



Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich · KBL
Centre de compétences en évaluation des formations et des acquis à l'Université de Zurich · CEA
Competence Centre for Educational Evaluation and Assessment at the University of Zurich · CEA



Lernstandserhebung bei Schülerinnen und
Schülern der dritten Klassen im Kanton Zürich

Überarbeitete Offerte zuhanden
der Bildungsdirektion des Kantons Zürich,
Bildungsplanung

Urs Moser & Judith Hollenweger
Zürich, 11. Oktober 2005

Anschrift der Verfasserinnen und Verfasser:

Kompetenzzentrum für Bildungsevaluation und Leistungsmessung
an der Universität Zürich

Wilfriedstrasse 15
CH-8032 Zürich

Tel. 043 268 39 61

Fax 043 268 39 67

E-Mail: Urs.Moser@access.unizh.ch

Pädagogische Hochschule Zürich
Departement Forschung und Entwicklung

Hirschengraben 28
CH-8090 Zürich

Tel. 044 305 51 54

Fax 044 305 54 55

E-Mail: Judith.Hollenweger@phzh.ch

Inhalt

1	Einleitung.....	4
2	Fragestellungen.....	4
3	Theoretische und bildungspolitische Grundlagen der Fragestellungen.....	7
3.1	Zur Erfassung fachlicher Kompetenzen.....	7
3.2	«Überfachliche Kompetenzen»: Kompetenzüberzeugungen, Lernfreude und soziale Integriertheit.....	9
3.3	Chancengleichheit des Bildungssystems	12
3.4	Der Einbezug von TaV und QUIIMS	13
3.5	Arbeitssituation in Reformprojekten	14
3.6	Sonderpädagogisches Angebot.....	15
4	Methodisches Vorgehen.....	17
4.1	Bildung der Stichprobe	17
4.2	Klassenwechsler.....	17
4.3	Erhebungsinstrumente	18
4.4	Untersuchungsablauf in den Klassen	19
4.5	Skalierung	20
4.6	Ergebnisrückmeldung für die Lehrpersonen.....	21
4.7	Auswertung von (Längsschnitt-)Daten aus Cluster-Stichproben	22
5	Zeitplan und Zuständigkeit	23
6	Projektorganisation	24
7	Literatur	25

1 Einleitung

Wenn Wirkungen von Schule und Unterricht überprüft werden sollen, führen (quasi)experimentelle Längsschnittstudien am ehesten zu wissenschaftlich zuverlässigen Ergebnissen. Die Bildungsdirektion des Kantons Zürich hat deshalb im Jahr 2003 den Lernstand der Schülerinnen und Schüler zu Beginn der ersten Klasse erheben lassen und damit die Basis für eine Längsschnittstudie geschaffen. Die beim Schuleintritt untersuchten Kinder sollen am Ende der dritten Klasse erneut in den fachlichen und überfachlichen Kompetenzen getestet werden. Der Lernfortschritt während den ersten drei Jahren der Primarschule soll in Abhängigkeit von verschiedenen Merkmalen der Schule und des Unterrichts erklärt werden.

Mit der vorliegenden Offerte bewerben sich das Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich (KBL) und die Pädagogische Hochschule Zürich (PHZH) für die Fortsetzung der Längsschnittstudie, welche die beiden Institutionen vor drei Jahren gemeinsam begonnen haben. Die Offerte greift die in der Ausschreibung vom 25. Mai 2005 festgehaltenen Erkenntnissinteressen der Bildungsdirektion des Kantons Zürich auf und fasst sie in Form von Fragestellungen zu sechs Themenschwerpunkten zusammen. Die sechs Themenschwerpunkte strukturieren das methodische Vorgehen sowie die Datenauswertung und Berichterstattung. Die Fragestellungen sind zum Teil eher bildungspolitisch, zum Teil eher pädagogisch beziehungsweise methodisch-didaktisch ausgerichtet. Mit dem geplanten Projekt sollen sowohl aktuelle Fragen der Bildungspolitik beantwortet als auch Erkenntnisse für die Schulentwicklung und die Lehrerinnen- und Lehrerbildung gewonnen werden.

2 Fragestellungen

Ziel der zweiten Erhebung im Rahmen der Zürcher Längsschnittstudie ist es, den Lernstand der Schülerinnen und Schüler in fachlichen und überfachlichen Kompetenzen am Ende der dritten Klasse zu messen und in Bezug zum Lernstand zu Beginn der ersten Klasse zu setzen. Von besonderem Interesse ist es, Erkenntnisse über die Wirkung von Schule und Unterricht auf den Lernerfolg während der ersten drei Jahre der Primarschule zu gewinnen. Auf Grund der Lernstandserhebung beim Schuleintritt sowie unter Berücksichtigung der Erkenntnissinteressen der Bildungsdirektion des Kantons Zürich sollen mit der zweiten Erhebung im Rahmen der Längsschnittstudie folgende Fragestellungen bearbeitet werden:

1. Fachliche Kompetenzen am Ende der 3. Klasse

- Welche Kompetenzen erreichen die Schülerinnen und Schüler am Ende der dritten Klasse in den Fächern Sprache und Mathematik?
- Wie gross ist der Anteil der Schülerinnen und Schüler, die die Ziele des Lehrplans in ausgewählten Fachbereichen erreichen?
- Wie gut werden die Kompetenzen der Schülerinnen und Schüler am Ende der dritten Klasse durch den Lernstand zu Beginn der ersten Klasse erklärt?

2. Unterricht, fachliche und überfachliche Kompetenzen

- Welchen Stand haben die überfachlichen Kompetenzen der Schülerinnen und Schüler am Ende der dritten Klasse erreicht?
- Welche Bedeutung haben überfachliche Kompetenzen für die Entwicklung fachlicher Kompetenzen?
- Besteht eine Wechselwirkung zwischen fachlichen und überfachlichen Kompetenzen?
- Hat die Unterrichtsgestaltung einen Einfluss auf die Entwicklung fachlicher und überfachlicher Kompetenzen?

3. Chancengleichheit des Bildungssystems

- Unterscheiden sich die Leistungen nach individuellen Merkmalen wie kognitive Leistungsfähigkeit (Intelligenz), Geschlecht, soziale Herkunft, Immigrationsstatus und Erstsprache der Kinder?
- Sind die zu Beginn der Schulzeit nachgewiesenen Zusammenhänge zwischen Leistungen und individuellen Merkmalen (kognitive Leistungsfähigkeit, Geschlecht, soziale Herkunft, Immigrationsstatus und Erstsprache) gleich geblieben, kleiner geworden oder grösser geworden?
- Sind die zu Beginn der Schulzeit nachgewiesenen Leistungsunterschiede zwischen den Klassen gleich geblieben, kleiner geworden oder grösser geworden?
- Wie gross ist der Effekt des Einzugsgebiets der Schule beziehungsweise des Anteils Schülerinnen und Schüler aus sozial benachteiligten Verhältnissen und des Anteils fremdsprachiger Kinder pro Klasse auf die Leistungen?

4. Wirkungen der Reformprojekte

- Unterscheiden sich die Leistungen zwischen den Schülerinnen und Schülern der Regelklassen, der TaV-Klassen und der QUIMS-Klassen?
- Sind die zu Beginn der Schulzeit nachgewiesenen Leistungsunterschiede zwischen Regelklassen und QUIMS-Klassen gleich geblieben, kleiner geworden oder grösser geworden?

5. Arbeitssituation in Reformprojekten

- Unterscheidet sich die (wahrgenommene) individuelle Arbeitssituation der Lehrpersonen in Reformprojekten von derjenigen in Regelklassen?
- Unterscheidet sich die (wahrgenommene) Organisationskultur in Reformprojekten von derjenigen in Regelklassen?

6. Sonderpädagogisches Angebot

- Welche sonderpädagogischen Massnahmen (Kleinklassen/Stütz- und Fördermassnahmen) erhalten die Schülerinnen und Schüler gegenwärtig und in welchem Umfang?
- Welche sonderpädagogischen Massnahmen haben die Schülerinnen und Schüler im Verlauf der ersten bis dritten Klasse erhalten?
- Welchen Einfluss haben sonderpädagogische Massnahmen auf fachliche Kompetenzen?
- Welchen Einfluss haben sonderpädagogische Massnahmen auf überfachliche Kompetenzen?

Neben den sechs aufgeführten Themenschwerpunkten wird im Rahmen der Längsschnittstudie in einem Zusatzprojekt der Themenschwerpunkt «Literale Förderung unter den Bedingungen einer heterogenen Schülerschaft» bearbeitet. Dieser Themenschwerpunkt ist Bestandteil einer separaten Offerte der PHZH.

3 Theoretische und bildungspolitische Grundlagen der Fragestellungen

3.1 Zur Erfassung fachlicher Kompetenzen

Leistungsmessungen in der Schweiz werden sich in Zukunft nach den Ergebnissen des Projekts HarmoS (Harmonisierung der obligatorischen Schule) der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren (EDK) zu richten haben. Mit HarmoS werden Bildungsstandards definiert und im Sinne von Mindeststandards politisch festgelegt. Die Bildungsstandards beruhen auf empirisch validierten Kompetenzmodellen, die eine Beschreibung des hierarchischen und kumulativen Aufbaus der Kompetenzen enthalten. Die Entwicklung hin zur Leistungsmessung auf der Grundlage von Kompetenzmodellen soll auch für die Züricher Längsschnittstudie genutzt werden.

Im Hinblick auf die Leistungsmessungen am Ende der sechsten und neunten Klasse sollen die Leistungen am Ende der dritten Klasse so erfasst werden, dass die Berichterstattung über den Lernstand zu allen drei Zeitpunkten anhand der gleichen Dimensionen erfolgen kann und die Kompetenzen jeweils auf der gleichen Skala beschrieben werden können. Ein einheitliches Reporting über die Kompetenzen der Schülerinnen und Schüler ist informativ und wünschenswert, weil es die Kommunikation der Ergebnisse erleichtert. Zudem ist es auch aus methodischen Gesichtspunkten notwendig, den Lernstand für die gleichen Dimensionen zu erfassen, weil dadurch differenziertere Analysen möglich werden und die Fortschritte der Schülerinnen und Schüler zuverlässig nachgewiesen werden können.

Die Erfassung der fachlichen Kompetenzen hat darüber hinaus zwei weitere Aspekte zu berücksichtigen. Zum einen handelt es sich um eine Fortsetzung der Lernstandserhebung beim Schuleintritt. Es gilt abzuklären, inwieweit zu einzelnen Teilen der ersten Erhebung der Längsschnittstudie eine direkte Verbindung hergestellt werden kann. Möglichkeiten dazu ergeben sich beim Wortschatz und bei der Arithmetik. Zum andern besteht die Möglichkeit, eine Verbindung zu den Ergebnissen der im Jahr 2001 durchgeführten Querschnittstudie in den dritten Klassen zu schaffen. Beide Verbindungen setzen voraus, dass eine bestimmte Anzahl Aufgaben dieser Studien erneut eingesetzt werden können.

Die Längsschnittstudie beschränkt sich auf die beiden Fächer «Mathematik» und «Sprache». Um Ergebnisse zu erhalten, die einerseits wissenschaftlich zu verantworten sind und andererseits von Praxis und Öffentlichkeit anerkannt werden, ist eine unterrichtsnahe Entwicklung der Messinstrumente unabdingbar. Die Instrumente zur Messung der fachlichen Kompetenzen werden deshalb gemeinsam mit Fachdidaktikerinnen und Fachdidakti-

kern sowie mit fachdidaktisch versierten Lehrpersonen des Kantons Zürich entwickelt. Dabei wird folgendes Vorgehen eingehalten:

- Auswahl und Definition der Inhalte auf der Grundlage des Lehrplans (Content Standards)
- Bestimmung von Leistungsanforderungen auf der Grundlage von Kompetenzmodellen (Performance Standards)
- Abklärungen über die Verbindung des geplanten Tests mit den Instrumenten der bisherigen Lernstandserhebungen beim Schuleintritt und am Ende der dritten Klasse
- Bestimmung von Bearbeitungsdauer, Anzahl Aufgaben und Aufgabenformaten
- Entwicklung der Aufgaben und redaktionelle Überarbeitung der Aufgaben
- Zusammenstellung der Aufgaben zu einem Pilottest
- Erprobung der Aufgaben bei rund 150 Schülerinnen und Schülern (Pilottest)
- Prüfung der Aufgaben auf Verständlichkeit, Schwierigkeit, Zuverlässigkeit und Übereinstimmung mit dem Kompetenzmodell
- Zusammenstellen der Aufgaben zu einem definitiven Test

Zur Beurteilung der Zielerreichung im Lehrplan wird bestimmt, welche Kompetenzen vorhanden sein müssen, damit die grundlegenden Lehrplanziele erreicht sind. Auf Grund der zur Verfügung stehenden Ressourcen werden die fachlichen Leistungen im Klassenverband erhoben, weshalb im Fach «Sprache» die mündliche Kommunikation (Teilbereich «Sprechen») nicht erfasst werden kann. Tabelle 3.1 gibt einen Überblick über die Lehrplanbereiche, die mit der Leistungsmessung erfasst werden sollen.

Tabelle 3.1: Ausgewählte Lehrplanbereiche in den Fächern Mathematik und Sprache

Lehrplanbereiche Mathematik	Inhalte
Mengen, Eigenschaften von Zahlen	Ordnen von Objekten, Mengenbildung nach ein, zwei oder drei Merkmalen
Zahlenbereich	Ordinalzahl, Kardinalzahl, Zahlwörter
Zahlenschreibweise	Zehner, Hunderter, Tausender
Operationen	Addition, Subtraktion, Multiplikation, Division
Grössen	Geld, Längen, Zeit
Geometrie	Raubegriffe, Räumliches Vorstellungsvermögen
Lehrplanbereiche Sprache	Inhalte
Texte lesen und verstehen Informationen verarbeiten	Sinnverständnis auf Grund eines Lesetextes
Sprachbetrachtung Wortschatz	Satzarten, Rechtschreibung, Interpunktion, Wortarten, Adäquate Wortwahl
Zuhören und Verstehen Texte für andere schreiben	Sinnverständnis auf Grund eines Hörtextes Schriftliche Sprachproduktion
Kommunikatives Handeln	

Die geplante Bearbeitungszeit für den Mathematik- und Sprachtest beträgt je 80 Minuten (viermal 20 Minuten mit Pausen zwischen den einzelnen Teilen). Weil mit dem Test Kompetenzen beschrieben werden sollen, nicht etwa das Lösen von Testaufgaben unter Zeitdruck, steht den Schülerinnen und Schülern genügend Bearbeitungszeit zur Verfügung. Je mehr ein Test zeitlich begrenzt wird, umso weniger lässt sich seine inhaltliche Validität in Bezug auf die erreichte Kompetenz nachweisen, weil bei der Bearbeitung nicht nur die Kompetenz, sondern auch andere Aspekte (Arbeiten unter Zeitdruck) geprüft werden.

3.2 «Überfachliche Kompetenzen»: Kompetenzüberzeugungen, Lernfreude und soziale Integriertheit

«Überfachliche Kompetenzen» ist ein weiter Begriff, unter dem die unterschiedlichsten Dinge subsumiert werden beziehungsweise werden können: unter anderem handlungsbezogene und kognitive Fähigkeiten aber auch Einstellungen und Grundhaltungen (Maag Merki & Grob, 2003). Auf der Basis einer inhaltsanalytischen Auswertung von Leitbildern beziehungsweise allgemeinen Bildungszielen von Lehrplänen extrahierten Grob, Maag Merki und Bühler (2003) 33 Zieldimensionen, denen 28 wissenschaftliche Konstrukte zugeordnet wurden. Fokussiert man auf diejenigen Zieldimensionen, die als primäre personale Ressourcen für erfolgreiches Lernen aufgefasst werden können, dann spielen auf Seiten der wissenschaftlichen Konstrukte insbesondere *Lern- und Arbeitsstrategien* sowie *Selbstwirksamkeit* eine prominente Rolle (vgl. Grob, Maag Merki & Bühler, 2003, S. 315).

Auf die Bedeutung von *Lern- und Arbeitsstrategien* (im Sinne u. a. von kognitiver und metakognitiver Regulation) für Lernprozesse und Leistungen wird immer wieder verwiesen (vgl. u. a. Weinert & Schrader, 1996). Dass im Rahmen der geplanten Lernstandserhebung 2006 auf die Erhebung von Strategien verzichtet wird, hat im Wesentlichen drei Gründe: *Zweifelhafter (prädiktiver) Nutzen selbstdeklarerter Strategien*: Zwischen mit Fragebogen erhobenen (allgemeinen) Strategien und Lernleistungen zeigen sich in der Regel nur schwache Zusammenhänge (Ainley, 1993; Patrick, Ryan & Pintrich, 1999; Pintrich & Garcia, 1993; VanderStoep, Pintrich & Fagerlin, 1996). Engere Beziehungen finden sich in Interviewstudien (Purdie & Hattie, 1996; Zimmerman & Martinez-Pons, 1986, 1988, 1990) und vor allem bei Interventionsstudien (Gürtler, 2003; Hattie, Biggs & Purdie, 1996). *Fragwürdige Validität von Selbstauskünften im Bereich von Strategien*: Zwischen via Fragebogen und mit anderen Messansätzen erhobenen Strategien existieren keineswegs immer hohe und signifikante Zusammenhänge (Beobachtungen: Wuttke, 2000; Computerprotokolle: Winne & Jamieson-Noel, 2002; Analyse von Videos: Ardel & Schellhas, 1996; Lerntagebücher: Mayring, 1999). Ursache hierfür könnte sein, dass die eingesetzten Fragebögen mit wenigen Ausnahmen (Pokay & Blumenfeld, 1990) situationsübergreifende Strategien erfassen (Boekaerts & Minnaert, 1999; VanderStoep, Pintrich & Fragelin, 1996) und auf potentielle Kompetenzen zielen, d.h. auf schülerseitiges Wissen über vernünftiges Lernverhalten. Solche Auskünfte erlauben nur beschränkt Rückschlüsse auf die Strategienutzung in der je spezifischen Situation (Mayring, 1999; Winne & Perry, 2000). Empfohlen

wird daher, in Zukunft Lernstrategien bereichsspezifisch zu erfassen (Boekaerts, 1999; Schiefele et al. 1995; Winne & Jamieson-Noel, 2002) und unterschiedliche Messansätze zu kombinieren (Mayring, 1999). *Problem eines altersgerechten Instrumentariums*: Fragebogenerhebungen zu Strategien wurden vorwiegend bei Jugendlichen (ab Sekundarstufe I) und jungen Erwachsenen durchgeführt. Grössere Fragebogenerhebungen mit jüngeren Kindern sind uns keine bekannt.¹

Selbstwirksamkeit ist im Sinne von *Kompetenzüberzeugungen* zu verstehen, das Zutrauen in die eigenen Möglichkeiten, Anforderungen bewältigen zu können (vgl. Maag Merki & Grob, 2003, S. 134). Wesentlich für die Initiierung und Aufrechterhaltung von Lernprozessen ist jedoch nicht allein die Überzeugung, etwas schaffen zu können, sondern auch der Anreiz, der persönliche Wert, den jemand einer Thematik, Tätigkeit usw. zuschreibt (vgl. u. a. Krapp & Ryan, 2002). Letztlich geht es hier auch um die positiven Affekte bei der Auseinandersetzung mit einem Gegenstand oder kurz: um die *Lernfreude*. Lernfreude wird von Helmke (1993) verstanden als «relativ überdauernde emotionale Besetzung beziehungsweise affektive Tönung des schulischen Lernens und fachlicher Inhalte» (S. 78).

Überfachliche Kompetenzen im eigentlichen Sinne liegen, wie Maag Merki und Grob (2003) schreiben, «quer zur herkömmlichen Fächerstruktur» (S. 125). Womit Kompetenzüberzeugungen und Lernfreude fachunspezifisch zu verstehen wären. Es zeigt sich indes immer wieder, dass solchermassen allgemeine Konstrukte jeweils einen deutlich geringeren Erklärungs- beziehungsweise Prognosewert hinsichtlich erfolgreichem schulischen Lernen haben als ihre fachspezifischen Gegenstücke (vgl. etwa Pietsch, Walker & Chapman, 2003). Ist man, wie in der Lernstandserhebung, an den Bedingungen erfolgreicher Bewältigung schulischer Anforderungen interessiert, erscheint es daher sinnvoller, Lernfreude und Kompetenzüberzeugungen *fachspezifisch* zu erheben. Angemessener wäre es, in diesem Falle den Begriff «überfachliche Kompetenzen» durch «motivationale und affektive Kompetenzen» (im Unterschied etwa zu konstitutionellen oder kognitiven Bedingungen beziehungsweise Determinanten erfolgreichen Lernens zu ersetzen (vgl. etwa Helmke & Weinert, 1997).

Neben personalen Kompetenzen wie etwa Kompetenzüberzeugungen lassen sich auch interpersonale und gesellschaftsbezogene Kompetenzen unterscheiden (Grob & Maag Merki, 2001). Zu den ersteren gehören etwa kooperative Fähigkeiten oder die Fähigkeit zur Verantwortungsübernahme. Mit anderen Menschen erfolgreich und zufrieden stellend interagieren zu können, ist eine der wichtigsten Lernaufgaben in der Entwicklung und von grundlegender Bedeutung für die Lebensbewältigung und -gestaltung. In der pädagogisch-psychologischen Forschung der letzten 15 Jahre hat dabei ein Aspekt zunehmende Bedeutung gefunden, welcher auch in motivationaler und Leistungshinsicht bedeutsam ist. Die subjektive Wahrnehmung der Beziehungen zu wichtigen Personen und Gruppen,

¹ Wir danken an dieser Stelle Rita Stebler für ihre Hilfe bei der Präzisierung der Offerte in diesem Bereich.

etwa als «sense of belonging» (Anderman & Freeman, 2004) oder als «relatedness» (Deci & Ryan, 1993; Furrer & Skinner, 2003) bezeichnet, wird dabei als eine wichtige Voraussetzung dafür angesehen, dass Schülerinnen und Schüler sich für schulische Ziele und Anforderungen engagieren. Zum einen hat dieses Gefühl, sozial eingebunden zu sein, einen positiven Effekt auf das globale Selbstkonzept, zum anderen stellt es eine wichtige Quelle von Sicherheit dar, welche die Exploration und Aneignung neuer Fähigkeiten begünstigt (Furrer & Skinner, 2003). Auch wenn es sich hier nicht um eine Kompetenz im engeren Sinne handelt, handelt es sich doch um eine zentrale personale Ressource mit vielfältigen Effekten u. a. auch auf das Lernen (vgl. unten). Im Folgenden wird hier für diesen Aspekt die Bezeichnung «soziale Integriertheit» verwendet.

Das Vertrauen in die eigenen Fähigkeiten, etwas erreichen zu können, wie auch das Interesse an fachlichen Inhalten sind zentrale Voraussetzungen für erfolgreiches Lernen. Dies wird auch im *expectancy-value model of achievement motivation* deutlich (Wigfield & Eccles, 2000). Leistungsbezogenes Handeln – z.B. Engagement, Persistenz, Wahl spezifischer Leistungskurse aber auch Leistungen – wird hier als abhängig von *Kompetenzüberzeugungen* und *affektiven Erinnerungen* gesehen. Letztere bilden sich aufgrund wiederholter positiver oder negativer Erfahrungen mit einem Gegenstand, einem Inhalt usw. Lernfreude im Sinne von Helmke (1993, S. 78) ist als eine zentrale Facette der affektiven Erinnerungen zu betrachten. Zahlreiche Studien belegen die Bedeutung von Kompetenzüberzeugungen und Lernfreude (beziehungsweise verwandter Konstrukte) für das Lernverhalten (im Sinne u. a. von Selbststeuerung) wie auch Leistungen (Buff, Reusser & Pauli, in Vorbereitung a; Helmke, 1993, 1997; Marsh, Trautwein, Lüdtke, Köller & Baumert, 2004). Günstige Kompetenzüberzeugungen und Lernfreude sind nicht nur wichtige *Voraussetzungen* für erfolgreiches Lernen, sondern auch *Folge und Ziel* schulischen Unterrichts (vgl. u. a. Ramseier, 2004). Bandura (2003) meint entsprechend: «The major goal of formal education should be to equip students with the intellectual tools, efficacy beliefs, and intrinsic interests needed to educate themselves in a variety of pursuits throughout their lifetime» (p. 214). Wie sich verschiedentlich zeigt, kann die Unterrichtsgestaltung die Entwicklung positiver Kompetenzüberzeugungen und Lernfreude begünstigen, aber auch hemmen (Buff, Reusser & Pauli, in Vorbereitung b; Prenzel, Kramer & Drechsel, 2001; Reeve, 2002; Satow & Schwarzer, 2003). Überfachliche Kompetenzen werden entsprechend nachfolgend verstanden als *motivational-affektive Merkmale* der Schülerinnen und Schüler.

Bezüglich sozialer Integriertheit zeigen verschiedene Studien Effekte auf die schulische Motivation und auf die Leistung. Je positiver soziale Beziehungen wahrgenommen werden, desto eher engagieren sich Schülerinnen und Schüler im Unterricht und für die Schule (Furrer & Skinner, 2003) und entwickeln eine intrinsische Motivation (Freeman & Anderman, 2002; zit. nach Anderman & Freeman, 2004). Verschiedene Studien zeigen zudem, dass das Gefühl, sozial integriert zu sein, direkt (Gutman & Midgley, 2000) oder indirekt – vermittelt durch das Selbstkonzept (Guay, Boivin & Hodges, 1999) – auch schulische Leistungen zu einem gewissen Grad vorherzusagen vermag.

Die drei im Rahmen überfachlicher Kompetenzen inhaltlich näher diskutierten Merkmale wurden bereits in der 1. Klasse erhoben, und es erscheint aus diesem Grund sinnvoll, in der Drittklass-Erhebung mit den gleichen Konstrukten weiterzuarbeiten. Es liegen Instrumente aus der Erhebung in den 1. Klassen sowie aus dem Projekt SEMOS (vgl. Projektbeschreibung; Buff et al., 2004) vor, welche in leicht angepasster Form übernommen werden können.

Tabelle 3.2: Übersicht über die überfachlichen Kompetenzen

Konstrukt	Dimensionen
Kompetenzüberzeugungen	Kompetenzüberzeugung Mathematik Kompetenzüberzeugung Deutsch
Schoolinvolvement	Einstellung zu Schule generell Lernfreude Mathematik Lernfreude Deutsch
Soziale Integriertheit	Integriertheit in der Klasse Beziehung zur Lehrperson

3.3 Chancengleichheit des Bildungssystems

Auf Grund der Beurteilung von Bildungssystemen anhand des Zusammenhangs zwischen sozialer Herkunft und Schulleistung, wie dies im Rahmen des internationalen Schulleistungsvergleichs PISA gemacht wird, ist die Diskussion über die Chancengleichheit des Bildungssystems neu entfacht worden (OECD [Organisation for Economic Co-Operation and Development], 2001, S. 222). Im Vergleich zu den 1960er- und 1970er-Jahren wird die Diskussion über Chancengleichheit allerdings vermehrt von der Bildungsökonomie bestimmt. Je mehr Individuen in die Ausbildung beziehungsweise in ihr Humankapital investieren, um so grösser ist die Chance, dass sie zu einem späteren Zeitpunkt einen Nutzen im Berufsleben erzielen und die Arbeitsproduktivität erhöhen können (Becker, 1975; Blossfeld, 1999, S. 30). Die soziale Wohlfahrt kann nur dann garantiert werden, wenn auch Kinder aus sozial benachteiligten Verhältnissen entsprechend ihrem Potenzial in ihr Humankapital investieren können und dieser Prozess durch die Schule nicht beeinträchtigt, sondern unterstützt wird (OECD, 2004).

Der Chancengleichheit beziehungsweise der Chancengerechtigkeit liegt das Prinzip zugrunde, wonach die Umverteilung der Chancen in einer Schullaufbahn nicht nach zugeschriebenen Kriterien wie Geschlecht, Religionszugehörigkeit oder sozialer Herkunft, sondern nach den schulischen Leistungen erfolgt (Fend, 1974, S. 102). Die Einteilung in unterschiedlich anspruchsvolle Schultypen, in Schultypen mit unterschiedlichem Prestige, wird nach dem Prinzip der Chancengleichheit nur dann als gerecht beurteilt, wenn sie

ausschliesslich auf der Basis unterschiedlicher Leistungen erfolgt. Das Prinzip setzt voraus, dass in der Schule alle die gleiche Chance haben, die erwarteten Leistungen zu erbringen und alle den gleichen Beurteilungskriterien ausgesetzt sind. Leistungsunterschiede sind deshalb einzig auf Begabungen und Neigungen, nicht aber auf persönliche Merkmale der Kinder zurückzuführen. Die Schule hat demnach die Schülerinnen und Schüler unterschiedlicher Herkunft so zu unterrichten und zu beurteilen, dass die beim Schulstart vorhandenen primären sozialen Ungleichheiten nicht grösser werden und dass keine sekundären sozialen Ungleichheiten durch sozialschichtbedingte Übergangsentscheidungen entstehen (Fend, 1974, S. 170).

Die demographischen Entwicklungen in der Schweiz haben allerdings dazu geführt, dass die grösste Gefahr für ein Bildungssystem, das Prinzip der Chancengleichheit nicht einzuhalten, längst nicht mehr nur bei den Selektionsentscheiden beziehungsweise bei den Schnittstellen des Bildungssystems liegt. Die Segregation der Schülerpopulation nach bildungsrelevanten Merkmalen ist vor allem in städtischen Gebieten, aber auch in der Agglomeration so weit fortgeschritten, dass Lehrpersonen in Bezug auf die Zusammensetzung der Klasse und die sprachlichen Kompetenzen der Kinder zu Beginn der Schule sehr unterschiedlich starten (Moser, 2005). Mit der Längsschnittstudie kann verfolgt werden, welche Folge diese gesellschaftlichen Entwicklungen für die Schule und den Lernerfolg der Kinder haben und wie weit es dem Bildungssystem gelingt, diese Ungleichheiten auszugleichen.

3.4 Der Einbezug von TaV und QUIMS

Die Projekte «Teilautonome Volksschule» (TaV) und «Qualität in multikulturellen Schulen» (QUIMS) gehören zu jenen Reformbestrebungen, durch welche die in Absatz 3.3 angesprochenen Themen «Chancengleichheit» und «Bildungsqualität» auf bildungspolitischer Ebene angegangen werden. Das übergeordnete Ziel von TaV besteht darin, dass Lehrpersonen und Schulbehörden die Schulqualität in eigener Verantwortung weiterentwickeln. Zu den Kernelementen der Reform gehört, dass (1) die Schulen eine grössere Handlungsfreiheit erhalten, um auf die Bedürfnisse in ihrer Gemeinde rechtzeitig und flexibel eingehen zu können, (2) Schulleitungen eingerichtet werden, (3) die Zusammenarbeit im Team (Schulleitung und Schulkonferenz) verstärkt wird und (4) pädagogische Schulentwicklung in Zusammenarbeit mit Eltern und unter Einbezug der Interessen der Schülerinnen und Schüler erfolgt. Die Stärkung der Schulen in ihrer Entscheidungskompetenz hat zur Folge, dass zugleich ihre Pflicht zur Rechenschaftslegung und die Verantwortung für die Schulqualität grösser werden. Dazu gehört auch, dass Schulen sich vermehrt über die Kompetenzen ihrer Schülerinnen und Schüler ein Bild machen und dafür besorgt sind, dass einerseits die Ziele des Lehrplans erfüllt werden und andererseits das Prinzip der Chancengleichheit eingehalten wird.

Mit dem Projekt QUIMS soll erreicht werden, dass (1) in den Schulen ein für alle Kinder leistungswirksamer Unterricht angeboten wird [Steigerung der Prozessqualität], (2) in allen Klassen ein hohes Leistungsniveau erreicht wird sowie die Zufriedenheit der Lehrpersonen und Eltern erhöht und das Image der Schulen verbessert werden [Steigerung der Ergebnisqualität] und (3) Kinder aus immigrierten und bildungsfernen Familien in ihrer Schullaufbahn unterstützt werden und die Chancengleichheit gewahrt wird [Steigerung der Systemqualität]. Wie weit die Potenziale der Kinder unter erschwerten Bedingungen genutzt und Chancengleichheit erreicht werden, lässt sich vor allem durch Transparenz über den Lernstand der Schülerinnen und Schüler in Erfahrung bringen.

Die Längsschnittstudie lässt zum ersten Mal eine Überprüfung der Wirkungen der beiden Reformen TaV und QUIMS auf fachliche und fachübergreifende Kompetenzen zu. Schulen der beiden Reformprojekte wurden bei der Stichprobenbildung für die Lernstandserhebung zu Beginn der ersten Klasse je als eigenes Stratum behandelt. Trotz methodisch guter Voraussetzungen für eine Wirkungsüberprüfung sollten aber folgende Punkte bei der Beurteilung der Ergebnisse berücksichtigt werden. Die Pflichten der teilautonomen Volksschulen in Bezug auf Transparenz über die Kompetenzen der Schülerinnen und Schüler sind nicht geregelt. Die Steigerung der Schulqualität gehört zwar zu den Zielen der teilautonomen Schulen. Die Leistungssteigerung gehört allerdings nicht explizit zu den Kernelementen der Reform. Auch die Ziele von QUIMS sind weit vielfältiger, als dass deren Erreichung mit der Lernstandserhebung am Ende der dritten Klasse hinreichend geprüft werden könnte. Die Förderung der Qualität in multikulturellen Schulen schliesst zwar die Förderung der Schülerinnen und Schüler in den Kulturtechniken mit ein. Darüber hinaus werden mit QUIMS aber sehr viele weitere Ziele verfolgt. Wie gut diese Ziele erreicht werden, wird im Rahmen des Zürcher Längsschnitts nicht beantwortet.

3.5 Arbeitssituation in Reformprojekten

Die Reformprojekte TaV und QUIMS sollen Schulen befähigen, auf den gesellschaftlichen Wandel und die damit verbundenen Anforderungen an die pädagogische und organisatorische Praxis zu reagieren und die Schulqualität zu erhalten und zu verbessern. Mit den durch die Reformprojekte implementierten Neuerungen verändern sich die Aufgabengebiete aller Mitarbeiterinnen und Mitarbeiter einer Schule. Das Schulkollegium («Schulkonferenz») ist für die Qualitätssicherung und die pädagogische Entwicklung verantwortlich («Schulprogrammarbeit»). Es setzt eigene pädagogische Schwerpunkte und erarbeitet eine Jahresplanung. Zur Qualitätsentwicklung führen die Lehrkräfte gemeinsam pädagogische Projekte durch, die durch Arbeitsgruppen vorbereitet werden. Schulintern werden Weiterbildungen organisiert und die Zusammenarbeit wird durch Teamentwicklung unterstützt. Die Eltern sowie die Schülerinnen und Schüler (TaV) werden stärker in die Gestaltung des Schullebens einbezogen. Am Ende eines Jahres evaluiert das Schulkollegium die Zielsetzungen (vgl. dazu Sträuli 2000).

Die *Organisationskultur* der Schule und die *Arbeitssituation* der einzelnen Lehrperson ändern sich damit grundlegend, da die Schule sich gegen innen und nach aussen öffnet und neben dem «Kernauftrag» des Unterrichts in der eigenen Klasse neu gemeinsame Aufgaben und Verantwortungen der Schule als Organisation übernommen werden müssen. Die einzelne Lehrperson ist herausgefordert, sich nicht mehr als Einzelkämpferin ihrer Profession zu verstehen, sondern als Teil einer Schule, die in Kooperation mit allen Beteiligten für ihre Qualität verantwortlich ist. Diese neuen Aspekte im Tätigkeitsprofil einer Lehrperson sind in erster Linie als Ressourcen zu verstehen, welche die Lehrperson in ihrer Unterrichtstätigkeit unterstützen, entlasten und die Arbeitszufriedenheit verbessern bzw. verbessern sollen (Nakamura, Leemann und Hollenweger 2005). Es wird davon ausgegangen, dass Merkmale der Arbeitssituation eine wichtige Voraussetzung – aber noch keine Garantie – für einen wirksamen Unterricht der Lehrperson verkörpern (Rüesch 1999, S. 28).

Untersucht werden soll, inwiefern sich in den Reformschulen die Arbeitssituation der befragten Lehrpersonen von derjenigen in anderen Schulen unterscheidet. Dabei soll einerseits a) die *individuelle Arbeitssituation* der Lehrperson angesprochen, andererseits allgemeiner auch b) die *Organisationskultur in der Schule* fokussiert werden.

3.6 Sonderpädagogisches Angebot

Schülerinnen und Schüler, welche Schwierigkeiten bei Lernen und beim Zusammenleben in der Schule zeigen, können durch sonderpädagogische Massnahmen unterstützt werden. Diese Massnahmen werden einerseits zusätzlich zum Unterricht in der Klasse angeboten (Stütz- und Fördermassnahmen), andererseits werden gewisse Kinder zur Förderung auch besonderen Schulungsformen (Kleinklassen A, B, C, D und E oder ISF) zugewiesen. Die Bildungsstatistik des Kantons Zürich weist für 2003 eine Segregationsrate von gut 3% für die dritten Klassen aus, während insgesamt auf der Unterstufe in diesem Schuljahr fast 8% in besonderen Klassen unterrichtet wurden². Die statistische Erfassung der Stütz- und Fördermassnahmen befindet sich im Kanton Zürich noch in der Erprobungsphase. In der 2001 durchgeführten Evaluation der dritten Klassen zeigte sich im Kanton Zürich, dass über 50% der Schülerinnen und Schüler bis zur dritten Klasse bereits eine solche Massnahme erhalten hatten (Moser & Tresch, 2003).

Wie in den Statistiken der OECD (2003, 2004) dargestellt, unterscheiden sich die Förderraten in den verschiedenen Ländern sehr stark. Über die Wirkung dieser Massnahmen ist jedoch wenig bekannt. Gut dokumentiert ist die Tatsache, dass in der Schweiz und in Deutschland Kinder mit Migrationshintergrund häufiger in gesonderten Klassen geschult werden (Kronig, Haeberlin & Eckhart, 2000 und Wagner & Powell, 2003) sowie die generell höhere Förderrate bei Knaben (vgl. OECD, 2003). Es gibt auch Hinweise dafür, dass

² vgl. <http://www.egov.bista.zh.ch/defaultEgov.htm>

sozial benachteiligte Kinder häufiger sozial weniger erwünschten Massnahmen, insbesondere Sonderklassen, zugewiesen werden und diese Gruppe im Vergleich zu ihrem effektiven Förderbedarf weniger unterstützt werden als Kinder aus privilegierten Familien (Sacker, Schoon & Bartley, 2001). Zudem sind auch weitere systemische Faktoren, wie die Zusammensetzung der Schule/Klasse oder die Verfügbarkeit von fachlichem Personal, für die Zuweisung zu einer bestimmten Massnahme mitverantwortlich (Benkmann, 2003; Emanuelsson & Persson 1997). Die soziale Selektivität von Bildungssystemen (vgl. Coradi Vellacott et al. 2003) zeigt sich somit insbesondere auch bei der Zuweisung zu sonderpädagogischen Massnahmen.

Die Zuweisung zu den sonderpädagogischen Massnahmen wird im Kanton Zürich meistens durch die Lehrperson in Gang gesetzt. Die Frage der diagnostischen Kompetenz – neben der eigentlichen Unterrichtskompetenz – stellt sich deshalb mit grosser Dringlichkeit: Gelingt es den Lehrpersonen, Schülerinnen und Schüler zu identifizieren, welche tatsächlich unter den angenommenen Störungen leiden und von einer besonderen Förderung am meisten profitieren könnten? Aus Deutschland gibt es verschiedene Hinweise, dass dies nicht der Fall ist. Mand (2002) zeigt auf, dass in diesen Zuweisungsprozessen die Variable «Gutachterpersönlichkeit» die höchste Erklärungsmacht besitzt. Auch leiden die verwendeten «Diagnosen» und die genannten Gründe zur Zuweisung in besondere Klassen an einer hohen Selbstreferenz: Kinder mit einer Lernbehinderung werden einer Lernbehindertenklasse zugewiesen und erhalten dort eine besondere Pädagogik für Lernbehinderte (vgl. Langfeldt 1998).

Die Zusammenhänge zwischen Leistungs- und Verhaltensschwierigkeiten, sozioökonomischen sowie demografischen Faktoren sind sehr komplex. Klauer (2003) schlägt deshalb vor, zur Feststellung von substantiellen Lernschwierigkeiten alle Kinder in einer bestimmten Population regelmässig zu erfassen, so dass eine gerechte Auswahl der Kinder mit den grössten Schwierigkeiten und somit ihre Förderung möglich wird. Es ist weitgehend unbeantwortet, ob in einem bestimmten Bildungssystem tatsächlich diejenigen Kinder eine Unterstützung erhalten, welche sie am dringendsten nötig hätten. Die komplexen Zusammenhänge zwischen fachlichen und überfachlichen Kompetenzen und dem erfolgreichen Teilnehmen am regulären Primarschulunterricht sind bis heute kaum erforscht. Zur Klärung dieser Fragen kann die Lernstandserhebung in den dritten Klassen einen wesentlichen Beitrag leisten.

4 Methodisches Vorgehen

4.1 Bildung der Stichprobe

Für die Längsschnittstudie wurde eine repräsentative Stichprobe jener Schülerinnen und Schüler gebildet, die im Schuljahr 2003/04 im Kanton Zürich in die erste Regelklasse eintraten. Die Stichprobe wurde so gebildet, dass neben repräsentativen Ergebnissen für den Kanton Zürich auch die Überprüfung der Wirkungen der Projekte «TaV» und «QUIMS» auf fachliche und fachübergreifende Kompetenzen möglich wird. Zudem wurde eine repräsentative Stichprobe der Kleinklassen A (Einschulungsklassen) einbezogen. Die Schülerinnen und Schüler der Regelklassen (inklusive jene der Projekte «TaV» und «QUIMS») sowie die Schülerinnen und Schüler, die im Schuljahr 2003/04 in das erste Jahr der Kleinklasse A eintraten, gehören der Grundgesamtheit der Längsschnittstudie an [Schülerinnen und Schüler, die im Sommer 2003 im Kanton Zürich in die erste Klasse eintraten]. Es gilt zu beachten, dass die Schülerinnen und Schüler, die im Schuljahr 2003/04 in das erste Jahr der Kleinklasse A eintraten, ebenfalls am Ende der 3. Klasse und deshalb erst am Ende des Schuljahres 2006/07 untersucht werden. Die Schülerinnen und Schüler, die im Schuljahr 2003/04 in das zweite Jahr der Kleinklasse A eintraten, gehören nicht zur definierten Grundgesamtheit, weshalb sie an der Fortsetzung der Längsschnittstudie nicht mehr teilnehmen.

4.2 Klassenwechsler

Ein beachtlicher Teil der Schülerinnen und Schüler der (Ausgangs-)Stichprobe hat bis zum Ende der dritten Klasse als Folge eines Umzugs die Schule gewechselt, musste eine Klasse repetieren oder hat eine Klasse übersprungen. Vorausgesetzt, diese Schülerinnen und Schüler werden am Ende der dritten Klasse im Kanton Zürich unterrichtet, sollten sie an der Fortsetzung der Zürcher Längsschnittstudie teilnehmen. Ein Verlust dieser Schülerinnen und Schüler würde zum einen die Panelmortalität allzu stark erhöhen, zum andern die Beantwortung relevanter bildungspolitischer Fragestellungen verunmöglichen. In einem ersten Schritt muss deshalb die (Ausgangs-)Stichprobe aktualisiert werden. Dazu werden die Daten der Stichprobe der Lernstandserhebung mit der bestehenden Individualstatistik der Bildungsdirektion des Kantons Zürich verbunden. Die Schülerinnen und Schüler der (Ausgangs-)Stichprobe werden in der Individualstatistik der Bildungsdirektion gesucht. Weil die Stichprobe für die Lernstandserhebung zu Beginn der ersten Klasse über eine Klassenliste gebildet werden musste – die Schülerinnen und Schüler wurden erst nach Schuleintritt in die Individualstatistik der Bildungsdirektion aufgenommen –, wird die Verbindung über den Namen der Schülerinnen und Schüler erfolgen. Dieser Arbeitsschritt erfordert eine enge Zusammenarbeit mit der Bildungsstatistik der Bildungsdirektion des Kantons Zürich.

4.3 Erhebungsinstrumente

Leistungstests

Das Ziel, möglichst alle Schülerinnen und Schüler der Stichprobe der Lernstandserhebung zu Beginn der ersten Klasse bis zum Ende der dritten Klasse einzubeziehen, kann aus ökonomischen Gründen nur dann erreicht werden, wenn die *Leistungstests* in Mathematik und Deutsch in den Klassen von den Lehrpersonen selbstständig durchgeführt werden können. Die Erhebung der fachlichen Leistungen dauert pro Fach zwei Lektionen (je rund 90 Minuten inklusive standardisierte Anleitung zur Durchführung der Tests).

Unterrichts- und Selbstwahrnehmung der Schülerinnen und Schüler

Bezüglich der Unterrichts- und Selbstwahrnehmung durch die Schülerinnen und Schüler findet eine *Befragung* durch Testleiterinnen und Testleiter im Klassenverband statt. Für dieses Vorgehen spricht, dass insbesondere die Fragen zur Unterrichtswahrnehmung nach Möglichkeit nicht von der Lehrperson selber erhoben werden sollten. Im Bereich Selbstwahrnehmung werden neben den bereits erwähnten Merkmalen (siehe Tabelle 3.2) auch Fragen zur (Schul-)Zufriedenheit der Schülerinnen und Schüler aus der früheren Drittklasserhebung verwendet (vgl. 3. Klasse Primarschule, Fragebogen für Schülerinnen und Schüler, Juni 2002 / KBL, S. 13). Hinsichtlich der Unterrichtswahrnehmung wird auf Instrumente und Dimensionen zurückgegriffen, die schon im Rahmen von SEMOS verwendet wurden (wahrgenommene Über- bzw. Unterforderung, wahrgenommene Motivierung des Unterrichts sowie wahrgenommene Unruhe im Unterricht). Schülerinnen und Schüler, die während der ersten drei Jahre die Klasse gewechselt haben und sich nicht mehr in einer Klasse der Ausgangsstichprobe befinden, werden allerdings nicht durch Testleiterinnen und Testleiter, sondern durch ihre aktuelle Lehrperson getestet. In diesen Fällen soll nur die Selbstwahrnehmung, nicht jedoch die Unterrichtswahrnehmung erfragt werden. Die Erhebung der Unterrichts- und Selbstwahrnehmung dauert rund zwei Lektionen pro Halbklassen.

Befragung der Lehrpersonen

Die Erhebung von allgemeinen Unterrichtsmerkmalen geschieht aus ökonomischen Gründen primär durch eine schriftliche Befragung der Lehrpersonen. Es kann hier auch auf Instrumente zurückgegriffen werden, die im Rahmen von SEMOS verwendet wurden (z.B. Klarheit/Strukturiertheit, Individualisierung). Daneben werden die Lehrpersonen auch zu ihrer Arbeitssituation befragt. Im Fordergrund stehen hier Fragen zu Arbeitsstrukturen (z.B. funktionale Ausdifferenzierungen, Ressourcen), zur Zusammenarbeit im Kollegium (z.B. Austausch von Unterrichtsmaterialien, gemeinsames Vorbereiten des Unterrichts), zu Bestrebungen zur Qualitätsentwicklung (z.B. gemeinsame pädagogische Schwerpunktsetzungen, regelmässige Teamsitzungen mit pädagogischen Themen), und zur Schulkultur (z.B. Umgang mit Konflikten, gemeinsames pädagogisches Verständnis). Die definitive Auswahl in beiden Bereichen hängt vor allem auch von der zur Verfügung stehenden Zeit ab. Zusätzlich soll bei den Lehrpersonen eine Fremdeinschätzung zu jeder Schülerin und zu jedem Schüler hinsichtlich insbesondere deren Leistungsfähigkeit in Mathematik und Spra-

che erfolgen. Erfragt werden auch die sonderpädagogischen Massnahmen und zwar so, dass sie Antworten mit den Daten der Lernstandserhebung in den ersten Klassen, der 2001 durchgeführten Erhebung der sonderpädagogischen Massnahmen in den dritten Klassen und der Erhebung des sonderpädagogischen und unterrichtsergänzenden Angebots der Bildungsdirektion des Kantons Zürich (Piloterhebung 2005) möglichst vergleichbar sind. Es werden nur jene Lehrpersonen befragt, die eine Klasse der Ausgangsstichprobe unterrichten.

Befragung der Eltern

Die Eltern werden mit Hilfe eines kurzen *Fragebogens* zu ihrer Wahrnehmung der Schule und ihren Bildungsaspirationen sowie (dort wo diese Angaben fehlen) zum sozioökonomischen Hintergrund der Familie befragt.

4.4 Untersuchungsablauf in den Klassen

Die Untersuchungen in den Klassen beanspruchen drei Vormittage (vorzugsweise Dienstag, Mittwoch oder Donnerstag). Die zeitliche Aufteilung der verschiedenen Untersuchungsteile ist in Tabelle 4.1 enthalten.

Tabelle 4.1: Zeitliche Planung der Untersuchungen in den Schulen

1. Vormittag		2. Vormittag	3. Vormittag	Zeit in Minuten
Einführung		Einführung	Einführung	5
Halbklasse 1: Selbstwahrnehmung Unterrichtswahrnehmung	<i>Während eine Testleiterin die Befragung der Schülerinnen und Schüler durchführt, führt eine zweite Testleiterin mit der Lehrperson das strukturierte Interview durch.</i>	Mathematik 1. Teil	Deutsch 1. Teil	20
		Pause	Pause	5
		Mathematik 2. Teil	Deutsch 2. Teil	20
		Pause	Pause	15
Halbklasse 2: Selbstwahrnehmung Unterrichtswahrnehmung		Mathematik 3. Teil	Deutsch 3. Teil	20
		Pause	Pause	5
		Mathematik 4. Teil	Deutsch 4. Teil	20
				110

Die Testleiterinnen und Testleiter führen die Erhebung der Selbst- und Unterrichtswahrnehmung am ersten Vormittag durch. Pro Klasse werden zwei Testleiterinnen und Testleiter eingesetzt. Während eine Testleiterin die Fragebogen ausfüllt, führt eine zweite Testleiterin mit der Lehrperson das strukturierte Interview durch. Zugleich überreichen die Testleiterinnen und Testleiter den Lehrpersonen die Tests zur Erfassung der Leistungen in Ma-

thematik und Deutsch sowie die Fragebögen für die Lehrpersonen und für die Eltern. Die Fragebögen für die Eltern werden von den Lehrpersonen den Schülerinnen und Schülern abgegeben. Die Schülerinnen und Schüler bringen den Fragebogen zum Ausfüllen und Zurückschicken nach Hause.

4.5 Skalierung

Skalierung der Leistungsdaten

Die geplante Untersuchung in den dritten Klassen des Kantons Zürich soll Auskunft darüber geben, inwieweit die Schülerinnen und Schüler die Ziele des Lehrplans erreicht haben. Um dieses Ziel erreichen zu können, müssen die Tests entsprechend den Anforderungen des Lehrplans entwickelt und die Daten auf der Basis der probabilistischen Testtheorie (Item Response-Theory) skaliert werden. Die Anwendung dieser Testtheorie erlaubt es, die Testergebnisse in Bezug zu einer expliziten Anforderung (Lehrplanziel, Kompetenzanforderung, Standard) darzustellen. Das methodische Vorgehen bei der Skalierung von Leistungsdaten nach der Item Response-Theory und bei der Bestimmung von Kompetenzniveaus ist in der Publikation zur Lernstandserhebung in den ersten Klassen dargestellt (Moser, Berweger & Stamm, 2005).

Die Anwendung der Item-Response-Theory hat zwei weitere Vorteile. Zum einen wird es möglich sein, die Fachbereiche Mathematik und Deutsch in ihrem ganzen Umfang zu prüfen, in dem ein Multi-Matrix-Design eingesetzt wird. Das heisst, dass die Schülerinnen und Schüler unterschiedliche Testaufgaben bearbeiten, die allerdings der gleichen Skala angehören. Ein solches Design wird beispielsweise im internationalen Schulleistungsvergleich PISA eingesetzt. Zum andern können die Ergebnisse in den Leistungstests anhand bestehender Skalen dargestellt werden, vorausgesetzt, es die Tests zweier Untersuchungen haben eine genügend grosse Anzahl exakt gleicher Testaufgaben.

In Abschnitt 3.1 wurde darauf hingewiesen, dass die Leistungstests der geplanten Untersuchung so entwickelt werden können, dass einerseits eine Verbindung der Ergebnisse mit den Ergebnissen der Lernstandserhebungen beim Schuleintritt (Moser, Stamm & Hollenweger, 2005), andererseits eine Verbindung der Ergebnisse mit den Ergebnissen der Lernstandserhebung am Ende der dritten Klasse im Jahr 2001 (Moser, Keller & Tresch, 2003) grundsätzlich möglich ist. Eine Verbindung verlangt, dass ein Anteil an Testaufgaben in beiden Untersuchungen eingesetzt wird.

Eine Verbindung der geplanten Lernstandserhebung mit der Lernstandserhebung beim Schuleintritt ist allerdings nur beschränkt möglich, weil davon auszugehen ist, dass nahezu alle Kinder am Ende der dritten Klasse beispielsweise die Buchstaben und die Zahlen kennen. Etwas anders sieht es aus bei den Wortschatz- und Arithmetikaufgaben. Ob sich allerdings eine Verbindung der beiden Tests und dadurch die Beschreibung der Ergebnisse auf einer gleichen Skala (sozusagen mit einem gleichen Massstab) für die Bereiche Wortschatz und Mathematik möglich ist, muss vorgängig mit dem Pilottest abgeklärt werden.

Eine Verbindung der geplanten Lernstandserhebung mit der Lernstandserhebung am Ende der dritten Klasse im Jahr 2001 lässt sich ohne Probleme realisieren, weil die Aufgaben für die gleiche Population entwickelt wurden. Die Itemparameter der im Jahr 2001 eingesetzten Leistungstests sind bekannt und könnten zur Bildung der gleichen Skala als so genannte «anchors» – zur Verankerung der Testaufgaben auf der bestehenden Skala – genutzt werden. Die Verbindung hätte den Vorteil, dass ein Vergleich der Ergebnisse beider Studien anhand einer gleichen Skala durchgeführt werden könnte. Dadurch könnten die Ergebnisse im Jahr 2006 direkt im Vergleich zu jenen im Jahr 2001 beurteilt werden. Der Nachteil dieses Vorgehens wäre, dass die aktuellen Entwicklungen im Projekt HarmoS nur beschränkt berücksichtigt werden könnten, weil der geplante Test auf den bestehenden abgestimmt werden müsste.

Wie die Leistungstests entwickelt, zusammengestellt und skaliert werden, sollte zum einen entsprechend den Wünschen der Bildungsdirektion und in Absprache mit dem Projektbeirat bestimmt werden. Andererseits müsste die Machbarkeit in einem Pilottest überprüft werden.

Skalierung der Befragungsdaten

Im Vordergrund steht hier eine Skalierung auf der Basis der klassischen Testtheorie. Angesichts der Grösse der Stichprobe bei Schülerinnen und Schülern sowie bei Eltern soll hinsichtlich der Frage der faktoriellen Validität der Konstrukte wie schon in der LeZH1 und in SEMOS wiederum mit einer Explorations- und einer Validierungsstichprobe gearbeitet werden. Im ersten Falle werden explorative, im zweiten basierend auf diesen, konfirmatorische Faktorenanalysen durchgeführt. Hinsichtlich sämtlicher Angaben der Lehrpersonen sind angesichts der Grösse des Samples lediglich explorative Analysen geplant. Die definitive Entscheidung hängt hier u. a. von der theoretischen Fundierung der Konstrukte wie auch von der Komplexität der Modelle ab.

4.6 Ergebnismrückmeldung für die Lehrpersonen

Den beteiligten Klassen wird eine ausführliche Rückmeldung der Ergebnisse zugestellt, die einerseits einen Vergleich mit Parametern (Mittelwerte und Verteilungen) des Kantons Zürich zulässt, andererseits die Testleistungen aller Schülerinnen und Schüler der Klasse aufzeigt. Die Ergebnisse werden differenziert nach Lehrplanbereichen beziehungsweise Kompetenz ausgewiesen. Zudem werden die Ergebnisse im Vergleich zu den Ausgangsleistungen dargestellt, so dass die Entwicklung der Klasse sowie der einzelnen Schülerinnen und Schüler sichtbar wird. Die Klassenergebnisse werden als Nettoeffekte, also nach statistischer Kontrolle von Intelligenz, sozialer Herkunft und Immigrationsstatus dargestellt. Für die Berechnung von Klassenergebnissen werden nur jene Schülerinnen und Schüler berücksichtigt, die bereits zum zweiten Mal am Zürcher Längsschnitt teilgenommen haben. Die Ergebnismrückmeldung enthält eine ausführliche Beschreibung, wie die Ergebnisse zu interpretieren sind. Die Ergebnisse werden einzig den Klassenlehrpersonen zugestellt.

4.7 Auswertung von (Längsschnitt-)Daten aus Cluster-Stichproben

Die Auswertung der Daten erfolgt entsprechend der Stichprobe beziehungsweise der hierarchischen Struktur des Bildungswesens. Angewendet werden Mehrebenenmodelle beziehungsweise die hierarchische Analyse von Cluster-Daten, bei der die Abhängigkeit der Individuen innerhalb einer Klasse berücksichtigt und in Rechnung gestellt wird.

Die hierarchische Analyse basiert zum Teil auf den gleichen Annahmen wie die Regressionsanalyse. Der Unterschied besteht darin, dass innerhalb jeder Einheit eine eigene Regressionsgleichung formuliert wird. Bei der Längsschnittstudie werden drei Einheiten unterschieden: Zeitpunkt, Individuum und Klasse. Die Konstante und die Steigung der Regressionsgerade können zwischen den Zeitpunkten oder zwischen den Klassen variieren und werden deshalb als Zufallsvariablen angesehen. Die Unterschiede zwischen den Zeitpunkten werden mit Merkmalen der Schülerinnen und Schüler erklärt, indem auf der Individual Ebene eine neue Regressionsgleichung formuliert wird. Die Konstanten und die Steigungen der Regressionen zwischen den einzelnen Zeitpunkten werden ihrerseits als abhängige Variablen, Klassenmerkmale als unabhängige Variablen betrachtet. Der Vorteil der hierarchischen Analyse liegt darin, dass nicht einfach die Abhängigkeit der Schülerinnen und Schüler durch den Besuch der gleichen Klasse berücksichtigt wird, sondern im Zentrum des Interesses steht. Für jede Klasse können Mittelwerte und Zusammenhangsmasse berechnet und regressionstechnisch kontrolliert werden. Dies ermöglicht zuverlässige Schätzungen von Klassenmittelwerten (Intercept, Effektivität) und von Zusammenhängen zwischen Merkmalen innerhalb der Klassen (Slopes, Selektivität beziehungsweise Chancengleichheit des Bildungswesens). Ein Beispiel dafür bildet der Zusammenhang zwischen sozialer Herkunft und Leistung innerhalb einer Klasse.

Insbesondere im Zusammenhang mit der Frage von Wechselwirkungen zwischen fachlichen und überfachlichen Kompetenzen im Längsschnitt (LeZH1 – LeZH03) wird mit Strukturgleichungsmodellen gearbeitet, wobei – in Abhängigkeit von der Komplexität der Modelle – mehrere Varianten zur Auswahl stehen: a) Variante, die «lediglich» den Cluster-Effekt kontrolliert, b) Variante, basierend auf der pooled-within Matrix und c) two-level Modelle. Letzterem dürften jedoch auch bei rund 100 Einheiten auf Klassenebene enge Grenzen gesetzt sein.

5 Zeitplan und Zuständigkeit

LERNSTANDSERHEBUNG		
Zeitpunkt	Meilenstein	Zuständigkeit
Sept. 05	Feinkonzeption	KBL/PHZH/Helmke/BID
Nov. 05 – Feb. 06	Entwicklung der Fachleistungstests	KBL/Fachdidaktik und Lehrpersonen
Nov. 05 – Feb. 06	Bestimmung der Konstrukte und Auswahl der Instrumente für die Schüler-, Lehrer- und Elternfragebögen	PHZH
Jan. 05	Bildung der Stichprobe	KBL
Jan. – Feb. 06	Zusammenstellung und Layout der Fragebögen (Lehrpersonen, Schülerinnen und Schüler, Eltern), Entwicklung strukturiertes Interview	PHZH
März 06	Pilottest	KBL/PHZH
April 05	Arbeitsbericht 1: Instrumente und Planung der Erhebung	KBL/ PHZH
Mai 06	Anschrift der Lehrpersonen	KBL/PHZH/BID
April / Mai 06	Rekrutierung und Schulung Testleiterinnen und Testleitern	PHZH
Juni 06	Durchführung der Erhebungen in den Klassen	KBL/PHZH
Oktober 06	Klassenrückmeldung	KBL
Oktober 06	Zwischenbericht	KBL/PHZH
Oktober 07	Schlussbericht	KBL/PHZH
ZUSATZOPTION (KLASSENWECHSLER / KLEINKLASSEN)		
Zeitpunkt	Meilenstein	Zuständigkeit
Mai 07	Anschrift der Lehrpersonen für die Nacherhebung	KBL/PHZH/BID
Juni 07	Durchführung der Nacherhebungen der Schülerinnen und Schüler, die in die Kleinklassen A eingetreten sind oder eine Klasse repetiert haben	KBL/PHZH
Oktober 07	Schlussbericht	KBL/PHZH

6 Projektorganisation

Die Lernstandserhebung am Ende der dritten Klasse wird vom Kompetenzzentrum für Bildungsevaluation und Leistungsmessung (KBL), vertreten durch Dr. Urs Moser, und der Pädagogischen Hochschule Zürich (PHZH), vertreten durch Prof. Dr. Judith Hollenweger, durchgeführt. Die Leitung des Projekts liegt bei Urs Moser. Die Zuständigkeiten der beiden Institutionen sind klar geregelt, sowohl was die Projektorganisation betrifft als auch für die Bearbeitung der Fragestellungen. Das KBL ist für die Bearbeitung der Fragestellungen 1, 3 und 4 zuständig und übernimmt das Projektmanagement sowie die Redaktion der Berichterstattung. Die PHZH ist für die Bearbeitung der Fragestellungen 2, 5 und 6 zuständig und übernimmt die Erfassung der überfachlichen Kompetenzen sowie die Befragung der Lehrpersonen.

Vom KBL ist geplant, dass lic. phil. Florian Keller (Schwerpunkt Bildungssoziologie) am Projekt mitarbeiten wird. Florian Keller hat bereits das Projektmanagement bei der Evaluation der dritten Primarschulklassen übernommen. Von der Pädagogischen Hochschule Zürich sind Prof. Dr. Alex Buff und Dr. Yuka Nakamura (Schwerpunkt überfachliche Kompetenzen) sowie Dr. Regula Leemann (Schwerpunkt Arbeitssituation der Lehrpersonen in den Reformprojekten) an der geplanten Erhebung beteiligt.

Für verschiedene fachliche Beratungen sind folgende Personen vorgesehen:

- *Fachdidaktische Beratung bei der Entwicklung der Leistungstests im Fach Sprache:*
Prof. Dr. Thomas Lindauer, Fachhochschule Aargau, Zentrum Lesen (angefragt)
- *Fachdidaktische Beratung bei der Entwicklung der Leistungstests im Fach Mathematik:*
Prof. Dr. Beat Wälti, Fachhochschule Aargau, Dozent für Fachdidaktik Mathematik (angefragt)
- *Methodische Beratung bei der Skalierung von Leistungsdaten und der Auswertung von Längsschnittdaten*
Prof. Dr. Ray Adams, Australian Council of Educational Research
- *Methodische Beratung bei der Entwicklung von Befragungsinstrumenten und der Konzeption der Längsschnittstudie*
Prof. Dr. Andreas Helmke, Universität Landau (angefragt)

7 Literatur

- Adams, R. (2002). Scaling PISA Cognitive Data. In R. Adams & M. Wu (Eds.), *PISA 2000 Technical Report* (pp. 99–108). Paris: Organisation for Economic Co-operation and Development.
- Ainley, M. D. (1993). Styles of engagement with learning: Multidimensional assessment of their relationship with strategy use and school achievement. *Journal of Educational Psychology, 85*(3), 395-405.
- Anderman, L.H., Freeman, T.M. (2004). Students' sense of belonging in school. In P. Pintrich, & M.L. Maehr (Eds.) *Motivating students, improving schools: the Legacy of Carol Midgley*, Elsevier, 27-63.
- Artelt, C. & Schellhas, B. (1996). *Zum Verhältnis von Strategiewissen und Strategieanwendung und ihren kognitiven und emotional-motivationalen Bedingungen im Schulalter*. Interdisziplinäres Zentrum für Lern- und Lehrforschung. Universität Potsdam.
- Bandura, A. (2003). *Self-efficacy: The Exercise of Control* (7nd ed.). New York: Freeman and Company.
- Becker, G.S. (1975). *Human Capital*. New York: Columbia University Press.
- Benkmann, R. (2003): Diagnose der Bedingungen: Sozialisation bei Lernbeeinträchtigungen. In: Ricken, G.; Fritz, A.; Hofmann, Ch. (Hrsg.): *Diagnose: Sonderpädagogischer Förderbedarf*. Lengerich, Berlin u.a.: Pabst Science Publishers, 163-173
- Boekaerts, M. (1999). Self-regulated learning: Where are we today. *International Journal of Educational Research, 31*, 445-458.
- Boekaerts, M. & Minnaert, A. (1999). Self-regulation with respect to informal learning. *International Journal of Educational Research, 31*, 533-544 .
- Buff, A., Reusser, K. & Pauli, C. (in Vorbereitung a). Selbstvertrauen ist wichtig, aber nicht ausreichend! Die Bedeutung von Unterricht, Selbstvertrauen, Qualität der Lernmotivation für Engagement und Leistung im Fach Mathematik. In K. Reusser, C. Pauli & M. Waldis (Hrsg.). *Unterricht und Lernen in Schweizer Schulen. Ergebnisse einer internationalen und nationalen Video-Studie zum Mathematikunterricht*. Chur: Rüegger.
- Buff, A., Reusser, K. & Pauli, C. (in Vorbereitung b). Die Qualität der Lernmotivation in Mathematik auf der Basis freier Äusserungen: Welches Bild präsentiert sich bei Deutschschweizer Schülerinnen und Schüler im 8. und 9.Schuljahr? In K. Reusser, C. Pauli & M. Waldis (Hrsg.). *Unterricht und Lernen in Schweizer Schulen. Ergebnisse einer internationalen und nationalen Video-Studie zum Mathematikunterricht*. Chur: Rüegger.
- Buff, A., Nakamura, Y., Hollenweger, J., & Achermann, E. (2004). PHZH-Längsschnittstudie: Entwicklung des Selbstkonzepts und der Motivation. Projektbeschreibung. Unveröffentlichtes Papier. Zürich: Pädagogische Hochschule Zürich.

- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39 (2), 223-238.
- Dole, J., Brown, K. and Trathen, W. (1997). The effects of strategy instruction on the comprehension performance of at-risk students. *Reading Research Quarterly*, Vol. 31, No. 1, 62-88.
- Emanuelsson, I., Persson, B. (1997). Who is considered to be in need of special education: why, how and by whom? In: *European Journal of Special Needs Education*, 12, 2, 127-136.
- Furrer, C., & Skinner, E.. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 95, 148-162 .
- Grob, U., Maag Merki, K. & Bühler, X. (2003). Young Adult Survey. Theoretische Begründung und empirische Befunde zur Validierung eines Indikatorensystems zu überfachlichen Kompetenzen. *Schweizerische Zeitschrift für Bildungswissenschaften*, 25, 309-330.
- Guay, F., Boivin, M., & Hodges, E.V.E. (1999). Predicting change in academic achievement: A model of peer experiences and self-system processes. *Journal of Educational Psychology*, 91, 105-115.
- Gutman, L.M., & Midgley, C. (2000). The role of protective factors in supporting the academic achievement of poor African American students during the middle school transition. *Journal of Youth and Adolescence*, 29, 223-248.
- Gürtler, T. (2003). *Trainingsprogramm zur Förderung selbstregulativer Kompetenz in Kombination mit Problemlösestrategien PROSEKKO*. Frankfurt a. M.: Peter Lang.
- Hattie, J., Biggs, J. & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, 66 (2), 99-136.
- Helmke, A. (1993). Die Entwicklung der Lernfreude vom Kindergarten bis zur 5. Klassenstufe. *Zeitschrift für Pädagogische Psychologie*, 7, 77-86.
- Helmke, A. (1997). Entwicklung lern- und leistungsbezogener Motive und Einstellungen: Ergebnisse aus dem SCHOLASTIK-Projekt. In In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 59-76). Weinheim: PVU.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule. Enzyklopädie der Psychologie, Bd. 3* (S. 71-176). Göttingen: Hogrefe.
- Klauer, K. J. (2003). Über internationale Entwicklungstendenzen sonderpädagogischer Diagnostik unter förderdiagnostischem Aspekt. In: Ricken, G.; Fritz, A.; Hofmann, Ch. (Hrsg.): *Diagnose: Sonderpädagogischer Förderbedarf*. Lengerich, Berlin u.a.: Pabst Science Publishers, 127-142.
- Krapp, A. & Ryan, R. N. (2002). Selbstwirksamkeit und Lernmotivation. In M. Jerusalem, & D. Hopf (Hrsg.), *Zeitschrift für Pädagogik. Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen. 44. Beiheft*, 54-82.

- Kronig, W., Haeberlin, U. & Eckhart, M. (2000). Immigrantenkinder und schulische Selektion. Pädagogische Visionen, theoretische Erklärungen und empirische Untersuchungen zur Wirkung integrierender und separierender Schulformen in den Grundschuljahren. Bern, Stuttgart, Wien: Haupt.
- Langfeldt, H.-P. (1998): Behinderte Kinder im Urteil ihrer Lehrkräfte: eine Analyse der Begutachterpraxis im Sonderschul-Aufnahme-Verfahren. Heidelberg: Universitätsverlag C. Winter.
- Maag Merki, K. & Grob, U. (2003). Überfachliche Kompetenzen: Zur Validierung eines Indikatorensystems. *Empirische Pädagogik*, 17, 123-147.
- Mand, J. (2002): Sonderschule oder Gemeinsamer Unterricht? Zum Einfluss von Gutachtervariablen auf Schullaufbahnentscheidungen für schulschwache und auffällige Kinder und Jugendliche. In: Zeitschrift für Heilpädagogik 53, 8-13.
- Marsh, H.W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2004). *Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering*. Paper presented at the 3th International Biennial SELF Research Conference, in Berlin, Germany - July 4 - 7, 2004.
- Mayring, P. (1999). Qualitativ orientierte Forschungsmethoden in der Unterrichtswissenschaft: Ein Anwendungsbeispiel aus der Lernstrategieforschung. *Unterrichtsforschung*, 27 (4), 292-309.
- Moser, U. (2005). Lernvoraussetzungen in Schulklassen zu Beginn der 1. Klasse. In U. Moser, M. Stamm und J. Hollenweger (Hrsg.), Für die Schule bereit? Lesen, Wortschatz Mathematik und soziale Kompetenzen beim Schuleintritt (S. 167–185). Oberentfelden: Sauerländer
- Moser, U., Berweger, S. und Stamm, M. (2005). Lesekompetenzen bei Schuleintritt. In U. Moser, M. Stamm und J. Hollenweger (Hrsg.), Für die Schule bereit? Lesen, Wortschatz Mathematik und soziale Kompetenzen beim Schuleintritt (S. 37–58). Oberentfelden: Sauerländer.
- Moser, U., Stamm, M. & Hollenweger, J. (Hrsg.). Für die Schule bereit? Lesen, Wortschatz Mathematik und soziale Kompetenzen beim Schuleintritt. Oberentfelden: Sauerländer.
- Moser, U., Keller, F. & Tresch, S. (2003). Schullaufbahn und Leistung. Bildungsverlauf von Zürcher Schülerinnen und Schülern am Ende der 3. Volksschulklasse: Bern: hep.
- Nakamura, Y.M., Leemann, R.J. und Hollenweger, J. (2005): Lebensqualität und berufliche Weiterbildung von Lehrpersonen. Die Zürcher MAB als Instrument der Qualitätsentwicklung. In: Sigrist, M., Wehner, T. und Legler, A. (Hrsg.): Schule als Arbeitsplatz. Mitarbeiterbeurteilung zwischen Absicht, Leistungsfähigkeit und Akzeptanz: S. 191-209.
- Organisation for Economic Co-Operation and Development (2001). *Lernen für das Leben. Erste Ergebnisse von PISA 2000*. Paris: Organisation for Economic Co-operation and Development.
- Organisation for Economic Co-Operation and Development (2001). *Lernen für das Leben. Erste Ergebnisse von PISA 2000*. Paris: Organisation for Economic Co-operation and Development.

- Organisation for Economic Co-Operation and Development (2003). Education Policy Analysis. Paris: Organisation for Economic Co-Operation and Development.
- Organisation for Economic Co-Operation and Development (2004). Equity in Education. Students with Disabilities, Learning Difficulties and Disadvantages. Paris: Organisation for Economic Co-Operation and Development.
- Patrick, H., Ryan, A. M. & Pintrich, P. R. (1999). The differential impact of extrinsic and mastery goal orientations on males' and females' self-regulated learning. *Learning and Individual Differences*, 11 (2), 153-171.
- Pietsch, J., Walker, R. & Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology*, 3, 589-603.
- Pintrich, P. R. & Garcia, T. (1993). Intraindividual differences in students' motivation and self-regulated learning. *Zeitschrift für Pädagogische Psychologie*, 7 (2/3), 99-107.
- Pokay, P. & Blumenfeld, P. C. (1990). Predicting achievement early and late in the semester: The role of motivation and use of learning strategies. *Journal of Educational Psychology*, 82, 41-50.
- Prenzel, M., Kramer, K. & Drechsel, B. (2001). Selbstbestimmt motiviertes und interessiertes Lernen in der kaufmännischen Erstausbildung – Ergebnisse eines Forschungsprojekts. In K. Beck & V. Krumm (Hrsg.), *Lehren und Lernen in der beruflichen Erstausbildung. Grundlagen einer modernen kaufmännischen Berufsqualifizierung* (S. 37-61). Opladen: Leske + Budrich.
- Purdie, N. & Hattie, J. (1996). Cultural differences in the use of strategies for self-regulated learning. *American Educational Research Journal*, 33 (4), 845-871.
- Ramseier, E. (2004). *Motivation als Ergebnis und als Determinante schulischen Lernens*. Dissertation, Universität Zürich.
- Reeve, J. (2002). Self-determination theory applied to educational settings. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 183-203). Rochester, NY: The University of Rochester Press.
- Rost J. (1996). *Lehrbuch Testtheorie Testkonstruktion*. Bern: Huber.
- Rüesch P. (1999). *Gute Schulen im multikulturellen Umfeld*. Orell Füssli. Zürich.
- Satow, L. & Schwarzer, R. (2003). Entwicklung schulischer und sozialer Selbstwirksamkeitserwartung. *Psychologie in Erziehung und Unterricht*, 50, 168-181.
- Schiefele, U., Wild, K.-P. & Winteler, A. (1995). Lernaufwand und Elaborationsstrategien als Mediatoren der Beziehung von Studieninteresse und Studienleistung. *Zeitschrift für Pädagogische Psychologie*, 9(3/4), 181-188.
- Sträuli, B. (2000). Sprache und Schulentwicklung. In: Lehrmittelverlag des Kantons Zürich (Hrsg.). *Schulerfolg: kein Zufall. Ein Ideenbuch zur Schulentwicklung im multikulturellen Umfeld*, 54-69.
- VanderStoep, S. W., Pintrich, P. R. & Fagerlin, A. (1996). Disciplinary differences in self-regulated learning in college students. *Contemporary Educational Psychology*, 21, 345-362.

- Wagner, S.J., Powell, J.W. (2003). Ethnisch-kulturelle Ungleichheit im deutschen Bildungssystem – Zur Überrepräsentanz von Migrantenjugendlichen an Sonderschulen. In: Cloerkes, G. (Hrsg.): *Wie man behindert wird. Texte zur Konstruktion einer sozialen Rolle und zur Lebenssituation betroffener Menschen*. Heidelberg: Universitätsverlag Winter, 183-208.
- Weinert, F. E. & Schrader, F.-W. (1996). Lernen lernen als psychologisches Problem. In F. E. Weinert & H. Mandl (Hrsg.), *Enzyklopädie der Psychologie* (Vol. 4, S. 296-335). Göttingen: Hogrefe.
- Wigfield, A. & Eccles, J.S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68-81.
- Winne, P. H. & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology*, 27, 551-572.
- Winne, P. H. & Perry, N. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531-566). San Diego: Academic Press.
- Wuttke, E. (2000). Lernstrategien im Lernprozess. *Zeitschrift für Erziehungswissenschaft*, 3(1), 97-110.
- Zimmerman, B. J. & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23, 614-628.
- Zimmerman, B. J. & Martinez-Pons, M. (1988). Construct validation of a strategy model of students self-regulated learning. *Journal of Educational Psychology*, 80(3), 284-290.
- Zimmerman, B. J. & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82(1), 51-59.