
- in einer zeitlich flexiblen Regelung von Übergängen: Kinder der Grundschulförderklasse sollen je nach individuellem Entwicklungsfortschritt gleitend auch während des Schuljahres in die 1. Klasse überwechseln können.

Aufgabe der wissenschaftlichen Begleitung war es daher, die jeweiligen Organisationsformen sowie die didaktischen und methodischen Konzepte der Förderung in dieser „spezifischen Lerngruppe“ zu untersuchen. Insbesondere sollte geprüft werden, ob dieser Ansatz einer Verzahnung der Grundschulförderklasse mit der ersten Klasse eine konzeptionelle Weiterentwicklung der Schuleingangsphase darstellt, vom Schulbesuch zurückgestellte Kinder angemessen fördert und Kindern während des laufenden Schuljahres einen Übertritt aus der Grundschulförderklasse - im Folgenden mit GFK abgekürzt - in die erste Klasse erlaubt.

Im Untersuchungszeitraum kristallisierten sich zwei Modelle der Förderung heraus:

Das oben beschriebene, im Verlauf der Untersuchung mit B1-Modell bezeichnete, und das sog. B2-Modell (siehe auch I.3).

In Abhebung zum Förderansatz im B1-Modell werden im B2-Modell Kinder, die allem Anschein nach vom Schulbesuch zurückgestellt werden, ein halbes Jahr vor Schulbeginn in einer dafür konzipierten Grundschulförderklasse auf die Schule vorbereitet. Die Erzieherin/Sozialpädagogin begleitet und fördert diese Kinder auch im ersten Halbjahr der ersten Grundschulklasse. Diese Modellvariante erspart einer Reihe von Kindern ein ganzes Jahr, das sie sonst mit dem Besuch der Grundschulförderklasse alten Stils hätten aufwenden müssen, ganz abgesehen von den emotionalen Problemen, die mit einer nicht „normalen“ Einschulung oder gar einer Ausschulung einhergehen können.

Da bis zum Erhebungszeitraum 1999/2000 nur drei in die Begleitung einbezogene Grundschulförderklassen als B2-Modelle durchgeführt wurden, konnten diese nicht in die Begleituntersuchung einbezogen werden; ihre Strukturdaten und die Aussagen der beteiligten Pädagogen sind weiter unten zusammenfassend aufgeführt.

### **Erhebungsinstrumentarium**

Die wissenschaftliche Begleitung entwickelte einen auf die Untersuchungsziele ausgerichteten Fragebogen, der im Verlaufe der Begleitung erstmals im Schuljahr 1997/1998 an B-Modellklassen mit einer angeschlossenen GFK zum Einsatz kam und zum zweiten Mal im Schuljahr 1999/2000 - allerdings in modifizierter Form für B1- oder B2-Modelle –an die Grundschulförderklassen versandt wurde. In beiden Fassungen waren den Fragen Auswahlantworten zugeordnet, bei manchen Fragen waren Mehrfachnennungen möglich; nur im Bogen für die Ersterhebung in 1997/1998 konnten Fragen nach den Vor- bzw. Nachteilen des Modellversuchs offen beantwortet werden.

Die Fragen zur spezifischen Lerngruppe sollten vom Team Erzieherin/Sozialpädagogin - Lehrerin/Lehrer gemeinsam beantwortet werden. Die nachfolgend dargestellten Untersuchungsergebnisse beruhen auf Selbstauskünften von Erzieherinnen/Sozialpädagoginnen und von Lehrerinnen und Lehrern der ersten Klassen.

### **Erhebungszeitpunkte und Stichprobe**

Der Fragebogen zur Unterrichtsgestaltung der Grundschulförderklassen im B-Modell: Grundschulförderklassen neuen Zuschnitts wurde in den Einschulungsjahrgängen 1997/1998 und 1999/2000 jeweils im Frühsommer versandt.

Im Einschulungsjahrgang 1997/1998 waren 29 Fragebogen Grundlage für die Auswertung. Die Befunde sind im Zwischenbericht (vgl. Arbeitskreis, 2000) niedergelegt.

Im Einschulungsjahrgang 1999/2000 standen 33 ausgefüllte Fragebogen zur Unterrichtsgestaltung an Grundschulförderklassen zur Verfügung: 30 Fragebogen aus dem B1-Modell, 3 Fragebogen aus dem B2-Modell.

Zunächst werden die Untersuchungsergebnisse zum B1-Modell aus der zweiten Befragung im Schuljahr 1999/2000 dargestellt und mit denen aus der Erstbefragung zusammengeführt; danach folgt eine Zusammenfassung der Einschätzungen der B2-Modelle durch die beteiligten Erzieherinnen/Sozialpädagoginnen. Hierzu liegen, wie bereits ausgeführt, nur Daten aus der Befragung im Einschulungsjahrgang 1999/2000 vor.

### **Ergebnisse: Förderung von Kindern im B1-Modell**

#### Fördergründe und Förderumfang

Die „spezifische Lerngruppe“ besuchen Kinder aus der schulvorbereitenden Einrichtung Grundschulförderklasse und dem ersten Schuljahr mit vergleichbarem Förderbedarf zu gemeinsamen Förderangeboten. Zu beiden Erhebungszeitpunkten wurden die Pädagogen der jeweiligen Einrichtung nach den Fördergründen befragt, Kinder in eine sLG aufzunehmen.

Im Erhebungszeitraum 1999/2000 gaben die Erzieherinnen/Sozialpädagoginnen der GFK als die vier häufigsten Gründe für die Aufnahme von Kindern in eine sLG - sozusagen eine Überweisung „nach oben“ - an:

- eine altersgerechte Entwicklung der Wahrnehmung (am häufigsten benannt),
- ein ausreichendes Konzentrationsvermögen und
- ein Interesse an den „Sachen“ des Umfeldes (mit gleichen Häufigkeiten an zweiter Stelle benannt),
- ein angemessenes Arbeitsverhalten/stabile Lernmotivation (am vierthäufigsten benannt).

Voraussetzungen wie ein „angemessenes Sozialverhalten“, eine „altersgerechte Entwicklung der Motorik“ sowie „Stärken in einer der Kulturtechniken“ wurden weniger häufig benannt.

**Abbildung 24: Gründe für die Aufnahme in die spezifische Lerngruppe, Schj. 99/00 (Erzieherinnen)**

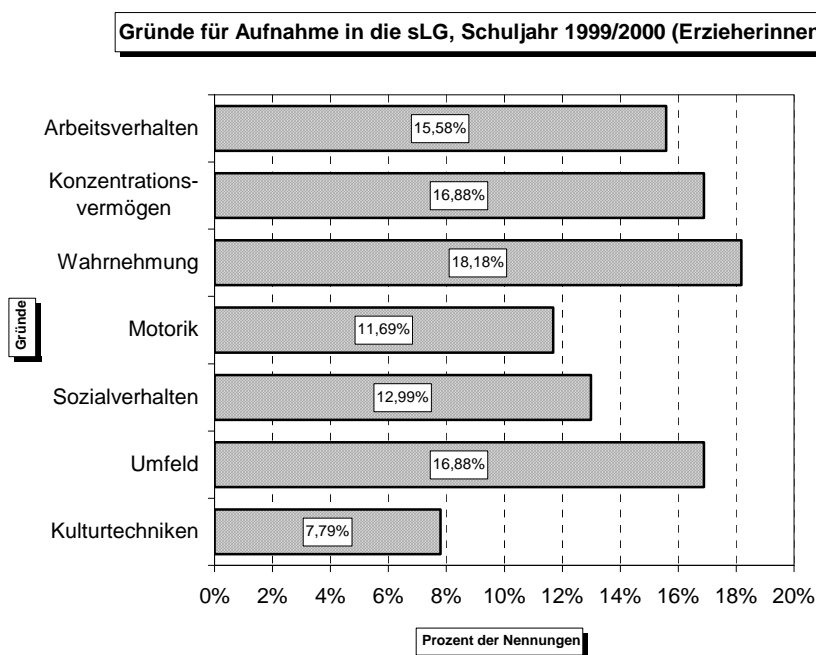
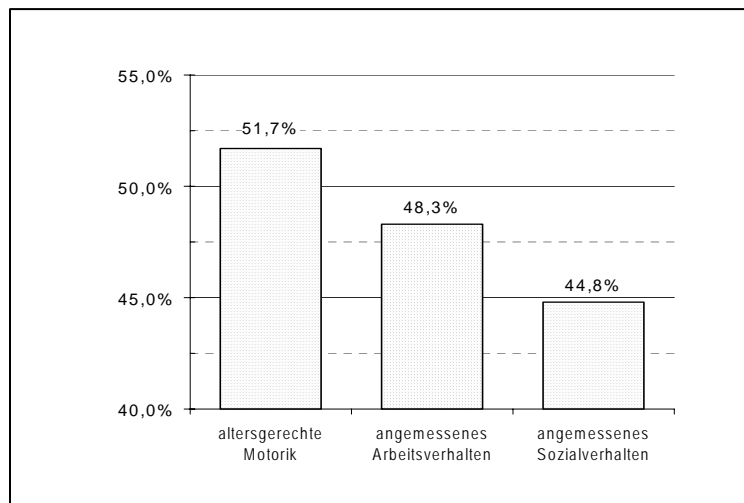


Abbildung 25 (nächste Seite) zeigt, dass in der Erhebung im Schuljahr 1997/1998 für Erzieherinnen/Sozialpädagoginnen die drei häufigsten Gründe für eine Aufnahme von GFK-Kindern in eine sLG waren:

- eine altersgerechte Motorik,
- ein angemessenes Arbeitsverhalten,
- ein angemessenes Sozialverhalten.

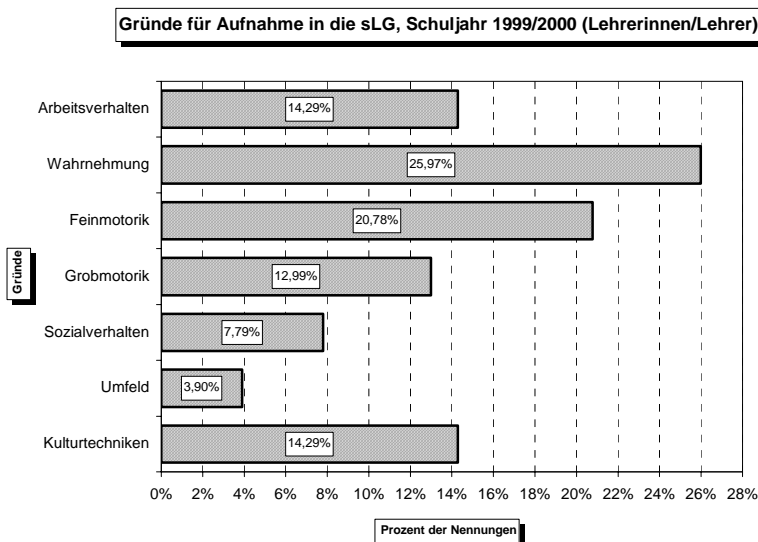
**Abbildung 25: Drei häufigsten Gründe für eine Aufnahme von GFK-Kindern in eine spezifische Lerngruppe (sLG), 1997/98**



Aus der folgenden Abbildung 26 geht hervor, dass im Erhebungszeitraum 1999/2000 die Erstklasslehrerinnen folgende Fördergründe für die Aufnahme von Schülerinnen und Schülern der 1. Klassen in eine sLG - sozusagen eine Überweisung „nach unten“ - anführten:

- Defizite in der Wahrnehmung, vor allem in den visuellen und akustischen Funktionen (weitaus am häufigsten benannt),
- Defizite bei Aufgaben, die feinmotorische Grundfertigkeiten voraussetzen,
- Arbeitsverhalten und Lernmotivation müssen aufgebaut werden und
- Schwächen in einer der Kulturtechniken sowie
- Entwicklungsrückstände in einzelnen grobmotorischen Grundfertigkeiten (mit gleichen Häufigkeiten als dritte benannt).

**Abbildung 26: Gründe für die Aufnahme in die spezifische Lerngruppe, Schj. 99/00 (Lehrerinnen/Lehrer)**



Weitaus weniger häufig wurde von den Lehrkräften genannt:

- Das manifeste soziale Verhalten erschwert die Bildbarkeit in der Gruppe.
- Mangelnde Ausgeschlossenheit gegenüber den „Sachen“ im persönlichen Umfeld (Inhalte und Themen des Heimat- und Sachunterrichts).

In 1997/1998 gaben Erstklasslehrerinnen als die drei häufigsten Gründe für eine Aufnahme von Erstklässlern in eine sLG „Konzentrationsprobleme“, „motorische Störungen“, „Wahrnehmungsstörungen“ an. „Mangelndes Sozialverhalten“ und „mangelndes Arbeitsverhalten“ wurden ebenfalls häufig genannt.

### Bereiche und Sozialformen der Förderung

Die Befragung in 1999/2000 verlangte von den Pädagoginnen der beiden Einrichtungen Angaben zu den Methoden bzw. zu den Sozialformen der Förderung der Kinder einer sLG. Mittels einer 4-stufigen Skala („überwiegend-häufig-selten-nie“) konnten die Erzieherinnen und die Lehrkräfte der 1. Klassen angeben, in welchen Bereichen und in welchen Sozialformen die Kinder der „spezifischen Lerngruppe“ gefördert wurden.

Fasst man die Kategorien „überwiegend“ und „häufig“ zusammen, ergibt sich folgende Rangreihe:

- Förderung in Kleingruppe(n) (87%)
- Abgesprochene Aktivität wird parallel durchgeführt (78%)
- Übernahme bestimmter Aktivitäten in einer sLG (68,8%)
- Einzelförderung (68,4%).

In 1997/1998 setzten die Erzieherinnen und Lehrkräfte der 1. Klassen folgende Methoden bzw. Sozialformen der Förderung in einer sLG ein (Kategorien „überwiegend“ und „häufig“ zusammengefasst):

- Erste Ansätze zur Gruppenarbeit (68,5%)
- Einsatz von Trainingsprogrammen (53,8%)
- Partnerarbeit mit einem Schüler/einer Schülerin aus Klasse 1 (57,7%).

### Förderung in Lernbereichen

In 1999/2000 wurde von den Pädagoginnen aus beiden Einrichtungen erfragt, in welchen Lernbereichen eine Förderung unter Einsatz dieser Methoden und Sozialformen stattfand. Werden die Kategorien „überwiegend“ und „häufig“ zusammengefasst, ergibt sich folgende Rangreihe:

- im Lernbereich Deutsch (18 Nennungen bei N=21)
- im Lernbereich Sport (15 Nennungen bei N=17)
- im Lernbereich Bildende Kunst/Textiles Werken (12 Nennungen bei N= 14)
- im Lernbereich Mathematik (11 Nennungen bei N=17)
- im Lernbereich Musik (10 Nennungen bei N=11)
- im Lernbereich Heimat- und Sachunterricht (9 Nennungen bei N=11).

(In 1997/1998 nicht erfragt.)



## Kooperation zwischen den beteiligten Pädagogen

In 1999/2000 wurden Erzieherinnen/Sozialpädagoginnen der GFK nach Kooperationspartnern und Kooperationsformen außerhalb der „spezifischen Lerngruppe“ befragt.

In der Zusammenarbeit mit Erzieherinnen der Kindertageseinrichtung führten sie als die drei häufigsten Kooperationsformen an (nur Kategorie „regelmäßig“ erfasst):

- Hospitation (40%),
- Fallbesprechungen (27,8%),
- Entwicklung von Diagnose- und Förderplänen (27,3%).

„Austausch von Materialien“ (10%) sowie „Gemeinsame Elterngespräche“ (9%) werden deutlich weniger benannt.

In der Zusammenarbeit mit Erstklasslehrerinnen führten sie als die drei häufigsten Kooperationsformen an (nur Kategorie „regelmäßig“ erfasst):

- Fallbesprechungen (64,3%),
- Entwicklung von Diagnose- und Förderplänen (61,5%),
- Austausch von Materialien (57,7%).

„Hospitation“ wurde mit 54,2% fast gleich häufig genannt. „Gemeinsame Elterngespräche“ mit 24% fielen dagegen ab.

In 1997/1998 wurden als die drei häufigsten Formen der Zusammenarbeit angeführt (nur Kategorie „regelmäßig“ erfasst):

- informeller Erfahrungsaustausch/Gespräche (79,3%),
- Austausch von Materialien (63%),
- Gemeinsame Vorbereitung von Förderangeboten (61,5%).
- „Gegenseitige Hospitationen“ (36%) wurden weniger oft genannt.

Zwei Drittel der Befragten (66,7%) gaben an, dass eine Kooperation mit anderen Lehrerinnen und Lehrern der Schule bestand.

## Organisationsformen

Zu beiden Erhebungszeitpunkten wurden Daten zur Organisation der Förderung in den „spezifischen Lerngruppen“ erfragt.

Einrichtung/Ort der Förderung:

In 1999/2000 war überwiegend (62%) die Grundschulförderklasse der Förderort.

In 1997/1998 fand die Förderung in mehr als der Hälfte der Fälle (54%) in einer „Extra-Gruppe“ statt.

An der Förderung beteiligte Pädagogen:

In 1999/2000 (Mehrfachnennungen möglich) wirkten an der Förderung der Kinder einer „spezifischen Lerngruppe“ hauptsächlich Erzieherinnen/Sozialpädagoginnen der Grundschulförderklasse (43,3%), Erstklasslehrerinnen (31,7%), Lehrkräfte einer anderen Klasse oder Lehrkräfte von Sonderschulen mit (zusammen 25% der Nennungen).

In 1997/1998 waren Erzieherinnen/Sozialpädagoginnen (65,8%), Lehrerinnen der Grundschule (18%) sowie Lehrer/Lehrerinnen von Sonderschulen (16%) an der Förderung in einer sLG beteiligt.

Anzahl der „spezifischen Lerngruppen“ in einer Einrichtung seit Schuljahresbeginn:

In 1999/2000 reichte im Verlauf eines Schuljahres die Anzahl der sLG von einer bis zu 10 Gruppen; im Durchschnitt bestanden 3 bis 4 „spezifische Lerngruppen“ je Einrichtung. (Aus 1997/1998 liegen keine Angaben vor.)

Durchschnittliche Gruppengröße:

In 1999/2000 lag diese zwischen 3 und 10 Kindern. Aus den Grundschulförderklassen nahmen bis zum Erhebungszeitpunkt (Mai 2000) durchschnittlich bis zu 5 Kinder an Fördermaßnahmen der sLG teil. Offensichtlich wechselten die Kinder je nach Förderangebot. In 1997/1998 lag diese zwischen 4 und 11 Kindern. Aus den Grundschulförderklassen ebenso wie aus den 1. Klassen nahmen bis zum Erhebungszeitpunkt (Frühsommer 1998) durchschnittlich 9 Kinder an Fördermaßnahmen der sLG teil. Offensichtlich wechselten die Kinder je nach Förderangebot.

Übertritt von Kindern der Grundschulförderklasse in die 1. Klasse der Grundschule

#### **Übertritt während des Schuljahres**

In 1999/2000 trat ein Kind aus der Grundschulförderklasse, das an der Förderung in einer „spezifischen Lerngruppe“ teilnahm, nach den Osterferien in die 1. Klasse der Regelschule über. Nach Einschätzung der Fachkraft hatte dieses - Erstsprache Amerikanisch - bis dahin genügend Kenntnisse in der deutschen Sprache erworben, um am Unterrichtsgeschehen der 1. Klasse teilnehmen zu können.

(In 1997/1998 nicht erfragt.)

#### **Übertritt im Anschluss an den einjährigen Besuch der Grundschulförderklasse**

In 1999/2000 wurden die Erzieherinnen/Sozialpädagoginnen gefragt, ob am Schuljahresende alle Kinder aus der GFK in die erste Klasse der Grundschule eingeschult werden. Nach ihren Angaben werden aus 10 Grundschulförderklassen (34,5%) jeweils alle Kinder in die erste Klasse eingeschult, aus 19 (65,5%) Grundschulförderklassen jedoch nicht alle Kinder. Als Bildungsstätte für diese Kinder wird am häufigsten die Förderschule angegeben.

Hier ist anzumerken, dass die bearbeiteten Fragebogen im Zeitraum Anfang Juli bis Mitte August 2000 beim Landesinstitut eingingen, d.h. einige vor Schuljahresende ausgefüllt wurden. Dennoch dürfte bereits zu diesem Zeitpunkt entschieden sein, welche Bildungseinrichtung die Kinder im Anschluss an die GFK besuchen werden.

(In 1997/1998 nicht erfragt.)

## Einstellungen der teilnehmenden Erzieherinnen/Sozialpädagoginnen

### **Vor- und Nachteile des Modellversuchs aus der Sicht der teilnehmenden Erzieherinnen/Sozialpädagoginnen**

Im Erhebungszeitraum 1999/2000 sollten die Erzieherinnen/Sozialpädagoginnen aus einer 12 (Vorteile) bzw. 15 (Nachteile) Items umfassenden Liste jeweils vier Vor- bzw. vier Nachteile des Modellversuchs ankreuzen.

Als Vorteile des Modellversuchs wurden am häufigsten benannt (Rangreihe der Nennungen):

- Durch die Zusammenarbeit von Erzieherin/Sozialpädagogin und Lehrerin in der 1. Klasse gehen eine breitere Fachkompetenz und Erfahrung in den Unterricht ein.
- Es ist eine kontinuierliche Beobachtung von Kindern möglich; unmittelbar anschließend sind Gespräche über einzelne Kinder sowie das Einleiten von Fördermaßnahmen möglich.
- Die Teamarbeit der beiden Pädagogen bringt mehr Ideen, methodische Abwechslung und Reflexion in den Unterrichtsvormittag. Und:
- Es wird ein problemloser Übergang von der Grundschulförderklasse in die Grundschule möglich.

In 1997/1998 – hier waren keine Auswahlantworten vorgegeben - hoben die Erzieherinnen/Sozialpädagoginnen als Vorteile hervor:

- (Es sind) eine gezieltere Beobachtung der Kinder und Gespräche über schwierige Kinder möglich.
- (Sie erhalten) Einblick in Lerninhalte, Lernmethoden und Lerntechniken der ersten Klassen.
- (Es sind) gezieltere Fördermaßnahmen möglich.

In 1999/2000 wurden als Nachteile des Modellversuchs am häufigsten benannt (Rangreihe der Nennungen):

- Für die Planung/Vorbereitung der Teamarbeit und anderer Abstimmungsmaßnahmen ist ein großer Zeitaufwand erforderlich. Und:
- Zu wenig Stunden für die Kooperation mit der Erstklasslehrerin.
- Die Mitarbeit in mehreren 1. Klassen bzw. die Kooperation mit mehreren Klassenlehrerinnen der 1. Klassen erfordert einen großen Organisationsaufwand.
- Erstklässler machen schnellere Fortschritte als die Kinder der Grundschulförderklasse; dadurch ist es schwierig, lernhomogene Gruppen zu bilden.

Das unter „Nachteile der B 1-Modelle“ aufgeführte Item „Der Besuch der Grundschulförderklasse verlängert die Dauer des Schulbesuchs“ wurde nicht angekreuzt.

In 1997/1998 wurden von den Erzieherinnen/Sozialpädagoginnen als Nachteile des Modellversuchs genannt (Rangreihe der Nennungen):

- Großer Zeitaufwand durch notwendige Absprachen, und:
- Erstklässler machen schnellere Fortschritte als die Kinder der Grundschulförderklasse, dadurch ist es schwierig, lernhomogene Gruppen zu bilden.
- Koordinationsprobleme der beteiligten Pädagoginnen, und:

- Manche Beschäftigungen der Grundschulförderklasse, z.B. Rollen-, Sing- und Bewegungsspiele, können wegen zu geringer Schülerzahl nicht mehr stattfinden.

In beiden Erhebungszeiträumen überwogen bei den Nennungen der Vor- und Nachteile des Modellversuchs die Vorteile.

Als Vorteile des Modellversuchs heben die Erzieherinnen/Sozialpädagoginnen in beiden Erhebungszeiträumen heraus, dass eine kontinuierliche und gezieltere Beobachtung der Kinder möglich wird, unmittelbar daran anschließend Fallbesprechungen stattfinden und „passende“ Förderangebote eingeleitet werden können.

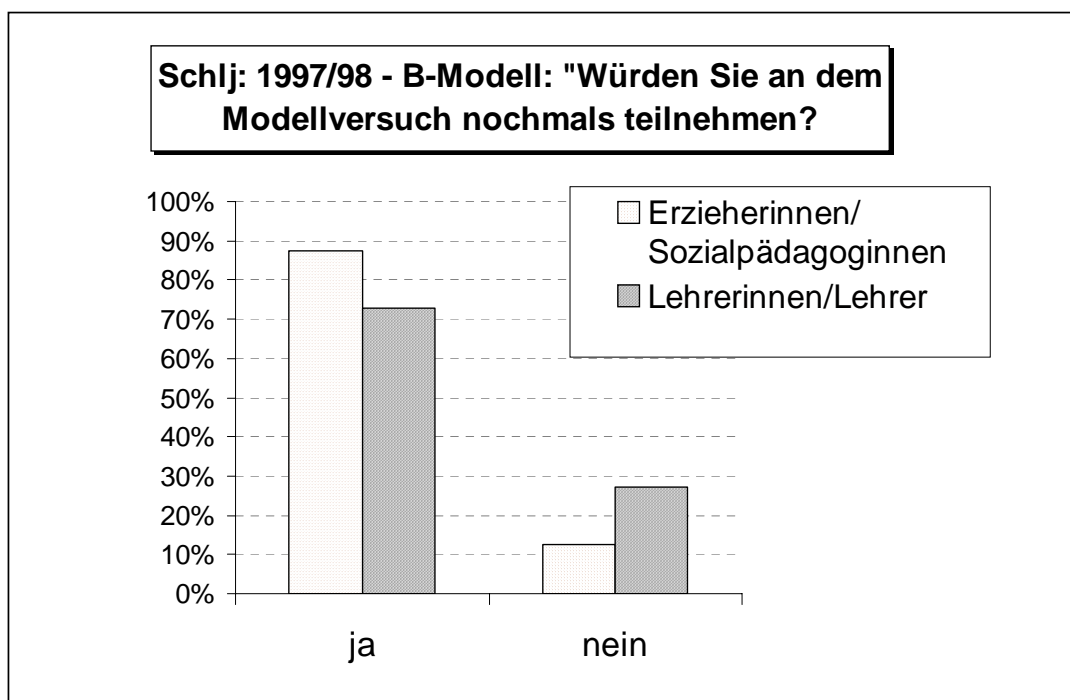
Die in beiden Erhebungszeiträumen am häufigsten genannten Nachteile lassen sich bündeln in einen hohen Zeitaufwand für Teamarbeit und Abstimmungsmaßnahmen/Kooperation mit den Erstklasslehrerinnen und daraus folgend dem Wunsch nach mehr Stunden(nachlass) für solche Koordinierungstätigkeiten.

### Erneute Teilnahme am Versuch

In 1997/1998 wurden die Erzieherinnen/Sozialpädagoginnen gefragt, ob sie noch einmal am Modellversuch teilnehmen würden. 86,7% der Befragten antworteten mit ja; als Gründe für ihr positives Votum führten sie an (nach Häufigkeiten geordnet):

- Neue Ideen, mehr Informationen, mehr Erfahrungsaustausch, mehr Reflexion
- Kindgerechtere Schule (bessere Fördermöglichkeiten, Entstehen von Freundschaften)
- Gute Resonanz bei Eltern und Kindern.

**Abbildung 27: Erneute Teilnahme am Modellversuch bei Erzieherinnen und Lehrerinnen**



## **Förderung von Kindern in B2-Modellen**

Hinweise auf Kinder, die zu Beginn des letzten Kindergartenhalbjahres in eine Förderklasse aufgenommen werden können, kommen von der Erzieherin der Kindertageseinrichtung, der Kooperationslehrerin, aus Befunden der schulärztlichen Untersuchung, von Eltern. Manche Eltern beabsichtigen bereits zu diesem frühen Zeitpunkt ihr Kind vom Schulbesuch zurückzustellen, andere Eltern dagegen wollen ihr Kind im kommenden Herbst einschulen (sog. Korridor-Kind).

Kriterien für die Aufnahme eines Kindes in die Grundschulförderklasse im B2-Modell sind u. a. ein genereller oder ein spezifischer Entwicklungsrückstand, etwa im kognitiven oder sozialen oder motorischen Bereich.

Die Angebote der Erzieherinnen/Sozialpädagoginnen zielen auf die Förderung von personalen, motivationalen, kognitiven und sozialen Basiskompetenzen mittels der traditionellen inhaltlichen Schwerpunkte von Tageseinrichtungen für Kinder bzw. von Grundschulförderklassen.

Die Kooperation zwischen den Pädagogen der Einrichtungen setzt in Abhängigkeit von Förderzeitraum und Förderort Schwerpunkte. In der ersten Halbzeit der Förderung, d.h. im Halbjahr vor der Einschulung, arbeitet die Erzieherin/Sozialpädagogin der GFK mit den Pädagogen aus Kindergarten und Grundschule gleichermaßen zusammen. Fallbesprechungen, Entwicklung von Diagnose- und Förderplänen, Hospitation, gemeinsame Elterngespräche werden hierbei als regelmäßige Kooperationsformen berichtet.

Vom Zeitpunkt der Einschulung bis zum Ende des 1. Schulhalbjahres arbeiten die beiden für die erste(n) Klasse(n) verantwortlichen Pädagogen intensiv zusammen. Auch in diesem Abschnitt heben die Erzieherinnen/Sozialpädagoginnen die oben angeführten regelmäßigen Kooperationsformen besonders hervor.

Die Mitarbeit der Erzieherinnen/Sozialpädagoginnen im ersten Schulhalbjahr der 1. Klasse richtet sich auf die Kinder der ehemaligen Grundschulförderklasse und auf Kinder mit besonderem Förderbedarf. Sie erfolgt auf der Grundlage von individuellen Förderplänen, durch Trainingsprogramme, in Form von Spiel- und Lernangeboten für Kleingruppen mit Niveaudifferenzierung. Inwieweit Erzieherinnen/Sozialpädagoginnen diesen Kindern Angebote auch in den Lernbereichen (Fächern der Grundschule) unterbreiten, kann aufgrund der Angaben der Befragten nicht schlüssig beantwortet werden; laut ihren Aussagen kommen Angebote in Deutsch, Mathematik, Musik, Bildende Kunst/Textilem Werken in Betracht.

Zum Befragungszeitpunkt Mai 2000 prognostizierten die Erzieherinnen/Sozialpädagoginnen, dass voraussichtlich nicht alle Schüler aus der GFK im kommenden Schuljahr in die 1. Klasse der Regelschule eingeschult werden.

Die Vorteile der Förderung nach dem B2-Ansatz sehen die Erzieherinnen/Sozialpädagoginnen darin, dass (tendenziell) in diesem Modell eine kontinuierliche Beobachtung von Kindern möglich ist; unmittelbar anschließend können Gespräche über den Förderbedarf dieser Kinder geführt und die Maßnahmen eingeleitet werden,

die Teamarbeit der beteiligten Pädagogen mehr Ideen, mehr methodische Abwechslung und Reflexion in den Unterrichtsvormittag bringt,

die der Einschulung vorausgehende Förderung in der Regel eine Zurückstellung verhindert, d.h. diese Kinder erleben einen positiven und erfolgreichen Start in der Schule.

Die Nachteile dieses Förderkonzepts liegen nach Meinung der Befragten darin, dass (tendenziell) für die Planung und die Vorbereitung der Teamarbeit und anderer Abstimmungsmaßnahmen ein großer Zeitaufwand erforderlich ist, die Mitarbeit in mehreren ersten Klassen bzw. die Kooperation mit mehreren Klassenlehrerinnen von 1. Klassen einen erheblichen organisatorischen Aufwand verlangt, für Kinder mit Entwicklungsrückständen in mehreren Persönlichkeitsbereichen die Verweildauer in der GFK nicht ausreichend ist.

Dieser Förderansatz B2 hat den Vorteil, dass er bereits in der Vor-Schul-Zeit Fähigkeiten fördern und Erfahrungsfelder öffnen kann, um den Übergang vom Kindergarten in die Grundschule zu erleichtern und vor allem Kindern mit Entwicklungsrückständen eine elementare Grundbildung und Förderung zukommen zu lassen. Er erspart manchem Kind ein ganzes Jahr, das es sonst in der Grundschulförderklasse hätte verbringen müssen. Die Verschränkung von sozialpädagogischer Kompetenz und schulpädagogischer Kompetenz sowie die Integration der Sozialpädagogin in das Unterrichtsgeschehen der 1. Klassen kann, sollte dieser Ansatz Verbreitung finden, neue Perspektiven eröffnen.

## Diskussion B1-Modell

Angesichts der eingangs formulierten Erwartungen an diesen Förderansatz ist zunächst die Frage nach dem Erfolg dieses Modells zu beantworten. Der Erfolg kann im Blick auf die berichtete intensive Kooperation des Personals beider Einrichtungen, also zwischen Erzieherinnen/Sozialpädagoginnen und Lehrerinnen/Lehrern beschrieben werden, er kann an der Qualität der erhobenen didaktischen und methodischen Konzepte festgemacht werden und schließlich an der Zahl der Kinder, die in die Regelschule überwechselten.

Insgesamt bestand eine sehr zeitintensive und im Blick auf die Kooperationsformen äußerst vielfältige Zusammenarbeit zwischen Erzieherinnen/Sozialpädagoginnen und Lehrerinnen/Lehrern. Die Förderung der Kinder der „spezifischen Lerngruppe“ stellte allerdings hohe Anforderungen an die soziale sowie an die didaktische und methodische Kompetenz der beteiligten Pädagogen. Die neuen Arbeitsbedingungen, vor allem die zu entwickelnden Maßnahmen der Differenzierung und Individualisierung sowie die Teamarbeit wurden von den teilnehmenden Erzieherinnen/Sozialpädagogen und Lehrerinnen/Lehrern in der Rückschau als Erweiterung der eigenen Professionalität bewertet. Der von ihnen als nachteilig angeführte hohe Zeitaufwand für die Vorbereitung der Teamarbeit, die Kooperation mit den Pädagogen der anderen Einrichtung und für andere Abstimmungsmaßnahmen kann als Indiz dafür gewertet werden, dass diese pädagogischen Aufgabenbeschreibungen der sLG in Zukunft in der Unterrichtsverpflichtung berücksichtigt werden müssen.

Betrachtet man die berichteten Übergangsquoten aus der Grundschulförderklasse in die Regelschule unter Ausschluss der „spezifischen Lerngruppe“, so kann das Förderangebot zunächst als erfolgreich bewertet werden: In 1999/2000 wurden in 10 von 29 Fällen alle Kinder einer Grundschulförderklasse in die erste Klasse der Grundschule übernommen. In 19 Einrichtungen war dies nicht der Fall (siehe dazu Anmerkung zum Erhebungszeitpunkt Seite 111/weiter oben). Jedoch kann angenommen werden, dass auch in den zuletzt genannten Einrichtungen bei einer Anzahl von Kindern die kognitiven, sozialen und sonstigen Voraussetzungen in dem einen Jahr verbessert wurden, so dass diese ebenso die Grundschule besuchen dürften.

Werden folglich die günstigen Startvoraussetzungen als Erfolgskriterien herangezogen, gilt es zu bedenken, dass diese Kinder schließlich auch ein Jahr älter geworden sind. Fragwürdig erscheint jedoch der Zeitpunkt, an dem der Erfolg des Modells festgemacht werden soll: am Schuleintritt oder am Schulerfolg dieser Kinder nach zwei Jahren oder am Ende der Grundschulzeit?

Offen bleibt die Frage, ob dieser Erfolg nicht auch bei sofortiger Einschulung dieser Kinder in eine jahrgangsgemischte Gruppe und einer ggf. längeren Verweildauer bis zu drei Jahren hätte erreicht werden können.

Würde zur Bewertung des Ansatzes, Kinder in „spezifischen Lerngruppen“ zusammenzufassen und zu fördern, einzig die Anzahl der Übertritte aus der GFK in die Regelschule herangezogen, müsste dessen Scheitern konstatiert werden: nur ein GFK-Kind trat während des Schuljahres aus einer „spezifischen Lerngruppe“ in die erste Klasse der Grundschule über.

Im Rahmen dieses Förderansatzes sind aus systemischen Gründen kaum zeitlich flexible Übergänge von zurückgestellten Kindern in die Regelschule zu erwarten, vor allem dann, wenn die aufnehmenden Grundschulen als B-Modell oder als Vergleichsschule in Form von Jahrgangsklassen geführt werden. Dies bedingt der andersartige Bildungsauftrag der Grundschule mit den stetig fortschreitenden Lehrgängen zum Schriftspracherwerb und in Mathematik. Je weiter die Kinder der ersten Klasse hier im Wochenrhythmus fortschreiten und je später im laufenden Schuljahr der Übertritt vollzogen wird, desto kleiner dürften die Chancen von Kindern aus der GFK werden, in diesen Lernbereichen die Anschlüsse zu finden. Aus Sicht der Erzieherinnen/Sozialpädagoginnen wird das Fortschreiten der Kinder in der sLG in zwei Geschwindigkeiten wahrgenommen; in beiden Einschulungsjahrgängen führen sie als Nachteil des B1-Modells an, dass Erstklässler schnellere Fortschritte machen als die Kinder der Grundschulförderklasse, dadurch sei es schwierig, lernhomogene Gruppen zu bilden.

Aus Sicht des Arbeitskreises wäre es ratsam zu klären, welche längerfristigen Auswirkungen verschiedene Einschulungsansätze auf den Lern- und Entwicklungsverlauf der Kinder bis zum Ende der Grundschulzeit haben. Der Erfolg dieses Modells kann nicht am Zeitpunkt des Schuleintritts vordem zurückgestellter Kinder bzw. der dabei erzielten Übertrittsquote festgemacht werden<sup>54</sup>. Aufschluss über die Qualität des B1-Modellansatzes könnten eher die Analysen der Lern- und Entwicklungsverläufe von Kindern über einen ganzen Grundschuljahrgang geben, hier von zurückgestellten im Vergleich zu denen, die zurückgestellt worden wären, jedoch ein A-Modell besuchten und dort gefördert wurden. Als Bewertungsgrundlagen könnten dabei die schulische Leistungsentwicklung in Lesen, Rechtschreiben und Mathematik, die Entwicklung kognitiver Kompetenzen und deren Stützfaktoren sowie des Sozialverhaltens neben der Analyse der individuellen Schullaufbahnen dienen.

---

<sup>54</sup> Aus Gründen des Untersuchungsdesigns hätte es zur Klärung dieser Fragen einer zusätzlichen Kohorte (Kohorte 4) bedürft.

## II.10 Die Elternmeinung zu den Modellen

### Einführung

Für die Beurteilung einer Übertragbarkeit von Modellen ist es wichtig, auch die Elternmeinung zu den jeweiligen Modellen zu erfragen. Modelle, die auf erhöhten Widerstand der Eltern stoßen, sind natürlich kaum in die Fläche auszuweiten, selbst wenn ihr pädagogischer Ertrag mit dem anderer Modelle vergleichbar sein sollte.

Festzuhalten ist hier allerdings, dass die Elternmeinung, wenn sie mit einem Fragebogen gemessen wird, weniger mit einer unmittelbar aus der Beobachtung von Unterricht resultierenden Schlussfolgerung zu tun hat, sondern mit einem eher diffus kognitiven, allerdings erheblich im Emotionalen verankerten „Wissen“, dessen Quellen kaum zu ergründen sind und deren Ausgestaltung von der Attitüden- bzw. der Vorurteilsforschung theoretisch zu erfassen gesucht wird.

### Methodische Vorbemerkung

Dieses Kapitel hat die Auswertung des „Fragebogens zur Schulzufriedenheit der Eltern“ zum Inhalt. Der Fragebogen wurde im Juli (in Ausnahmefällen im September) des Jahres 2001 bei den Eltern der sogenannten Kohorte 3 (regulärer Einschulungstermin: September 1999) eingesetzt.

Um einen vergleichbaren „Erfahrungszeitraum“ zu gewährleisten, wurden nur solche Eltern befragt, deren Kind im September 1999 eingeschult worden war und das sich zum Zeitpunkt der Erhebung am Ende der Klasse 2 (vor den Sommerferien) bzw. am Anfang der Klasse 3 (nach den Sommerferien) befand.

Für die Auswertung lagen insgesamt 705<sup>55</sup> Fragebogen vor, die sich wie folgt auf die Untersuchungsgruppen verteilten:

- A1-Modell = 246 Fragebogen
- A2-Modell = 109 Fragebogen
- B-Modell = 139 Fragebogen
- Vergleichsschulen = 211 Fragebogen

Die Unterschiede zwischen den Modellen wurden je nach Skalenniveau der Antworten mit dem Chi<sup>2</sup>-Test und der einfaktoriellen Varianzanalyse mit anschließendem LSD bzw. dem t-Test für unabhängige Stichproben auf statistische Signifikanz überprüft ( $\alpha \leq .05$ ). Das Ergebnis der Signifikanzprüfung ist jeweils am Fuß der Tabellen festgehalten.

Die Antworten zu einigen Einzelfragen zur Förderung der Schüler werden anhand eines aus diesen generierten Skalenwertes (Modellindex) zum Vergleich der Modelle herangezogen.

### Ergebnisse

#### Fragen zum Schulanfang

- Etwa die Hälfte der Eltern hat zum ersten Mal die Einschulung eines ihrer Kinder erlebt.

---

<sup>55</sup> Es handelt sich bei einer Gesamtschülerzahl der Kohorte 3 von 1040 Schülern um eine Rücklaufquote von 67,8 %, für Elternbefragungen ein durchaus zufriedenstellender Wert.



- Bei den Fragen nach der Einstellung der Kinder zur Schule zeigten sich keine signifikanten Unterschiede zwischen den Eltern der einzelnen Modelle.
- Danach haben dem Schulbeginn 86,9% der Kinder freudige und 6,3% ängstliche Erwartungen entgegengebracht. 6,8% der Kinder standen dem Schulanfang eher gleichgültig gegenüber.
- Soweit bei den Kindern Ängste vorherrschten, wurden diese im Laufe der ersten beiden Schuljahre weitgehend abgebaut (74,4%), allerdings nahm bei einem Teil der Kinder auch die anfängliche Freude ab (27,1%).

### **Fragen zur Förderung der grundlegenden Kulturtechniken und des musisch-sportlichen Bereichs durch die Grundschule**

- Hinsichtlich der klassischen Kulturtechniken Lesen, Schreiben, Rechnen ergaben sich einige, wenn auch unsystematische signifikante Unterschiede bei der Meinung der Eltern zum jeweiligen Modell:
- Beim Lesen lernen trauten die Eltern der A2-Schulen ihrem Modell weniger Förderung zu als die Eltern der A1- und B-Schulen.
- Beim Schreiben lernen gab es keine Unterschiede.
- Beim Rechnen billigten die Eltern der A1-Schulen ihrem Modell die höchste Förderwirkung zu, die Eltern der A2-Modelle die niedrigste (nur die Unterschiede der Eltern dieser Modelle sind signifikant).
- Hinsichtlich sportlicher und musischer Aktivitäten gab es wieder keine Unterschiede in der Meinung der Eltern.
- Bei der Meinung über die musikalische Förderung waren die Eltern der A1-Schulen signifikant positiver eingestellt als die der Vergleichsschulen.
- Insgesamt ergibt sich also ein Bild, bei dem die Eltern der A1-Schulen die am meisten positive Meinung von der Förderwirkung ihres Modells hatten, die größte Skepsis zeigten diesbezüglich die Eltern der A2-Schulen.

### **Die Meinung der Eltern zur Förderung bestimmter Fähigkeiten, die in die Richtung von Schlüsselqualifikationen deuten**

- Hinsichtlich der Kompetenz zum Lernen und zur Kreativität, d.h.  
Fähigkeit zum selbständigen Lernen  
Fähigkeit, eigene Ideen zu entwickeln  
erteilen die Eltern des A1-Modells die höchsten Werte.
- Hinsichtlich der sozialen Kompetenz, d.h.  
Fähigkeit, mit anderen Kindern möglichst störungsfrei zusammenzuleben  
Fähigkeit, mit anderen Kindern zu spielen, zu arbeiten und zu lernen  
Fähigkeit, sich auch gegen die Gruppe zu behaupten  
Fähigkeit, anderen Kindern in der Schule zu helfen  
urteilen die Eltern des A1-Modells bei allen vier Items am meisten positiv, die der Vergleichsschulen am wenigsten positiv.

### **Zur Förderung der Kinder durch die Schule**

Von besonderem Interesse sind die beiden Fragenbereiche, von denen schon oben bei der Vorstellung der Detailergebnisse zu den Einzelfragen berichtet wurde, also die nach der Förderung der Kinder durch die Schule bzw. das jeweilige Modell:

- Fragenbereich: Welche Förderung erfuhr Ihr Kind in den genannten Lernbereichen?
- Fragenbereich: Wie wurden die genannten Fähigkeiten bei Ihrem Kind gefördert?

Fasst man die in diesem Zusammenhang aufgelisteten Einzelbereiche zu einer Skala (Modellindex) zusammen, so erhält man ein relativ messscharfes Instrument.<sup>56</sup>

Die Messskala beinhaltet drei Aspekte:

1. Klassische Kulturtechniken (Lesen, Schreiben, Rechnen)
2. Sport und Musik
3. Soziale Kompetenz

Der so definierte Modellindex (die gemittelte Einstufung der Eltern eines bestimmten Modells von allen 11 Einzelfragen) kann zwischen 11 und 33 liegen. Je höher der Wert ist, umso besser schneidet das Modell im Urteil der Eltern ab.

Wie die folgende Tabelle und die Graphik auf der nächsten Seite zeigt, nimmt der Mittelwert (Modellindex) in der Reihenfolge

A1-Modell, A2-Modell, B-Modell und Vergleichsschulen ab, während die Standardabweichung<sup>57</sup> zunimmt.

**Tabelle 29: Modellindizes**

	<b>A1-Modell</b>	<b>A2-Modell</b>	<b>B-Modell</b>	<b>Vergleichsschulen</b>
Anzahl der Fälle	220	96	117	197
Mittelwert (Modellindex)	<b>26,91</b>	<b>26,52</b>	<b>25,49</b>	<b>25,33</b>
Standardabweichung	3,91	4,31	4,46	4,51

**Signifikante Unterschiede zwischen den Modellen**

LSD (Mittelwertvergleiche):

- A1 > B
- A1 > VS
- A2 > VS

*Ergebniszusammenfassung zur „Förderung der Kinder durch die Schule“*

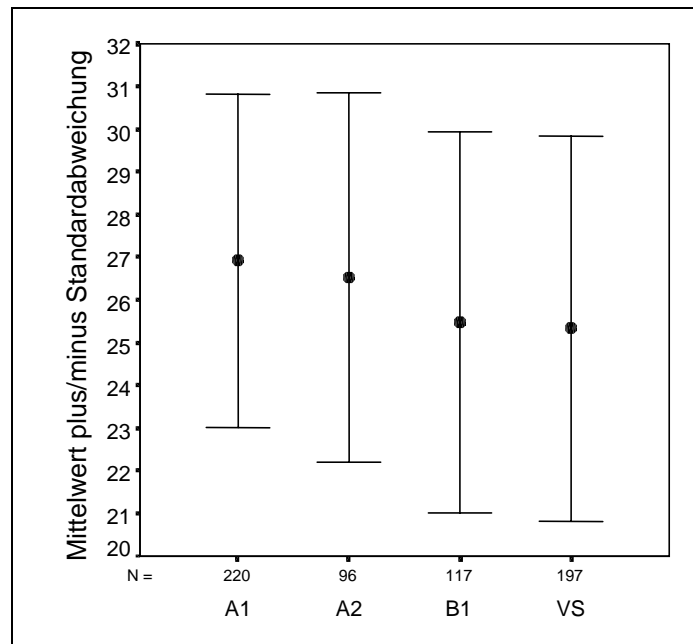
- Das Urteil der Eltern aller Modelle fällt insgesamt positiv aus, wobei die Eltern der A-Modelle allerdings der Schule ihres Kindes (bzw. dem dahinterstehenden Modell) ein noch besseres Urteil ausstellen – und sich darin auch einiger sind – als die Eltern des B-Modells und der Vergleichsschulen.
- Schaut man sich jetzt die weiter oben angeführten Ergebnisse zu den Einzelbereichen unter Anlegung der aufgezeigten Faktorenstruktur noch einmal an, so gilt dies insbesondere für die Förderung der sozialen Kompetenz. Das heißt, die betroffenen Eltern

<sup>56</sup> Reliabilität (Cronbachs Alpha) = .84 (N = 630)

<sup>57</sup> Hier: Maß für die Homogenität (kleiner Wert) bzw. Heterogenität (großer Wert) der Einstufungen. Die Werte sind nicht absolut, sondern relativ zueinander zu lesen.

ordnen der in den A-Modellen praktizierten Jahrgangsmischung einen diesbezüglich positiven Effekt zu.

- Hinsichtlich der klassischen Lernbereiche unterscheiden sich allerdings die Eltern der beiden A-Modelle. Hier fällt das Urteil der A2-Eltern signifikant schlechter aus als das der A1-Eltern.



**Abbildung 28: Der Mittelwert der einzelnen Modelle bezogen auf den Modellindex**

### **Gibt es modellbedingte Unterschiede des Gesprächsbedarfs von Eltern mit der Lehrerin/dem Lehrer?**

Es gibt gewisse (signifikante) Unterschiede, was die Gesprächskontakte der Eltern und Lehrer betrifft:

- Die Eltern der A-Modelle wenden sich häufiger an den Lehrer/die Lehrerin Ihrer Kinder als die der B-Modelle und der Vergleichsschulen.
- Allerdings geben die Eltern der A2-Modelle den höchsten Gesprächsbedarf an, auch signifikant höher als die der A1-Modelle; die Eltern der B-Modellschulen und der Vergleichsschulen unterscheiden sich nicht signifikant, was ihren Gesprächsbedarf mit den Lehrern ihrer Schule betrifft.

### **Modellspezifische Fragen**

#### Fragen zum A1- und A2-Modell

(Bewertung der Jahrgangsmischung, Zustimmungstendenz zum Modell am Kriterium „nochmalige Einschulung“ und Weiterführung des Modellansatzes auch in den Kl. 3 und 4)

- Zur Jahrgangsmischung zeigten sich keine signifikanten Unterschiede in den Antworten der A1- und A2-Eltern. Die überwiegende Mehrheit der Eltern gibt an, dass ihr

Kind damit zufrieden ist, im Unterricht mit älteren und jüngeren Kindern zusammenzuarbeiten. Nur 11,5% sind danach eher unzufrieden.

- 58,6% der Eltern würden sich vorbehaltlos nochmals für die Einschulung ihres Kindes in eine jahrgangsgemischte Klasse entscheiden.
- 44,3% der Eltern sind der Meinung, dass der Modellansatz auch in Klasse 3 fortgeführt werden sollte, in Bezug auf Klasse 4 sinkt die Zustimmungsrage auf 36,9%.

#### Eine Frage speziell zum A2-Modell

(Bewertung des zweiten Einschulungstermins)

Der zweite Einschulungstermin wird von fast der Hälfte der A2-Eltern als problematisch angesehen, nur 53,6% der Eltern bewerten ihn als eine begrüßenswerte Chance für viele Kinder.

#### Eine Frage speziell zum B-Modell

(Modellbewertung)

Die Eltern des „klassischen“ B-Modells (einzelne Projekte einer gemeinsamen Gruppe von Schülern der ersten Grundschulklasse und der Grundschulförderklasse) halten sich ziemlich bedeckt, was die Zustimmungstendenz zum Modell betrifft. Diese Zurückhaltung kann möglicherweise auch darauf zurückzuführen sein, dass die Mehrzahl der Grundschüler nicht oder nur selten in einer solchen gemeinsamen Gruppe war.

### **Zusammenfassung**

*Bei fast allen Fragen schneiden die A-Modelle und hier besonders die A1-Modelle in den Antworten der Eltern gut ab, was die Förderung der Kinder im Erwerb von Fertigkeiten (Kulturtechniken und Fertigkeiten im Bereich Sport und Musik betrifft); noch besser, was die Einübung von Fähigkeiten des selbstständigen Lernens, der Förderung kreativer Tendenzen und der Sozialkompetenz betrifft. Auch hier schneidet das A1-Modell noch eine Spur besser ab. Ein sog. Modellindex, der alle entsprechenden Urteile zusammenfasst, zeigt, bezogen auf den durchschnittlichen „Förderwert“, diese Reihenfolge: 1. A1-, 2. A2-, 3. B-Modell und 4. Vergleichsschulen.*

*Eine Jahrgangsmischung, die die Klassenstufen 1 und 2 zusammenfasst, würde eine gewisse Mehrheit der Eltern vorbehaltlos nochmals wählen, wenn eine erneute Einschulung anstünde. Allerdings sinkt die Zustimmung zur Fortführung des Modellansatzes in der dritten Klasse ein wenig unter die 50%-Marke; und für die Fortführung in Klassenstufe 4 unter die 40%-Marke. Der zweite Einschulungstermin wird von 54% der A2-Eltern als „begrüßenswerte Chance“ gesehen.*

*Die Zustimmung zum „klassischen“ B-Modell (B1-Modell) ist eher von Zurückhaltung geprägt.*

*Insgesamt scheinen die Antworten der Eltern – zumindest teilweise – das widerzuspiegeln, was ihnen von den Schulen als Vorteil ihres jeweiligen Modells vermittelt worden ist. Offensichtlich ist die Aufklärungsarbeit der A-Schulen hier besonders fruchtbar gewesen<sup>58</sup>.*

---

<sup>58</sup> So ist die Diskrepanz zwischen den Ergebnissen zur Leistungsförderung im Bereich des kognitiven Entwicklungsstands, der Lernfreude, des Sozialverhaltens und der schulischen Leistungsfähigkeit der Grundschüler (siehe die entsprechenden Kapitel) und der Elternmeinung wohl am wenigsten dissonant zu erklären.

### **III. Bilanz der Ergebnisse über die verschiedenen Modelle in der Eingangsstufe der Grundschule und Konsequenzen für die Bewertung der Ergebnisse**

#### **III.1 Eingangsvoraussetzungen der Schüler**

##### Kognitive Entwicklung

###### Phonologische Bewusstheit und Aufmerksamkeit

*Zusammengefasst ergeben sich folgende Befunde: Zu Beginn von Kl.1 bringen die Schüler der A-Modelle eher einen niedrigeren Entwicklungsstand mit als die der B-Modelle und der V-Schulen (dies zeigt das Bielefelder-Screening). Zu Beginn des zweiten Schuljahres, etwa im Herbst 2000, zeigten sich dann bei kovarianzanalytischen Berechnungen in zwei von vier Untertests der phonologischen Bewusstheit signifikante Unterschiede zwischen den A-Modellen und den Vergleichsschulen, die zu Lasten der A-Modelle ausfallen. Dabei muss allerdings der erheblich größere Ausländeranteil in den A-Modellen berücksichtigt werden.*

###### (Phonologisches) Arbeitsgedächtnis

*Zusammengefasst gilt daher auch für die Entwicklung des Arbeitsgedächtnisses, dass die Modellvariante, also auch die jahrgangsübergreifende Beschulung im A-Modell des „Schulanfangs auf neuen Wegen“, keinerlei systematischen Einfluss auf die Entwicklung im Verlauf des ersten Schuljahres hat.*

###### (Induktives) Denken

*Für die zwei untersuchten Bereiche des Denkens (Zahlabstraktion, Zählen von Objekten) fielen die kovarianzanalytisch errechneten Befunde wie folgt aus: Zum Zeitpunkt der Einschulung zeigten sich bei den Schülern des B-Modells höhere Leistungen im Untertest Zahlabstraktion. Ansonsten waren bei allen Modellen deutliche, voneinander nicht unterscheidbare Entwicklungszuwächse bis etwa Mitte des ersten Schulhalbjahrs von Kl. 2 zu verzeichnen.*

##### Motivationale Entwicklung

*Insgesamt lässt sich zur motivationalen Entwicklung festhalten, dass die Kinder das hohe durchschnittliche Motivationsniveau am Schulbeginn auch noch ein Jahr später zeigen; allerdings nicht ganz unabhängig vom Modell, das sie besuchen. Die Schüler der A-Modelle und der Vergleichsschulen zeigten etwa gleiche Werte, die der B-Modelle gaben im ersten Schulhalbjahr von Klasse 2 eine signifikant höhere Lernfreude an.*

## **III.2 Soziales Verhalten und soziale Einstellungen**

Soziales Verhalten und soziale Einstellungen in der Grundschuleingangsstufe (Kl. 1 und 2)

*Insgesamt lässt sich zur Entwicklung des Sozialverhaltens festhalten, dass die Kinder das hohe durchschnittliche Einstellungsniveau zum Schulbeginn auch noch ein Jahr später zeigen; allerdings nicht ganz unabhängig vom Modell, das sie besuchen, wobei es in den Ergebnissen auch Unterschiede zwischen den beiden Kohorten (Schülerjahrgänge 1 und 3) gibt. Diese Befunde weisen auf eine eher modellunabhängige als auf eine modellabhängige Förderung des Sozialverhaltens hin.*

Soziales Verhalten und soziale Einstellungen am Ende von Klasse 4

*Bei den Querschnittsvergleichen, unter Einbeziehung der fünf SVG-4.1-Skalen und des Testgesamtwerts, ergeben sich sowohl bei den generellen Gruppenvergleichen (Mehrfachvergleichen) als auch bei den Einzelvergleichen bei Anwendung der Post-hoc-Tests nur wenige signifikante Modellunterschiede.*

*Bei den Längsschnittvergleichen, die nur mit dem Testgesamtwert durchgeführt wurden, sind sowohl alle Mehrfachvergleiche als auch alle Einzelvergleiche nicht signifikant.*

*Man muss daher von einer im hohen Maße modellunabhängigen Förderung des Sozialverhaltens bzw. der Entwicklung der Einstellung zur Schule während der Grundschulzeit ausgehen.*

*Dennoch sollte nicht unerwähnt bleiben, dass die Lehrerinnen und Lehrer der A-Modelle mit einer etwas schwierigeren Ausgangslage konfrontiert wurden: In ihren Eingangsklassen befanden sich deutlich mehr Schülerinnen und Schüler, die in den B-Modellen und in den Vergleichsschulen vom Schulbesuch zurückgestellt worden wären; statistisch gesehen hatten ihre Eingangsklassen auch einen signifikant höheren Ausländeranteil und einen deutlich höheren Anteil an vorzeitig eingeschulter Kindern.*

## **III.3 Schulische Leistungen Kl. 1 bis Kl. 4**

Schulische Leistungen gemessen mit standardisierten Verfahren

### Lesen und Rechtschreiben

*In diesem Bereich sind kaum Unterschiede zwischen den drei Gruppen auszumachen. Eine leichte Überlegenheit im Bereich des Lesetempos ist für die Kinder der A-Modelle und B-Modelle am Ende der Grundschulzeit feststellbar. Weitergehende Analysen zeigen, dass die schriftsprachlichen Schuleingangsvoraussetzungen für die Schülerschaft in den A-Modellen die ungünstigsten Werte aufweisen. Angesichts dessen ist es umso bemerkenswerter, dass für diese Kinder am Ende der vierten Klasse bessere Leseleistungen als in den Vergleichsschulen resultieren.*

## Rechnen (Mathematik)

*Zwar zeigen die Schüler des A-Modells numerisch eine signifikant niedrigere Durchschnittsleistung im Rechnen gegenüber den B-Modellen, nicht aber gegenüber den Vergleichsschulen; weitergehende Analysen zeigen, dass die Ergebnisse teilweise auf die unterschiedliche Zusammensetzung der Schülerschaft in den Modellen zurückzuführen ist. Dennoch zeigt sich ein Mathematikvorteil der B-Modelle am Ende der Grundschulzeit. Am Ende von Klasse 4 erreichen alle drei untersuchten Gruppen (A-Modelle, B-Modelle, Vergleichsschulen) im Bundesvergleich überdurchschnittlich gute Leistungen.*

Schulische Leistungen gemessen mit landesweiten Vergleichsarbeiten Fächern Deutsch (Lesen, Rechtschreibung und Aufsatz) und Mathematik am Ende von Kl. 3

*Vergleicht man die Ergebnisse der landesweiten Diagnosearbeiten mit den schulischen Leistungen, gemessen mit standardisierten Verfahren, so bestätigt sich auch hier der Mathematikvorteil der B-Modelle. Ein Sprachvorteil der A-Modelle kann zwar ebenfalls numerisch bestätigt werden, lässt sich jedoch nicht statistisch absichern.*

Leistungsentwicklung von drei Schülergruppen „potentiell leistungsstarken“, „unauffälligen“ Schülern und von „Risikoschülern“ in Abhängigkeit vom Modell

*Insgesamt zeigen sich kaum differentielle Auswirkungen der drei miteinander verglichenen Gruppen. Lediglich in zwei Fällen ergaben die kovarianzanalytischen Auswertungen Leistungsunterschiede zwischen den Gruppen für die Risikokinder. In beiden Fällen zeigten sich Vorteile der A-Modelle. Im Bereich der mit den DEMAT-Verfahren erfassten Mathematikleistungen erwiesen sich die A-Modelle gegenüber den Vergleichsschulen für die Risikokinder als günstiger. Im Bereich des Leseverständnisses am Ende der Grundschulzeit zeigte sich für die Risikokinder ein tendenzieller Vorteil der A-Modelle im Vergleich zu den B-Modellen.*

### **III.4 Noten in den Fächern Deutsch und Mathematik (Kl. 2 bis Kl.4)**

*Trotz numerischer Unterschiede der Noten bewegen sich dennoch alle modellspezifischen Notendurchschnitte derselben Klassenstufe im selben Bereich, d.h. es gibt keine signifikanten Unterschiede zwischen den Modellen, was die Durchschnittsnoten in Deutsch und Mathematik am Ende der Kl. 2, Kl.3 und Kl. 4 betrifft.*

### **III.5 Das Durchschnittsalter der Schüler am Ende von Kl. 4**

*Die Schüler der A-Modelle sind am Ende der Kl. 4 signifikant jünger (um ca. 3 Monate) als die der B- Modelle und der Vergleichsschulen.*

### **III.6 Bildungsempfehlung und Modell**

*Der signifikante Mittelwertsunterschied zwischen den vier Modellen (A1-, A2-, B- Modelle und Vergleichsschulen) ist auf die unterschiedliche Struktur der Bildungsempfehlung der Vergleichsschulen zurückzuführen. Diese sprechen erkennbar mehr Hauptschulempfehlungen und weniger Gymnasialempfehlungen aus als die Schulen der drei Schuleingangsmodelle.*

### **III.7 Unterrichtsgestaltung**

In der Schuleingangsstufe (Klassenstufe 1 und 2)

*In den Vergleichsklassen ohne Jahrgangsmischung wird signifikant häufiger frontal unterrichtet als in den Erprobungsklassen. Positiv formuliert bedeutet dies, dass trotz fachbedingter Variationen in den Erprobungsklassen eher Unterrichtsformen praktiziert werden, die ein binnendifferenziertes Lernen begünstigen. Dieser statistisch bedeutsame Unterschied in Kohorte 1 bestätigt sich auch in Kl. 2 der Kohorte 3. Demnach kann auf der Grundlage der bisherigen Befragungsergebnisse dem Modellversuch ein innovationsauslösender und -stabilisierender Effekt im Bereich der didaktisch-methodischen Gestaltung des Anfangsunterrichts bescheinigt werden, der bei den A-Modellklassen stärker ausfällt als bei den B-Modellklassen.*

*Die Tatsache, dass die in B-Modellklassen (mit deutlich nachlassender Tendenz in Kohorte 3) und mehr noch die in A-Modellklassen unterrichtenden Lehrkräfte in der überwiegenden Mehrzahl eine erneute Teilnahme an dem Schulversuch bejahen (Kohorte 1 und 3), kann als Hinweis auf die überaus große Akzeptanz der intendierten Neuerungen gewertet werden. Neben den lernbedeutsamen Vorteilen für die Kinder bewerten die Lehrkräfte der Erprobungsklassen für ihre eigene Berufsarbeit die Fortentwicklung ihrer didaktischen Kompetenzen als eine positive Wirkung des Schulversuchs, dem damit ein bedeutsamer Effekt für die Erhöhung der Professionalität zugeschrieben wird.*

*Die über die Selbstauskunft der Lehrerinnen und Lehrer erhobenen Ergebnisse zum Sozialverhalten der Schülerinnen/Schüler zeigen, dass den Schülerinnen/Schüler der A-Modelle am häufigsten ein positives Sozialverhalten, insbesondere wegen der durch die Jahrgangsmischung hervorgerufenen Helfereffekte zuerkannt wird; dies ergibt auch die Elternbefragung. Allerdings lassen sich in der Schülerbefragung die Rangreihen zwischen den Modellen nicht bestätigen. In Bezug auf die Lernentwicklung der Kinder wird von den Lehrerinnen/Lehrern am häufigsten positiv bewertet, dass die Jahrgangsmischung das Spektrum der Lernmöglichkeiten und Lernwege (z.B. durch Mithören, Zuschauen) verbreitert.*

Anschlussuntersuchung in der dritten und vierten Jahrgangsstufe

*Wie die Anschlussuntersuchung in der dritten und vierten Jahrgangsstufe gezeigt hat, nehmen die in der jahrgangsgemischten Eingangsstufe festgestellten Unterschiede in der Unterrichtsgestaltung mit aufsteigender Jahrgangsstufe deutlich ab. Eine Fortsetzung der Jahrgangsmischung über die Eingangsstufe hinaus wird von etwa der Hälfte der Dritt- und ca. 43% der Viertklasslehrerinnen befürwortet.*



### **III.8 Unterrichtsgestaltung Grundschulförderklasse**

*In den „spezifischen Lerngruppen“ werden Kinder überwiegend in Kleingruppen und häufig in Einzelarbeit unterrichtet. Die Angebote werden im ersten Durchgang überwiegend, im zweiten Durchgang häufig von Erzieherinnen / Sozialpädagoginnen erteilt. Eine zeitintensive und im Blick auf die Kooperationsformen sehr vielfältige Zusammenarbeit zwischen Erzieherinnen / Sozialpädagoginnen und Lehrerinnen / Lehrern wirkt sich nach Einschätzung der erstgenannten nicht nur positiv auf den problemlosen Übergang von Kindern in die Grundschule aus, sondern erweitert auch die eigene Professionalität. Fallbesprechungen im Anschluss an kontinuierliche Beobachtung von Kindern, im Team entwickelte Förderpläne, gemeinsame Vorbereitung der Angebote sowie ein Austausch von Materialien bringen mehr Ideen, methodische Abwechslung und Reflexion in den Unterrichtsvormittag ein.*

### **III.9 Die Elternmeinung zu den Modellen**

*Bei fast allen Fragen schneiden die A-Modelle und hier besonders die A1-Modelle in den Antworten der Eltern gut ab, was die Förderung der Kinder im Erwerb von Fertigkeiten (Kulturtechniken und Fertigkeiten im Bereich Sport und Musik betrifft); noch besser, was die Einübung von Fähigkeiten des selbstständigen Lernens, der Förderung kreativer Tendenzen und der Sozialkompetenz betrifft. Auch hier schneidet das A1-Modell noch eine Spur besser ab. Ein sog. Modellindex, der alle entsprechenden Urteile zusammenfasst, zeigt, bezogen auf den durchschnittlichen „Förderwert“, diese Reihenfolge: 1. A1-, 2. A2-, 3. B-Modell und 4. Vergleichsschulen.*

*Eine Jahrgangsmischung, die die Klassenstufen 1 und 2 zusammenfasst, würde eine gewisse Mehrheit der Eltern vorbehaltlos nochmals wählen, wenn eine erneute Einschulung anstünde. Allerdings sinkt die Zustimmung zur Fortführung des Modellsatzes in der dritten Klasse ein wenig unter die 50%-Marke; und für die Fortführung in Klassenstufe 4 unter die 40%-Marke. Der zweite Einschulungstermin wird von 54% der A2-Eltern als „begrüßenswerte Chance“ gesehen.*

*Die Zustimmung zum „klassischen“ B-Modell (B1-Modell) ist eher von Zurückhaltung geprägt.*

*Insgesamt scheinen die Antworten der Eltern – zumindest teilweise – das widerzuspiegeln, was ihnen von den Schulen als Vorteil ihres jeweiligen Modells vermittelt worden ist. Offensichtlich ist die Aufklärungsarbeit der A-Schulen hier besonders fruchtbar gewesen.*

### **III.10 Generelles Fazit**

Auf den ersten Blick scheint es, dass die Bilanz der modellspezifischen Wirkungen eher ernüchternd ausfällt. Nur in einigen wenigen Fällen lassen sich modellspezifische Effekte nachweisen; d.h. auch, die Modelle unterscheiden sich im Wesentlichen nicht oder nur kaum in der individuellen Förderung der Schüler hinsichtlich ihrer kognitiven, motivationalen, sozialen Fähigkeiten; genauso wenig unterscheiden sie sich wesentlich in der Förderung des schulischen Leistungsvermögens ihrer Schüler.

Aus internationalen Vergleichen ist bekannt, dass ein Problem des deutschen Schulwesens im höheren Durchschnittsalter der Schüler bei vergleichbaren Abschlüssen liegt. Legt man zur Bewertung der Datenbefunde den Maßstab des möglichst niedrigen Altersdurchschnitts der Schüler und einer mindestens gleichwertigen Förderung der Schüler am Ende der Grundschulzeit an, dann gibt das für die Kohorte 3 (Hauptuntersuchung) vorliegende Datenmuster durchaus Anlass zu einer positiven Bilanzierung des jahrgangsübergreifenden Anfangsunterrichts, wie er in den A-Modellen verwirklicht ist:

Die Kinder der A-Modelle waren schon zum Zeitpunkt der Einschulung jünger, am Ende der Grundschulzeit sind sie durchschnittlich 2 bis 3 Monate jünger. Und das, obwohl sie - zumindest in dieser Stichprobe - ungünstigere Startbedingungen als die Kinder der Vergleichsschulen hatten: Sie verfügten zu einem deutlich höheren Prozentsatz über geringere Deutschkenntnisse, und damit wohl verbunden, hatten sie auch schlechtere Ausgangsbedingungen in der phonologischen Bewusstheit und im phonologischen Speicher des Arbeitsgedächtnisses. Dennoch erreichten sie bis zum Ende der zweiten Klasse das gleiche Leistungsniveau beim Lesen, Rechtschreiben und Rechnen und behielten dieses bis zum Ende der Grundschulzeit.

Eine ebenfalls für die schulischen Leistungsentwicklung wichtige Bilanz für die A-Modelle lautet: Insbesondere für Risikokinder, also für Kinder mit ungünstigen kognitiven Entwicklungsvoraussetzungen, ergaben sich gegenüber den B-Modellen Vorteile hinsichtlich des Leseverständnisses (in Klasse 3) sowie gegenüber den Vergleichsschulen Vorteile in der Entwicklung mathematischer Leistungen.

So kann das vom Kultusministerium aufgestellte Ziel zur Bewertung des Erfolgs des jahrgangsübergreifenden Prinzips bei der Bildung der Klassen und beim Unterricht als erreicht gelten. Sowohl die Lehrerinnen/Lehrer als auch die Eltern beurteilen die A-Modelle am günstigsten, vor allem was die Einübung positiver sozialer Kompetenzen betrifft. Von den Lehrerinnen/Lehrern scheint auch der von der Jahrgangsmischung ausgehende Druck zur inneren Differenzierung und mit ihr im Gefolge die deutlicher erlebte Notwendigkeit, eine breitere Palette von Sozialformen des Unterrichts und Unterrichtsmethoden einzusetzen zu müssen, überwiegend positiv erlebt und reflektiert zu werden.

Greift man die durch die PISA-Studien wieder ins Bewusstsein gekommene stärkere Beachtung der Unterrichtsqualität als Bewertungskriterium auf, so wird deutlich, dass die für die A-Modelle typischen Elemente weitaus mehr dem Erkenntnisstand der heutigen Grundschulpädagogik entsprechen als die Elemente des B-Modells und denen der Vergleichsschulen.

Interessanterweise wurde in den Klassenstufen drei und vier, also in den versuchsnachfolgenden Grundschulklassen, wieder eine Unterrichtsgestaltung praktiziert, die sich nur geringfügig von derjenigen in den Vergleichsklassen unterscheidet. So gleicht sich im Gegensatz zu der Schuleingangsstufe der Anteil frontaler Unterrichtsphasen an der wöchentlichen Unterrichtszeit in den Fächern Deutsch, Mathematik und Heimat- und Sachunterricht an. Während dort je nach Unterrichtsfach signifikante Unterschiede zwischen den A-Modellklassen und den Vergleichsklassen zu verzeichnen waren, entfielen diese bei den versuchsnachfolgenden Klassen.

Sowohl Lehrer als auch Eltern der A-Modelle befürworten in der überwiegenden Mehrheit das jahrgangsübergreifende Prinzip für die Schuleingangsstufe, für die Klassenstufen drei und vier geht diese Befürwortung deutlich zurück.

Nicht ganz einfach gestaltet sich die Bewertung des B-Modells. Der Kern dieses Modells besteht darin, dass durch die Einrichtung von „spezifischen Lerngruppen“ eine engere Verzahnung der Grundschulförderklassen mit den ersten Grundschulklassen erreicht werden soll. Unter dem Aspekt der Schulzeitverkürzung muss dieses Modell als gescheitert angesehen werden, weil praktisch kein Kind aus der Grundschulförderklasse gleitend auch während des Schuljahres in die erste Grundschulklasse gewechselt ist, d.h., diese Kinder verbringen nach dem „normalen“ Einschulungstermin erst einmal ein Jahr in der Grundschulförderklasse. Aus diesem B1-Modell hat sich das B2-Modell<sup>59</sup> entwickelt, das eher in der Lage zu sein scheint, den Kindern ein Jahr Verweilzeit in einer Grundschulförderklasse zu ersparen.

Für die Kinder der Grundschulklassen im B-Modell unterscheidet sich der Unterricht mit Ausnahme derer, die im ersten Schuljahr an den spezifischen Lerngruppen teilnehmen, vom Konzeptionellen her gesehen durch nichts von dem der Vergleichsschulen. Und dennoch hat man beim Durchschauen der Ergebnisse den Eindruck, dass sich die B-Modellklassen die gesamte Grundschulzeit eher den „Erprobungsklassen“ zuordnen lassen als den Vergleichsklassen. Bemerkenswert ist das gute Abschneiden dieser Klassen beim Rechnen (Mathematik). Dies kann nur teilweise durch bessere Werte bei der Eingangsuntersuchung im deduktiven Denken erklärt werden.

Die Vergleichsschulen haben am Ende der Grundschulzeit die ältesten Schüler, bei der Förderung der „Problemschüler“ schneiden sie weniger günstig ab, darüber hinaus unterscheiden sie sich deutlich von den „Erprobungsschulen“ bei der Bildungsempfehlung: sie sprechen mehr Hauptschulempfehlungen und weniger Gymnasialempfehlungen aus.

Als abschließendes Fazit soll noch einmal hervorgehoben werden: Die Befürchtungen, dass die Jahrgangsmischung in der Schuleingangsstufe die Lehrerinnen und Lehrer überfordere und damit sich ungünstig auf die Schullaufbahn der Grundschülerinnen und –schüler auswirke, konnten nicht bestätigt werden.

---

<sup>59</sup> Beschreibung siehe I.3

## Literatur

### Zur wissenschaftlichen Begleitung

Arbeitskreis wissenschaftliche Begleitung „Schulanfang auf neuen Wegen“ (2000). Zwischenbericht und erste Ergebnisse. Teil I: Ergebnisbericht, Teil II: Tabellenband. Stuttgart.

Arbeitskreis wissenschaftliche Begleitung „Schulanfang auf neuen Wegen“ (2002). Vorläufiger Abschlussbericht zur Eingangsstufe der Grundschule. Stuttgart.

Birkel, P. (1994). Weingartener Grundwortschatz Rechtschreib-Test für dritte und vierte Klassen (WRT 3+). Göttingen: Hogrefe.

Birkel, P. (1995). Weingartener Grundwortschatz Rechtschreib-Test für 1. und 2. Klassen (WRT 1+). Göttingen. Hogrefe.

Bjorklund, D.F., Gaultney, J.F. & Green, B.L. (1993). „I watch, therefore I can do“: The development of meta-imitation during the preschool years and the advantage of optimism about one's imitative skills. In M.L. Howe & R. Ptasnik (Eds.), *Emerging themes in cognitive development*. Vol. II. Competencies (pp. 79-99). New York: Springer.

Faust-Siehl, G. (2001). Die neue Schuleingangsstufe in den Bundesländern. In: Faust-Siehl, G. & Speck-Hamdan, A. (Hrsg.): *Schulanfang ohne Umwege. Mehr Flexibilität im Bildungswesen*. Frankfurt/M., 194-252.

Fölling-Albers, M. (2003). Grundschulpädagogik, Grundschulforschung und Kindheit. In: Panagiotopoulou, A. & Brügelmann, H. (Hrsg.): *Grundschulpädagogik meets Kindheitsforschung. Zum Wechselverhältnis von schulischem Lernen und außerschulischen Erfahrungen im Grundschulalter*, S. 34-43. Opladen. Leske & Budrich.

Gathercole, S.E. & Baddeley, A.D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, 8, 259-272.

Gölitz, D., Roick, T. & Hasselhorn, M. (in Druck). Deutscher Mathematiktest für vierte Klassen (DEMAT 4). Göttingen: Beltz Test.

Götz, M. / Neuhaus-Siemon, E. (1999). Schulanfang auf neuen Wegen - Der Modellversuch in Baden-Württemberg. In: Brügelmann, H. u. a. (Hrsg.): *Jahrbuch Grundschule. Fragen der Praxis-Befunde der Forschung*, Seelze/Velber, 35 - 41.

Götz, M. (2003). Die neue Schuleingangsstufe in Deutschland: Alter Wein in neuen Schläuchen? In: Faust, G.; Götz, M.; Hacker, H. & Rossbach H.-G. (Hrsg.): *Anschlussfähige Bildungsprozesse im Elementar- und Primarbereich*. Bad Heilbrunn, 254-272.

Hasselhorn, M. & Grube, D. (1997). Entwicklung der Intelligenz und des Denkens: Literaturüberblick. In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 15-26). Weinheim: Psychologie Verlags Union.

Hasselhorn, M. & Körner, K. (1997). Nachsprechen von Kunstwörtern: Zum Zusammenhang zwischen Arbeitsgedächtnis und syntaktischen Sprachleistungen bei Sechs- und Achtjährigen. *Zeitschrift für Entwicklungspsychologie u. Pädagogische Psychologie*, 29, 212-224.

Hasselhorn, M. & Mähler, C. (1998). Wissen, das auf Wissen baut. Entwicklungspsychologische Erkenntnisse zum Wissenserwerb und zum Erschließen von Wirklichkeit im Grundschulalter. In J. Kahler (Hrsg.), *Wissenserwerb in der Grundschule: Perspektiven erfahren, vergleichen, gestalten* (S. 73-89). Bad Heilbrunn: Klinkhardt.

Hasselhorn, M. & Schumann-Hengsteler, R. (1998). Arbeitsgedächtnis. In D.H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 14-17). Weinheim: Psychologie Verlags Union.

- Hasselhorn, M. (1996). Grundschulfähigkeit im Lichte neuerer pädagogisch-psychologischer Forschung: Rechtzeitig eingeschulte Kinder profitieren von natürlicher Lernfreude. *Eltern-Journal* 3/96, S. 3.
- Helmke, A. & Weinert, F.E. (1997). Bedingungsfaktoren schulischer Leistungen. In F.E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (Enzyklopädie der Psychologie, Themenbereich D, Serie I, Bd. 3) (S. 71-176). Göttingen: Hogrefe.
- Helmke, A. & Weinert, F.E. (1997). Bedingungsfaktoren schulischer Leistungen. In F.E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (Enzyklopädie der Psychologie, Themenbereich D, Serie I, Bd. 3) (S. 71-176). Göttingen: Hogrefe.
- Helmke, A. (1993). Entwicklung der Lernfreude vom Kindergarten bis zur 5. Klassenstufe. *Zeitschrift für Pädagogische Psychologie*, 7, 77-86.
- Horn, W.F. & Packard, T. (1985). Early identification of learning problems: A meta-analysis. *Journal of Educational Psychology*, 77, 597-607.
- Jansen, H., Mannhaupt, G., Marx, H. & Skowronek, H. (1999). Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC). Göttingen: Hogrefe.
- Klauer, K.J. (1993). Induktives Denken beeinflusst das Rechtschreib-Lernen. *Zeitschrift für Entwicklungspsychologie u. Pädagogische Psychologie*, 25, 352-365.
- Knörzer, W. (Hrsg.) (1985). Sind Schüler in kombinierten Grundschulklassen benachteiligt? *Baltmannsweiler*.
- Kossakowski, A. (1961). Wie überwinden wir die Schwierigkeiten beim Lesen- und Schreibenlernen, insbesondere bei Lese-Rechtschreibschwäche? Berlin: Volk und Wissen.
- Krajewski, K., Küspert, P. & Schneider, W. (2002). Deutscher Mathematiktest für erste Klassen (DEMAT 1+). Göttingen: Beltz Test.
- Krajewski, K., Liehm, S. & Schneider, W. (2004). Deutscher Mathematiktest für zweite Klassen (DEMAT 2+). Göttingen: Beltz Test.
- Laging, R. (1995): Altersgemischte Gruppen in der Grundschule. In: Eberwein, H. & Mand, J. (Hrsg.): *Forschen für die Schulpraxis. Was Lehrer über Erkenntnissequalitativer Sozialforschung wissen sollten*. Weinheim, 117-136.
- Lehmann, R.H., Peek, R. & Poeschke, J. (1997). HAMLET 3-4. Hamburger Lesetest für 3. und 4. Klassen. Weinheim: Beltz.
- Marx, H. (1992). Vorhersage von Lese-Rechtschreibschwierigkeiten in Theorie und Anwendung. Unveröffentlichte Habilitationsschrift, Universität Bielefeld.
- Marx, H. (1997). Erwerb des Lesens und des Rechtschreibens: Literaturüberblick. In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 85-111). Weinheim: Psychologie Verlags Union.
- Marx, H. (1998). Knuspels Leseaufgaben (KNUSPEL-L). Göttingen: Hogrefe.
- Marx, P. & Schneider, W. (2000). Entwicklung eines Tests zur phonologischen Bewusstheit im Grundschulalter. In M. Hasselhorn, W. Schneider & H. Marx (Hrsg.), *Diagnostik von Lese-Rechtschreibschwierigkeiten. Tests & Trends*, N.F.1. Göttingen: Hogrefe.
- Melchers, U. & Preuß, U. (1993). Kaufman Assessment Battery for Children (K-ABC) (Deutsche Version). Lisse: Swets & Zeitlinger.
- Melchers, U. & Preuß, U. (1993). Kaufman Assessment Battery for Children (K-ABC) (Deutsche Version). Lisse: Swets & Zeitlinger.
- Piaget, J. (1930). *The child's conception of physical causality*. London: Routledge & Kegan Paul (franz. Orig. 1927).
- Roick, T., Gölitz, D. & Hasselhorn, M. (2004). Deutscher Mathematiktest für dritte Klassen (DEMAT 3+). Göttingen: Beltz Test.

Scheerer-Neumann, G. (1997). Lesen und Leseschwierigkeiten. In F.E. Weinert (Hrsg), Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Themenbereich D, Serie I, Bd. 3) (S. 279-325). Göttingen: Hogrefe.

Schneider, W. (1997). Rechtschreiben und Rechtschreibschwierigkeiten. In F.E. Weinert (Hrsg), Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Themenbereich D, Serie I, Bd. 3) (S. 327-363). Göttingen: Hogrefe.

Skowronek, H. & Marx, H. (1989). The Bielefeld longitudinal study on early identification of risks in learning to write and read: Theoretical background and first results. In M. Brambring, F. Lösel & H. Skowronek (Hrsgs.), Children at risk: Assessment, longitudinal research, and intervention (S. 268-294). New York: de Gruyter.

Stern, E. (1997). Mathematik. In F.E. Weinert (Hrsg), Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie, Themenbereich D, Serie I, Bd. 3) (S. 398-426). Göttingen: Hogrefe.

Terhart, E. (1997<sup>2</sup>). Lehr- und Lernmethoden. Weinheim, München.

Tewes, U. (Hrsg.) (1983). Hamburg-Wechsler-Intelligenztest für Kinder - Revision 1983 (HAWIK-R). Bern: Huber.

Wang, M.C., Haertel, G.D. & Walberg, H.J. (1993). Toward a knowledge base for school learning. Review of Educational Research, 63, 249-294.

Weinert, F.E. & Helmke, A. (Hrsg.) (1997). Entwicklung im Grundschulalter. Weinheim: Psychologie Verlags Union.

Weinert, F.E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In: Weinert, F. E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim, Basel, 17-31.

Winkelmann, W. (1976). Testbatterie Menge und Zahl. Unveröffentlichtes Testverfahren, Pädagogische Hochschule Rheinland, Abteilung Köln

## **Zur Konzeption der Modelle**

Engemann, Ch. (1997). „Schulanfang auf neuen Wegen. Konzeption und operative Planung der Erprobungen an Grundschulen in Baden-Württemberg ab dem Schuljahr 1997/98“, Ministerium für Kultus, Jugend und Sport, Baden-Württemberg, Ref. 33.

Engemann, Ch. (1997). Schulanfang auf neuen Wegen. Für eine kindgerechte Einschulung. In Elternjournal. Zeitschrift für Eltern von Schülerinnen und Schülern der Klassenstufen 1 – 10 (S. 3).

Engemann, Ch. (1998). Richtige Antwort auf veränderte Kindheitsbedingungen: Schulanfang auf neuen Wegen. Positive Antworten auch mit dem zweiten Einschulungstermin. In Elternjournal. Zeitschrift für Eltern von Schülerinnen und Schülern der Klassenstufen 1 – 10 (S. 4-5).

Ministerium für Kultus, Jugend und Sport (Hrsg.) (1996). Schulanfang auf neuen Wegen. Dokumentation Symposium Grundschule, 6. Mai 1996, Stuttgart.

Ministerium für Kultus, Jugend und Sport (Hrsg.) (2001). Schulanfang auf neuen Wegen. Dokumentation eines Projekts der Schuleingangsstufe. Informationen für Schulen, Kindergärten und Eltern. Diese Broschüre enthält u.a. folgende Aufsätze (Nachdrucke aus anderen Publikationen) zum Thema:

- Engemann, Ch. Schulanfang auf neuen Wegen, Grundsatzüberlegungen (S. 50-56).
- Meiers, K. Zum Problem der Schulfähigkeit. Politik und Pädagogik sind in gleicher Weise gefordert (S. 57-59).
- Götz, M., Neuhaus-Siemon, E. Schulanfang auf neuen Wegen - Der Modellversuch in Baden-Württemberg (S. 60-63).

- Schulanfang auf neuen Wegen – Was ändert sich für Erzieherinnen? Interview mit Kultusministerin Annette Schavan (S. 64-66).
- Sollen Kinder schon ab 5 Jahren in die Schule gehen? Beitrag von Kultusministerin Annette Schavan (S. 67-68).
- Engemann, Ch. Schulanfang auf neuen Wegen macht Schule (S. 72-73).

Ministerium für Kultus, Jugend und Sport Baden-Württemberg (Hrsg.). Pressemitteilungen

- Nr. 47/99 vom 25. März 1999: Schulanfang auf neuen Wegen. Schlussfolgerungen aus dem Zwischenbericht der Erprobungsschulen
- Nr. 93/99 vom 25. Mai 1999: SWR3 – Film „Schultüten zweimal im Jahr - Schulanfang auf neuen Wegen“
- Nr. 167/2000 vom 4. August 2000: Einschulungspraxis verändert sich – Senkung der Zurückstellungsquoten, Erhöhung der Vorzeitigenquoten.

