

14. Effekte der NMS-Konzeptmerkmale auf die fachlichen Schülerleistungen

Christoph Helm, David Kemethofer, Herbert Altrichter & Christoph Weber

1. Fragestellungen

Haben die speziellen Konzeptmerkmale der Neuen Mittelschule (siehe Kapitel 2 und 3) einen Einfluss auf die fachlichen Leistungen der Schüler/innen? Auf diese Leitfrage versucht der vorliegende Beitrag Antworten zu geben.¹ Fachliche Leistungen stellen zentrale Ziele von Schule und Unterricht dar (Weinert, 1998). Daher wird in diesem Beitrag untersucht, ob die fachlichen Schülerleistungen durch

- 1) die Lernumgebung entsprechend den Konzeptmerkmalen der NMS (auf Klassenebene aggregierte Angaben der Schüler/innen),
- 2) die Intensität der Umsetzung dieser Konzeptmerkmale in den jeweiligen Schulklassen (Angaben der Klassenvorstände) und
- 3) die Intensität der Umsetzung dieser Konzeptmerkmale in den jeweiligen Schulen (Angaben der Schulleitungen) vorhergesagt werden können.

Die fachliche Leistung wird in den Fächern Deutsch, Mathematik und Englisch durch Leistungsmessungen in den Jahren 2012 und 2013 erhoben (vgl. Kapitel 5). Damit etwaige Effekte den Konzeptmerkmalen der NMS zugerechnet werden können, wird für zentrale Schülermerkmale (wie z. B. deren kognitive Leistungsfähigkeit und sozialer Hintergrund) kontrolliert.

Die im Titel dargestellte Fragestellung geht von der Annahme direkter Effekte der verschiedenen Ebenen (Schule, Klasse, Schüler/innen) auf die Schülerleistungen aus. Es ist jedoch anzunehmen, dass die NMS-Konzeptmerkmale vor allem indirekt über unterschiedliche Mediatoren auf die Schülerleistungen wirken. Konsequenterweise wird in diesem Beitrag somit neben der Frage nach den direkten Effekten der NMS-Konzeptmerkmale auch der Frage nach möglichen indirekten Effekten über Variablen, die sich als Mediatoren der Wirkung von Konzeptmerkmalen der NMS identifizieren lassen, nachgegangen.

Zur adäquaten Beantwortung der Forschungsfrage wird zunächst ein Mehrebenen-Regressionsmodell zur Schätzung der direkten Effekte der NMS-Konzeptmerkmale auf die fachliche Leistung berechnet. In weiterer Folge dient ein Mehrebenen-Pfadmodell zur Prüfung indirekter Effekte (Hox, 2013). Die Ergebnisse sowie die aus ihnen abgeleiteten Handlungsimplicationen werden vor dem Hintergrund der Grenzen der Studie diskutiert.

¹ In diesem Kapitel werden die Leistungsunterschiede von Schüler/inne/n durch NMS-Konzeptmerkmale erklärt, während in Kapitel 12 auf Schulebene Leistungsänderungen und deren Prädiktoren im Zentrum des Interesses stehen. Diese unterschiedliche Analyseperspektive ist Grund für teils unterschiedliche Ergebnisse.

2. Theoretische Ausgangsüberlegungen

Da die „Implementationslogiken und -erwartungen“ der Einführung der Neuen Mittelschule an anderer Stelle in diesem Band ausführlich dargestellt sind, wird in diesem Beitrag auf eine Wiederholung verzichtet und der/die Leser/in auf Kapitel 2 verwiesen. Die konkret vermuteten Wirkungszusammenhänge zwischen den NMS-Konzept- und Schülermerkmalen werden im Folgenden, im Anschluss an eine Beschreibung der NMS-Konzeptmerkmale, dargestellt.

2.1 Implementierung von NMS-Konzeptmerkmalen

Jene Konzeptelemente der NMS, die aus den vom BIFIE durchgeführten Datenerhebungen abgeleitet werden konnten, sind aus theoretischer Sicht auf der Schul- und Klassenebene zu verorten. Operationalisierung, deskriptive Statistiken sowie Reliabilitätskennzahlen der einzelnen Skalen sind im Online-Dokument (Kemethofer & Helm, 2015) nachzulesen.

Auf Schulebene (Angaben der Schulleitung): Strukturelle Veränderungen

Ein Merkmal des NMS-Konzepts ist, dass neben PH-ausgebildeten Lehrpersonen auch solche mit einer universitären Lehrerausbildung unterrichten. Die Realisierung dieses Konzeptmerkmals wird durch die Variable „relativer Anteil der AHS-Lehrer/innen“ gemessen. Darüber hinaus wurde das Vorhandensein folgender NMS-Konzeptmerkmale erfasst: (1) Jahrgangskordinator/in, (2) Nachmittagsbetreuung, (3) Schülerpartizipation, (4) individuelle Fördervereinbarungen für Schüler/innen. Weiters geben zwei Indizes Auskunft über das Ausmaß, in dem (1) die Lehrer/innen frei über die Zusammensetzung beim Teamteaching entscheiden können und (2) die Schule an Wettbewerben teilnimmt.

Auf Klassenebene (Angaben der Klassenvorstände und Schüler/innen): Veränderungen der Unterrichtsprozesse

Das Konzeptmerkmal *Individualisierung und Differenzierung* wurde durch zwei Faktoren abgebildet, die sich entsprechend den Itemformulierungen als „selbstreguliertes Lernen“ und „Schülerorientierung“ betiteln lassen. Der Faktor zum selbstregulierten Lernen ist ein Indikator für das Ausmaß an Entscheidungsspielräumen, die Schüler/innen beim Lernen im Unterricht vorfinden, z. B. im Hinblick auf die Lernmaterialien, die Zeiteinteilung, die Aufgabenwahl, das Lerntempo und die eigenen Interessen etc. Der Faktor zur Schülerorientierung fokussiert stärker auf die Einbeziehung der Schüler/innen in den Unterricht, z. B. indem nach Schülerfeedback gefragt wird. Weiters wurde *innere Differenzierung* im Fach allgemein und im Rahmen der Leistungsbeurteilung durch Items erfasst. *Teamteaching* wurde durch drei Indikatoren erfasst, wobei diese sowohl die Quantität des Teamteachings (im Verhältnis zu den Wochenstunden je Unterrichtsfach) als auch zwei Aspekte der Qualität des Teamteachings (a. verschränktes Teamteaching; b. Teamteaching, bei dem die Klassen in Gruppen geteilt werden, die je von einer Lehrkraft unterrichtet werden) messen. Zusätzlich wurden das Ausmaß des *Einsatzes alternativer Leistungsrückmeldungen* und das *Vorhandensein von Kind-Eltern-Lehrer/in-Gesprächen* erfasst.

Weiters wurden die Skalen „Anomie“, „Fördernder Unterricht“, „Überfordernder Unterricht“, „Klassenführung“, „Wahlmöglichkeiten (fachspezifisch)“ und „Individuelle Lernunterstützung (fachspezifisch)“ (vgl. Kapitel 7 und 8 in diesem Band) als Prädiktoren der fachlichen Leistung auf Klassenebene berücksichtigt. Dabei handelt es sich um individuelle Schülereinschätzungen, die auf Klassenebene aggregiert wurden und damit Unterrichtsmerkmale darstellen, die über die geteilte Wahrnehmung der Schüler/innen

Auskunft geben. Zwar handelt es sich hier um Urteile über Unterricht und nicht um Konzeptmerkmale der NMS, allerdings würde man erwarten, dass die „neue Lernkultur“ der NMS mit positiven Ausprägungen dieser Unterrichtsmerkmale in Zusammenhang steht.

Die zu erklärenden Variablen stellen die *Schülerleistungen* am Ende der 8. Schulstufe dar (vgl. Kapitel 5).

2.2 Wirkungsmodell

Die Evaluierung des NMS-Konzepts bringt die Herausforderung mit sich, konzeptgeleitet Wirkungsmechanismen zu identifizieren und den Konzeptelementen zuzuordnen. Herausfordernd deshalb, weil die zu prüfende Intervention nicht hinreichend präzise bestimmt ist. Andererseits hängen erhoffte Effekte von Konzeptmerkmalen nicht nur vom Vorhandensein bestimmter Konzeptelemente ab, sondern auch von der Qualität ihrer Umsetzung. Was eine qualitätsvolle Realisierung bestimmter Konzepte ausmacht, ist aber (v. a. fachunabhängig) schwierig zu bestimmen. Auch im Fall der NMS-Reform liegen kaum theoretisch ausformulierte Beziehungen zwischen Konzeptmerkmalen und Zielvariablen vor (siehe Kapitel 2). In der vorliegenden Untersuchung werden daher zwei Analyseschritte durchgeführt, um die Forschungsfragen zu beantworten.

Schritt 1: In einem explorativ angelegten Mehrebenen-Regressionsmodell werden auf den entsprechenden Ebenen die unter Abschnitt 2.1 beschriebenen NMS-Konzeptmerkmale aufgenommen. Von ihnen werden durchwegs positive, wenn auch aufgrund ihrer distalen Stellung zu den Schülerleistungen nur schwache Effekte auf die fachlichen Leistungen vermutet (siehe Abbildung 1).

Schritt 2: In einem konfirmatorischen Mehrebenen-Pfadmodell werden die indirekten Effekte ausgewählter NMS-Konzeptmerkmale untersucht. Auf Schulebene fiel die Auswahl auf (1) den Anteil an AHS-Lehrer/innen, (2) die Nachmittagsbetreuung und (3) den individuellen Förderplan sowie (4) die selbstständige Zusammensetzung der Lehrerteams beim Teamteaching, weil sie einen besonderen Stellenwert im pädagogischen Konzept der NMS haben (BMBF, 2014). Merkmale von denen anzunehmen ist, dass sie auch an den seinerzeitigen Hauptschulen forciert wurden (Schülerpartizipation, Teilnahme der Schule an verschiedenen Wettbewerben) wurden nicht ins Mehrebenen-Pfadmodell aufgenommen. Auch die Jahrgangsteamkoordination wurde nicht berücksichtigt, da sie vor allem nichtkognitive Aspekte des Unterrichts („beständige, verlässliche und belastbare menschliche Beziehungen“ zwischen Lehrer/innen und Schüler/innen; Expertenkommission, 2007, S. 27) unterstützen soll.

Auf Klassenebene wurden im finalen Mehrebenen-Pfadmodell die drei Treatment-Variablen der Klassenvorstandsbefragung Ausmaß *Teamteaching*, *Schülerorientierung* und *selbstreguliertes Lernen* aufgenommen. Die Pfade spiegeln die Annahme wider, dass der Anteil an AHS-Lehrer/innen, die freie Zuteilung zu Lehrerteams für das Teamteaching und das Ausmaß an Teamteaching an einer Schule die Schülerorientierung und die Gewährung selbstgesteuerter Lernmöglichkeiten durch die Lehrpersonen steigern. Dies wiederum sollte zu einer erhöhten Wahrnehmung eines individualisierten und differenzierten Unterrichts durch die Schüler/innen führen, was förderlich für die fachliche Leistungserbringung einer Klasse (kontrolliert für die kognitiven Fähigkeiten, den Sozialstatus und das Geschlecht der Schüler/innen) ist. Auch für die Strukturelemente *Nachmittagsbetreuung* und *Individuelle Förderpläne für Schüler/innen*, die – im ersten Fall – mehr Zeit für Individualisierung und – im zweiten Fall – ein systematischeres Konzept von Individualisierung signalisieren, wird jeweils ein positiver Effekt auf die fachlichen Schülerleistungen (je Schule) vermutet.

Aufgrund der großteils fachunabhängigen Formulierung der eingesetzten Items/Skalen wird davon ausgegangen, dass das in Abbildung 1 dargestellte Modell in gleicher Weise für alle drei Unterrichtsfächer (Deutsch, Mathematik und Englisch) Gültigkeit besitzt.

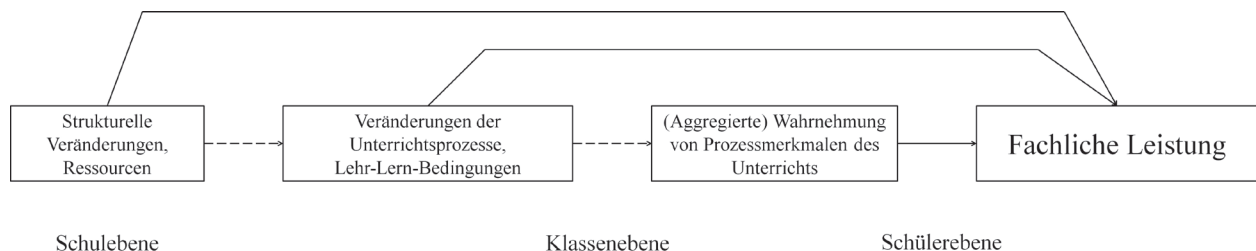


Abbildung 1: Angenommene Wirkungsmodelle direkter und indirekter Effekte der NMS-Konzeptmerkmale auf die Schülerleistungen in den Fächern Deutsch, Mathematik und Englisch

Anmerkung:

Abbildung 1 spiegelt die angenommenen Wirkungen der NMS-Konzeptmerkmale wider, die im explorativen Mehrebenen-Regres-sionsmodell (nur durchgehende Pfeile) und im konfirmatorischen Mehrebenen-Pfadmodell (alle Pfeile) geprüft werden.

3. Daten und methodisches Vorgehen

3.1 Datenbasis

Für die Beantwortung der Forschungsfragen stehen Daten für beide NMS-Generationen zur Verfügung (vgl. Kapitel 5). Im vorliegenden Kapitel werden jeweils die Schlusserhebungen, d. h. die Nachher-Messungen (G1: 2012, G2: 2013) herangezogen. Nach multipler Datenimputation (vgl. Weber, 2015) liegen für G1 Daten von insgesamt 3371 Schüler/inne/n aus 171 Klassen in 66 NMS vor. Für G2 liegen nach multipler Imputation Daten von insgesamt 6026 Schüler/inne/n aus 301 Klassen in 104 NMS vor.²

3.2 Datenanalyse

Um die in den vorangegangenen Abschnitten dieses Kapitels dargestellten Wirkungsannahmen zu prüfen, wurden die Daten in Dreiebenen-Regressionsmodelle und Dreiebenen-Pfadmodelle (Schul-, Klassen- und Schülerebene) – jeweils für die drei Fächer Deutsch, Mathematik und Englisch sowie für G1 und G2 – überführt. Dies ermöglichte eine korrekte Behandlung der Standardfehler und eine Überprüfung der Effekte auf den unterschiedlichen Ebenen. Es wurde jeweils ausschließlich mit manifesten Variablen gerechnet. Die Analysen wurden in Mplus 7.2 (Muthén & Muthén, 1998–2014) durchgeführt. Vorab wurden die verwendeten Prädiktorvariablen durch Inspektion der Korrelationsmatrix – vgl. Kemethofer und Helm (2015) – sowie durch Berechnung von Toleranz und Varianzinflationsfaktor auf Hinweise für Multikollinearität geprüft. Die Ergebnisse deuten darauf hin, dass Multikollinearität ausgeschlossen werden kann. Alle Kontrollvariablen der Schülerebene wurden am Gesamtmittelwert („grandmean“) zentriert (Enders & Tofghi, 2007). Im Ergebnisteil werden unstandardisierte Koeffizienten berichtet. Dort wo im Text Effektstärkenvergleiche zwischen Prädiktoren angeführt werden, werden zusätzlich standardisierte Koeffizienten

² Die Fallzahlen weichen von jenen in Kapitel 12 ab, da dort nur die Schulen, für die Vorher- und Nachher-Messungen verfügbar sind, in die Analyse einbezogen wurden.

berichtet. Für die Diskussion der berichteten Koeffizienten gilt, dass diese immer unter der Bedingung sonst gleicher Variablenausprägung zu interpretieren sind. Für die hier interessierenden Konzeptmerkmale auf Schul- und Klassenebene heißt das: Ein/e Schüler/in, der/die in eine Klasse/Schule geht, in der eine höhere Ausprägung eines bestimmten Unterrichtsmerkmals/NMS-Konzeptmerkmals vorliegt, erreicht eine bessere/schlechtere Leistung (je nach Vorzeichen des Koeffizienten) als ein/e Schüler/in mit den gleichen Individualmerkmalen, der/die in eine Klasse/Schule geht, in der dasselbe Unterrichtsmerkmal/NMS-Konzeptmerkmal geringer ausgeprägt ist (wenn für alle anderen angeführten Schüler- und Unterrichtsmerkmale kontrolliert wird).

4. Ergebnisse

Die Ergebnisse der explorativ angelegten Mehrebenen-Regressionsanalyse wie auch des hypothesenprüfenden Pfadmodells zeigen für alle drei Fächer, dass die NMS-Konzeptmerkmale (aus Sicht der Klassenvorstände und Schulleiter/innen) nicht direkt (mit wenigen Ausnahmen, die aber praktisch nicht relevant sind) auf die Schülerleistungen wirken. Auch indirekt sind die Effekte nur schwach und begrenzt auf die kollektive Wahrnehmung der Lernumgebung durch die Schüler/innen, welche die Schülerleistung nicht positiv oder negativ beeinflusst (mit Ausnahme der Effekte im Fach Englisch in G1).

4.1 Mehrebenen-Regressionsmodelle

Auf Schulebene

Als signifikante Konzeptmerkmale auf Schulebene ließen sich lediglich die „freie Zuteilung zum Team-teaching“ in G1 (Englisch) und die Nachmittagsbetreuung in G2 (Deutsch und Mathematik) identifizieren. Sie weisen allerdings entgegen den Erwartungen signifikant negative Effekte auf. Eine mögliche Erklärung könnte darin gesehen werden, dass in Städten (in Schulen mit niedrigerem Leistungspotenzial der Schüler/innen) häufiger Nachmittagsbetreuung angeboten wird; positive Korrelationen bestätigen diese Annahme. In beiden Generationen zeigt sich entsprechend, dass die Einwohnerzahl des Schulstandorts im Zusammenhang mit den Schülerleistungen steht. In G1 ist dieser Effekt in Mathematik signifikant negativ, in G2 in Deutsch und Englisch allerdings signifikant positiv.

Auf Klassenebene

Wie bereits erwähnt nehmen auf Klassenebene die NMS-Konzeptmerkmale keinen direkten Einfluss auf die fachlichen Schülerleistungen. Lediglich die auf Klassenebene aggregierte Schülerwahrnehmung der Unterrichtsmerkmale förrender Unterricht, Anomie, überfordernder Unterricht, Klassenführung, individuelle Lernunterstützung und Wahlmöglichkeiten wirken abhängig von Fach und Kohorte direkt auf die Leistungen. Dies deutet darauf hin, dass wie vermutet lediglich indirekte Effekte von NMS-Konzeptmerkmalen angenommen werden können.

Entgegen den Erwartungen wirkt der fördernde Unterricht *signifikant negativ* auf die Fachleistungen in Deutsch (G1) und Englisch (G2). Dieser unerwartete Effekt könnte möglicherweise im Sinne der umgekehrten Kausalität erklärbar sein: Wenn es vor allem leistungsschwächere Schulen sind, in denen Lehrpersonen häufiger angeben „fördernd“ zu unterrichten, dann ist ein solcher Effekte durchaus zu erwarten. Die Merkmale eines überfordernden und anomischen Unterrichts stehen in G2 erwartungsgemäß in fast allen Fächern im negativen Zusammenhang mit den Schülerleistungen. Der in anderen Studien vielerorts nach-

gewiesene positive Effekt einer effizienten Klassenführung bleibt dagegen in dieser Studie aus (dies kann natürlich auch auf die nicht immer idente Operationalisierung zurückzuführen sein). Nur in G1 im Fach Deutsch konnte dieser Effekt beobachtet werden. Auch die individuelle Lernunterstützung steht nur in G1 im Fach Englisch und in G2 im Fach Deutsch im positiven Zusammenhang mit der Schülerleistung. D. h., ein/e Schüler/in, der/die in eine Klasse geht, in der die Schüler/innen im Mittel eine höhere Ausprägung der individuellen Lernunterstützung durch die Lehrperson wahrnehmen, erreicht eine rund 30 bzw. 25 Punkte bessere Englisch- bzw. Deutschleistung als ein/e Schüler/in, der/die in eine Klasse geht, in der die Schüler/innen diese individuelle Lernunterstützung nicht so stark wahrnehmen. Schließlich weist das Angebot von Wahlmöglichkeiten in G2 im Fach Englisch einen negativen Effekt aus, der auch als „umgekehrte Kausalität“ interpretiert werden könnte, indem Wahlmöglichkeiten von Lehrpersonen eher in leistungsschwächeren Klassen als remediale Maßnahmen gedacht sind.

Schülerebene

Auf Schülerebene zeigen alle sechs Modelle (siehe Tabelle 1 und 2), dass die *kognitive Leistungsfähigkeit* der Schüler/innen (hier erfasst mit Subtests aus dem Kognitiven Fähigkeitstest, Heller & Perleth, 2000) am stärksten die Leistungen im Rahmen der Testungen (Deutsch, Mathematik und Englisch) vorhersagt. Die standardisierten Effektstärken (Beta) des verbalen und des numerischen Subtests liegen zwischen rund .20 und .40, während jene der restlichen Prädiktoren zwischen .00 und .10 liegen. Erwartungsgemäß ist der Einfluss des verbalen Subtests in den Fächern Deutsch und Englisch stärker als jener des numerischen; im Fach Mathematik ist dies wie erwartet umgekehrt. Ob der Einfluss der kognitiven Schülerfähigkeiten als niedrig oder hoch zu interpretieren ist, ist mangels (österreichischer) Vergleichsstudien schwierig zu sagen. Die Tatsache, dass für andere leistungsrelevante Prädiktoren, wie bspw. die soziale Herkunft oder die Wahrnehmung der Lernumgebung, kaum nennenswerte Effekte beobachtet werden konnten, rückt die Bedeutung der kognitiven Fähigkeiten in der vorliegenden Studie jedoch in den Mittelpunkt.

Deutlich geringer als der Effekt der KFT-Subtests, aber ebenfalls in allen drei Fächern beider Generationen (G1 und G2) statistisch signifikant, ist der Einfluss des Geschlechts auf die schulische Leistung (standardisierte Beta-Koeffizienten um +/- .10). Wiederum ist dieser Einfluss erwartungskonform: In den Fächern Deutsch und Englisch zeigen Mädchen bessere Leistungen, während Jungen in den Leistungserhebungen im Fach Mathematik signifikant besser abschneiden. Diese Effekte sind allerdings als gering einzustufen.

Die erfassten Aspekte des *Sozialstatus* der Schüler/innen wirken je nach Fach und Generation unterschiedlich. Zudem sind die signifikanten Ergebnisse eher ein Resultat des hohen Stichprobenumfangs und haben kaum praktische Bedeutung. Die schwachen Effekte der sozialen Schicht bzw. des Migrationshintergrunds (standardisiertes Beta zwischen .00 und .08) sind dadurch erklärbar, dass diese Merkmale indirekt über die kognitiven Fähigkeiten wirken. Kinder und Jugendliche in ungünstigeren sozialen Verhältnissen finden eine weniger anregende Umgebung vor, um ihre kognitiven Fähigkeiten entwickeln zu können (Bradley & Corwyn, 2002).

zu erklärende Variable ⇨	Deutsch			Mathematik			Englisch		
Prädiktoren ↓	β	S.E.	p	β	S.E.	p	β	S.E.	p
Schülerebene									
Geschlecht	23.07	3.13	0.00	-17.77	2.10	0.00	14.81	2.93	0.00
Höchste Bildung (Eltern)	3.49	1.66	0.04	4.59	1.46	0.00	5.30	1.76	0.00
Höchste Position (Eltern)	0.14	0.08	0.09	0.21	0.08	0.01	0.20	0.09	0.02
Migrationsgeneration 1	-9.00	6.93	0.19	4.86	7.04	0.49	17.84	7.85	0.02
Migrationsgeneration 2	3.36	6.44	0.60	0.61	5.87	0.92	18.48	7.26	0.01
zuhaus gespr. Sprache	-5.87	5.41	0.28	-6.44	4.47	0.15	9.87	6.97	0.16
KFT (verbal)	4.64	0.22	0.00	2.94	0.14	0.00	5.08	0.28	0.00
KFT (numerisch)	2.37	0.17	0.00	3.63	0.15	0.00	2.52	0.18	0.00
Residual Leistungsmessungen	5308.29	154.57	0.00	3359.31	119.61	0.00	5657.40	203.80	0.00
Klassenebene									
selbstreguliertes Lernen	-3.54	3.98	0.37	4.41	3.46	0.20	3.08	4.22	0.47
Schülerorientierung	9.11	5.56	0.10	-5.21	5.08	0.31	3.53	7.41	0.63
Diff. Leistungsbeurteilung	2.95	4.19	0.48	-2.19	5.00	0.66	-6.51	4.74	0.17
Diff. allgemein	-6.88	15.01	0.65	6.16	6.66	0.36	7.68	18.44	0.68
Teamteaching (Teilung)	0.39	2.79	0.89	0.98	2.76	0.72	1.44	2.98	0.63
Teamteaching (verschränkt)	-1.64	2.47	0.51	-1.24	2.43	0.61	0.29	2.28	0.90
Teamteaching Ausmaß	-5.69	9.40	0.55	-8.42	8.75	0.34	-0.17	10.53	0.99
Formen Leistungsbeur.	0.74	5.26	0.89	-4.62	5.36	0.39	-7.18	5.42	0.19
K-E-L-G	-4.01	6.74	0.55	3.09	5.78	0.59	0.22	7.61	0.98
fördernder Unterricht	-32.47	11.18	0.00	-15.96	15.36	0.30	-23.50	13.05	0.07
Anomie im Unterricht	4.59	9.82	0.64	-5.74	8.82	0.52	-7.76	9.87	0.43
überfordernder Unterricht	-16.63	6.92	0.02	-3.98	7.89	0.61	-12.38	7.02	0.08
Klassenführung	33.98	13.33	0.01	12.97	14.09	0.36	20.53	13.25	0.12
individ. Lernunterstützung	21.32	15.32	0.16	19.27	15.13	0.20	29.79	14.74	0.04
Wahlmöglichkeiten im Unt.	-18.44	16.28	0.26	1.79	18.27	0.92	-10.59	19.45	0.59
Residual Leistungsmessungen	90.03	63.14	0.15	167.41	55.39	0.00	144.89	66.04	0.03
Schulebene									
Anteil Oberstufenlehrer	25.64	23.54	0.28	45.80	27.98	0.10	48.14	34.28	0.16
Nachmittagsbetreuung	0.20	5.78	0.97	-3.60	6.28	0.57	-4.30	6.13	0.48
individueller Förderplan	0.45	5.92	0.94	-0.96	5.69	0.87	-1.27	5.65	0.82
Einwohnerzahl	-0.77	3.19	0.81	-11.30	3.16	0.00	-5.19	3.13	0.10
Jahrgangsteamkoordination	-9.67	6.30	0.13	-7.58	6.56	0.25	-10.38	6.72	0.12
Schülerpartizipation	3.76	5.94	0.53	8.06	7.56	0.29	3.56	6.69	0.59
Wettbewerbe	-0.16	1.64	0.92	-0.95	1.81	0.60	-1.18	1.52	0.44
TT freie Zuteilung	-2.74	1.67	0.10	2.03	2.00	0.31	-4.15	1.98	0.04
Intercept Leistungsmessungen	475.78	61.41	0.00	501.21	54.75	0.00	486.97	57.27	0.00
Residual Leistungsmessungen	154.56	80.81	0.06	330.46	86.08	0.00	189.68	83.16	0.02

Tabelle 1: Mehrebenen-Regressionsmodell zu direkten Effekten der NMS-Konzeptmerkmale auf die fachlichen Leistungen in Deutsch, Mathematik und Englisch für die Generation 1

Anmerkungen: β = unstandardisierter Koeffizient. S.E. = Standardfehler. p = Irrtumswahrscheinlichkeit.

zu erklärende Variable ⇨	Deutsch			Mathematik			Englisch		
Prädiktoren ⇩	β	S.E.	p	β	S.E.	p	β	S.E.	p
Schülerebene									
Geschlecht	20.68	2.39	0.00	-16.70	2.84	0.00	19.88	2.32	0.00
Höchste Bildung (Eltern)	3.64	1.29	0.01	3.95	1.65	0.02	6.02	1.24	0.00
Höchste Position (Eltern)	0.20	0.06	0.00	0.02	0.09	0.80	0.21	0.06	0.00
Migrationsgeneration 1	7.02	5.06	0.17	8.80	6.41	0.17	25.92	5.23	0.00
Migrationsgeneration 2	0.74	4.43	0.87	-1.25	4.95	0.80	10.87	4.26	0.01
zu Hause gespr. Sprache	-12.11	4.24	0.00	-7.26	4.90	0.14	9.65	4.53	0.03
KFT (verbal)	5.10	0.19	0.00	2.64	0.23	0.00	5.11	0.18	0.00
KFT (numerisch)	2.59	0.12	0.00	3.47	0.19	0.00	1.81	0.11	0.00
Residual Leistungsmessungen	5844.99	130.24	0.00	9896.16	244.84	0.00	4867.09	120.91	0.00
Klassenebene									
selbstreguliertes Lernen	-1.43	2.53	0.57	-1.82	2.93	0.53	-3.38	2.73	0.22
Schülerorientierung	-1.45	3.41	0.67	-0.70	5.09	0.89	4.64	4.79	0.33
Diff. Leistungsbeurteilung	-2.75	2.85	0.34	-3.53	3.36	0.29	-2.35	3.05	0.44
Diff. Allgemein	-8.12	6.04	0.18	4.70	8.23	0.57	-6.02	8.72	0.49
Teamteaching (Teilung)	1.16	2.00	0.56	1.54	2.64	0.56	0.29	2.26	0.90
Teamteaching (verschrän.)	0.59	1.74	0.74	0.91	1.90	0.63	-1.09	1.84	0.55
Teamteaching Ausmaß	0.71	6.63	0.92	-1.66	8.22	0.84	-6.18	8.99	0.49
Formen Leistungsbeur.	-1.23	1.82	0.50	-0.01	2.49	1.00	-1.06	1.90	0.58
K-E-L-G	0.52	5.97	0.93	1.44	8.11	0.86	1.14	5.55	0.84
fördernder Unterricht	-8.06	8.52	0.34	-1.43	11.53	0.90	-20.81	9.77	0.03
Anomie im Unterricht	-4.78	7.22	0.51	-29.22	8.66	0.00	-17.42	6.98	0.01
überfordernder Unterricht	-15.97	5.77	0.01	-12.85	6.62	0.05	-12.20	6.10	0.05
Klassenführung	1.58	11.21	0.89	-16.07	13.16	0.22	17.30	10.59	0.10
individ. Lernunterstützung	25.02	13.00	0.05	8.02	14.54	0.58	-3.75	13.06	0.77
Wahlmöglichkeiten im Unt.	-21.10	13.37	0.11	6.67	14.52	0.65	-53.06	14.69	0.00
Residual Leistungsmessungen	135.48	42.52	0.00	182.69	85.47	0.03	264.47	58.22	0.00
Schulebene									
Anteil Oberstufenlehrer	1.34	8.98	0.88	7.47	8.79	0.40	15.82	11.10	0.15
Nachmittagsbetreuung	-8.39	3.84	0.03	-13.00	5.68	0.02	2.58	4.12	0.53
individueller Förderplan	-3.70	3.61	0.31	-4.55	4.47	0.31	3.42	4.01	0.39
Einwohnerzahl	0.75	0.28	0.01	-0.01	0.40	0.98	1.11	0.40	0.01
Jahrgangsteamkoordination	-2.62	3.65	0.47	3.36	4.95	0.50	3.23	4.37	0.46
Schülerpartizipation	-2.80	2.93	0.34	-4.01	4.80	0.40	-3.60	4.63	0.44
Wettbewerbe	0.73	1.24	0.56	0.36	1.39	0.80	1.64	1.33	0.22
TT freie Zuteilung	0.43	1.48	0.77	-2.99	2.37	0.21	1.30	2.38	0.59
Intercept Leistungsmessungen	537.67	45.30	0.00	606.49	51.41	0.00	682.18	39.69	0.00
Residual Leistungsmessungen	12.90	29.90	0.67	92.79	66.51	0.16	140.94	89.99	0.12

Tabelle 2: Mehrebenen-Regressionsmodell zu direkten Effekten der NMS-Konzeptmerkmale auf die fachlichen Leistungen in Deutsch, Mathematik und Englisch für die Generation 2

Anmerkungen: siehe Tabelle 1

Die durch die angeführten Regressionsmodelle aufgeklärte Varianz auf den unterschiedlichen Analyseebenen beträgt für das Fach Deutsch in G1 und G2 39 % bzw. 40 % (Schülerebene), 69 % bzw. 44 % (Klassenebene) und 25 % bzw. 76 % (Schulebene).³ Für das Fach Mathematik sind das 48 % bzw. 20 % (Schülerebene), 29 % bzw. 38 % (Klassenebene) und 39 % bzw. 38 % (Schulebene). Im Fach Englisch beträgt die Aufklärungsrate 38 % bzw. 36 % (Schülerebene), 56 % bzw. 46 % (Klassenebene) und 43 % bzw. 42 % (Schulebene). Grundsätzlich zeigt sich, dass der Anteil der Varianz in den Schülerleistungen, der auf Klassenebene liegt und daher durch Prädiktoren auf dieser Ebene erklärt werden kann, relativ gering ist – die Intra-Klassen-Korrelation (ICC) auf Klassenebene für alle Fächer und beide Generationen beträgt rund 6 % – während der Anteil auf Schulebene der erwarteten Höhe entspricht (ICC Schulebene: Deutsch 16 % in G1 bzw. 9 % in G2, Mathematik 23 % bzw. 10 % und Englisch 13 % bzw. 9 %).

4.2 Mehrebenen-Pfadmodelle

Abbildung 2 und 3 zeigen die Ergebnisse der Mehrebenen-Pfadmodelle je Generation. Aus Platzgründen werden die Ergebnisse der drei untersuchten Schulfächer jeweils in einer Abbildung zusammengefasst dargestellt. Die Pfadkoeffizienten sind dabei in folgender Reihenfolge angeführt: Deutsch – Mathematik – Englisch. Aus Gründen der Übersichtlichkeit werden lediglich die Koeffizienten der statistisch signifikanten Pfade berichtet (durchgehende Pfeile).

Die Abbildungen zeigen, dass auch die Resultate der Analysen zu den indirekten Effekten der Konzeptmerkmale mittels Mehrebenen-Pfadmodellen ernüchternd sind. Vermutete indirekte Effekte der NMS-Konzeptmerkmale auf die Schülerleistungen in den standardisierten Erhebungen über die Schülerwahrnehmung der Lernumgebung lassen sich nicht bzw. nur im Fach Englisch in G1 bestätigen.

Immerhin kann für die *Generation 1* gezeigt werden, dass das Ausmaß an (1) Teamteaching und (2) der Unterstützung selbstgesteuerten Lernens positiv auf die Schülerwahrnehmung der Lernumgebung wirken: Je höher der Anteil von Teamteaching-Stunden an den Wochenstunden eines Fachs ist, desto häufiger geben Klassenvorstände an, dass in der Klasse verstärkt schülerorientiert unterrichtet wird (das gilt nur für Mathematik) und die Schüler/innen mehr Freiraum zum selbstständigen Lernen erhalten (das gilt für Mathematik und Englisch). Auch hängt das Ausmaß an Teamteaching mit der aggregierten Schülerwahrnehmung von Unterrichtsqualität zusammen: Teamteaching geht mit einer stärkeren individuellen Lernunterstützung, stärker förderndem Unterricht und vermehrten Wahlmöglichkeiten im Unterricht (aus aggregierter Schülersicht) einher. Dies gilt vor allem für die Fächer Mathematik und Englisch. Der Unterricht in Schulklassen, von denen die Klassenvorstände angeben, dass selbstreguliertes Lernen forciert wird, wird von den Schüler/innen im Mittel auch als stärker fördernd wahrgenommen. Auch nehmen diese Schüler/innen im Mittel mehr Wahlmöglichkeiten im Unterricht wahr. Diese Effekte selbstgesteuerten Lernens sind in allen drei Fächern beobachtbar. Die Wirkung der geteilt wahrgenommenen Merkmale der Lernumgebung, die auf Basis der Theorien zum selbstgesteuerten Lernen als lernförderlich angesehen werden können, ist allerdings auf das Fach Englisch beschränkt: Während das Merkmal individuelle Lernunterstützung mit positiven

³ Der hohe Unterschied in der Varianzaufklärung im Fach Deutsch auf Schulebene zwischen G1 (25 %) und G2 (76 %) ist auf zwei Ursachen zurückzuführen: (1) In G2 ließen sich zwei signifikante Prädiktoren auf Schulebene (Nachmittagsbetreuung und Einwohnerzahl am Schulstandort) identifizieren, die in G1 dagegen nicht zur Varianzaufklärung beitragen. (2) In G2 liegt im Vergleich zu G1 grundsätzlich weniger zu erklärende Varianz in den Deutsch-Leistungen auf Schulebene ($\sigma^2_{G1} = 865,34 / \sigma^2_{G2} = 1.584,75$) wodurch bereits wenige signifikante Prädiktoren reichen, um einen bedeutenden Anteil zu erklären. Dieselben Gründe sind für den Unterschied bzgl. der Varianzaufklärung in der Mathematikleistung auf Schülerebene verantwortlich.

Effekten auf die fachliche Leistung in Englisch einhergeht, hängt das Merkmal Wahlmöglichkeiten negativ mit den Schülerleistungen zusammen. Es gilt: Ein/e Schüler/in, der/die in eine Klasse geht, in der das Unterrichtsmerkmal individuelle Lernunterstützung/Wahlmöglichkeiten höher ausgeprägt ist, erbringt eine um rund 42 bzw. 38 Punkte bessere/schlechtere Leistung in Englisch, als ein/e Schüler/in, der/die in eine Klasse geht, in der das Merkmal weniger stark ausgeprägt ist. Effekte der Konzeptmerkmale individueller Förderplan und Nachmittagsbetreuung konnten in G1 nicht beobachtet werden.

Die Effekte der NMS-Merkmale sind über die beiden Generationen hinweg jedoch nicht konstant. In *Generation 2* verschwinden die positiven Effekte des Teamteachings und des selbstregulierten Lernens. Das Ausmaß an Teamteaching wirkt nun sogar negativ auf das Ausmaß an selbstreguliertem Lernen. Hingegen wirkt nun in G2, im Gegensatz zu G1, in allen Fächern das Ausmaß an Schülerorientierung positiv auf die von den Schüler/inne/n im Klassenmittel wahrgenommene individuelle Lernunterstützung und die im Unterricht angebotenen Wahlmöglichkeiten. Auch für die Prädiktoren der Schulebene zeigen sich in G2 v. a. im Fach Mathematik Effekte: Hier stehen allerdings entgegen den Erwartungen (1) die Nachmittagsbetreuung im negativen Zusammenhang mit den Schülerleistungen und (2) das Merkmal individueller Förderplan im negativen (allerdings knapp nicht signifikanten) Zusammenhang mit der auf Klassenebene aggregierten Wahrnehmung von individueller Lernunterstützung, wenn man für die anderen Variablen kontrolliert. In anderen Worten: Ein/e Schüler/in, der/die in eine Schule mit Nachmittagsbetreuung geht, erreicht eine um rund 16 Punkte schlechtere Leistung in Mathematik, als ein/e Schüler/in mit der gleichen Merkmalskombination, der/die jedoch eine Schule ohne Nachmittagsbetreuung besucht. Auch dieser Effekt ist wohl als umgekehrte Kausalität zu interpretieren: Schulen mit Nachmittagsbetreuung werden womöglich eher an herausfordernden Schulstandorten angeboten. Diese Annahme bleibt zu prüfen.

Auch berichten Schüler/innen der G2 in allen Fächern im Mittel von weniger förderndem Unterricht, wenn Schulleitungen berichten, dass an ihrer Schule nach individuellen Förderplänen gearbeitet wird. Dieses Ergebnis stellt einen einigermaßen überraschenden Zusammenhang dar, über dessen Erklärung man nur mutmaßen kann. Ähnlich wie in G1 konnten auch in G2 keine signifikanten Zusammenhänge zwischen den Schülerleistungen und den aggregierten Schülerurteilen über den Unterricht im Hinblick auf individuelle Förderung, fördernden Unterricht und gebotene Wahlmöglichkeiten identifiziert werden. Indirekte Effekte der NMS-Konzeptmerkmale bleiben daher auch in G2 aus.

Auf Schülerebene gilt für beide Generationen, dass, wie bereits in den Mehrebenen-Regressionsmodellen berichtet, allen voran die kognitiven Fähigkeiten zentralen Einfluss auf die fachlichen Schülerleistungen nehmen.

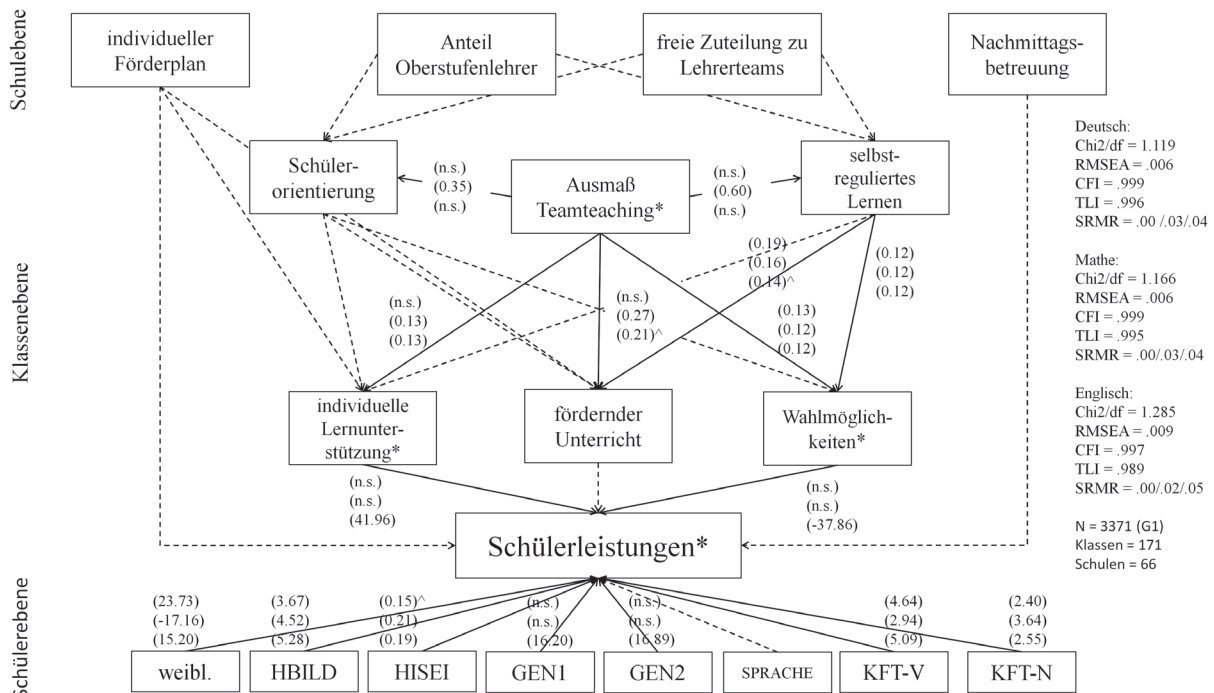


Abbildung 2: Empirisches Mehrebenen-Pfadmodell zur Analyse der indirekten Effekte der NMS-Konzeptmerkmale auf die fachlichen Leistungen in den Unterrichtsgegenständen Deutsch, Mathematik und Englisch für die Generation 1

Anmerkungen:

* Items wurden fachspezifisch formuliert bzw. beziehen sich auf die Leistungen im jeweiligen Fach.

n.s. = Effekte sind statistisch nicht signifikant

^ knapp nicht signifikante Effekte zwischen .05 und .06 Irrtumswahrscheinlichkeit

Pfeile mit unterbrochenen Linien zeigen Pfade an, die in allen Fächern statistisch nicht signifikant von null verschieden sind.

Chi2/df = Verhältnis von Chi2-Wert zu Freiheitsgraden

RMSEA = root mean square error of approximation

CFI = comparative fit index

TLI = Tucker-Lewis-index

SRMR = standardized root mean square residual (Werte beziehen sich auf: Individual-/Klassen-/Schulebene)

N = Stichprobengröße auf Schülerebene

HBILD = höchste schulische Bildung der Eltern

HISEI = höchste berufliche Position der Eltern

GEN1 = Migration erster Generation

GEN2 = Migration zweiter Generation

Sprache = zuhause gesprochene Sprache

KFT-V = kognitiver Fähigkeitstest (verbal)

KFT-N = kognitiver Fähigkeitstest (numerisch)

Zwischen den aggregierten Schülerwahrnehmungen wurden Residualkovarianzen spezifiziert.

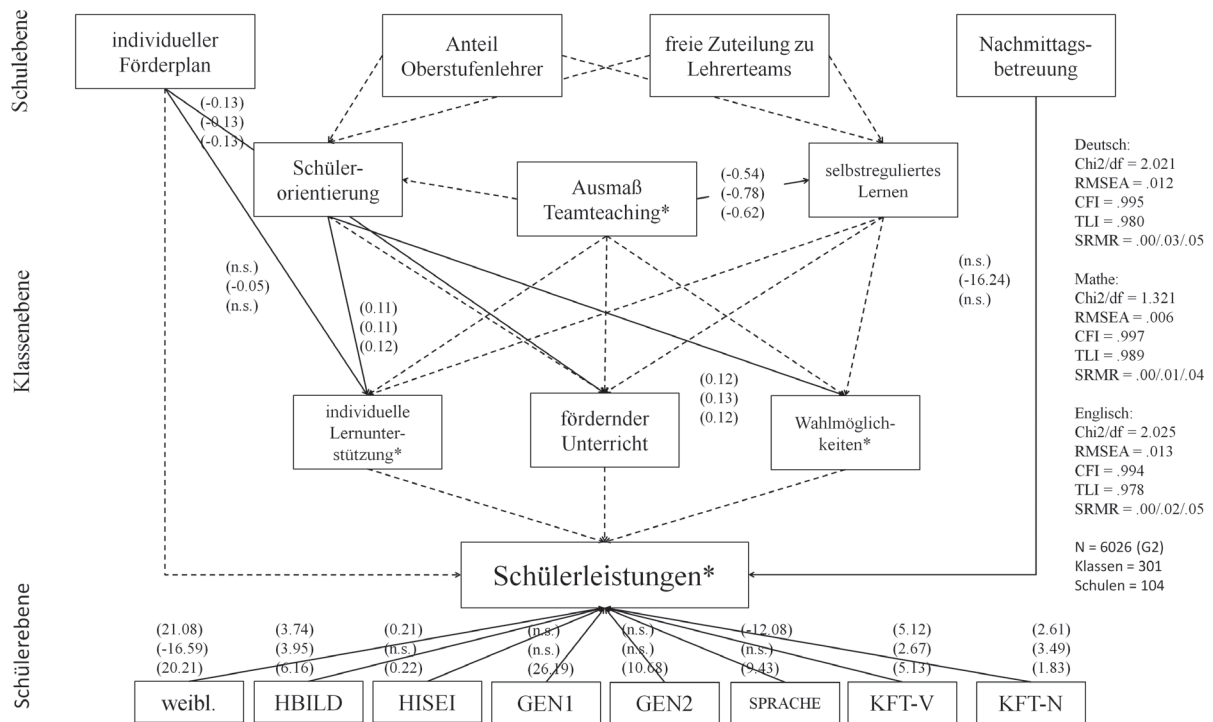


Abbildung 3: Empirisches Mehrebenen-Pfadmodell zur Analyse der indirekten Effekte der NMS-Konzeptmerkmale auf die fachlichen Leistungen in den Unterrichtsgegenständen Deutsch, Mathematik und Englisch für die Generation 2

Anmerkungen: Legende siehe Abbildung 2

5. Zusammenfassung und Diskussion

In diesem Beitrag wurden direkte und indirekte Effekte von zentralen NMS-Konzeptmerkmalen auf die Schülerleistungen in den drei Unterrichtsgegenständen Deutsch, Mathematik und Englisch untersucht. Die Resultate von Mehrebenen-Regressionsanalysen und Mehrebenen-Pfadmodellen zeigen für beide NMS-Generationen und für alle drei Schulfächer ernüchternde Ergebnisse. Angenommen wurde, dass auf Klassenebene v. a. das Teamteaching sowie eine neue Lehr-Lern-Kultur, die sich durch verstärkte Schülerorientierung und verstärkten Einsatz selbstregulierter Lernformen charakterisiert, positiv auf die Leistungserbringung der Schüler/innen wirken. Auf Schulebene sollten der Anteil an AHS-Lehrer/inne/n, das Angebot von Nachmittagsbetreuung und individuelle Förderpläne sowie das Vorhandensein von Jahrgangsteamkoordinator/inn/en positiv (auch über die Konzept-Elemente auf Klassenebene) wirken. Allerdings konnten kaum direkte oder indirekte Effekte der NMS-Merkmale identifiziert werden. Wo vereinzelt Effekte beobachtet werden konnten, wirken sie teils erwartungswidrig negativ bzw. wirken sie nicht konsistent über beide Generationen hinweg.

Für das Fehlen der erwarteten Effekte der NMS-Konzeptmerkmale sind unterschiedliche Erklärungen denkbar und plausibel:

- (1) Schrader und Helmke (2008; vgl. auch Grube & Hasselhorn, 2006) verweisen darauf, dass fachspezifisches Vorwissen und kognitive Leistungsfähigkeiten von Schüler/inne/n die stärksten Prä-

diktoren der fachlichen Leistungserbringung bilden. Auch in der vorliegenden Studie zeigt sich, dass die verbale und numerische kognitive Fähigkeit den stärksten Einfluss auf die fachliche Leistungserbringung hat, und das, obwohl mit den vorliegenden NMS-Stichproben im Hinblick auf die kognitiven Fähigkeiten und die soziale Zusammensetzung nur ein eingeschränktes Spektrum des Möglichen (es fehlt die AHS) abgebildet ist. Daraus lässt sich ableiten, dass der Förderung der kognitiven Schülerfähigkeiten in Zukunft mehr Aufmerksamkeit gewidmet werden sollte (siehe auch Kapitel 13), da Studien zeigen (Flach & Sandgren, 2011; Ceci, 1991), dass die Schule sehr wohl einen Einfluss auf die Entwicklung kognitiver Fähigkeiten (Intelligenz) haben kann.

- (2) Dass in der vorliegenden Untersuchung kaum Effekte der Lernumgebung nachgewiesen werden konnten, ist kein Einzelfall: Auch andere Studien, die Interventionen untersuchten, die – wie die NMS – die Selbststeuerung der Schüler/innen fördern sollen, konnten nur begrenzte Effekte der Lernumgebung und des Lehrerverhaltens nachweisen (z. B. Hattie, 2009; Hedges & Gaiconia, 1982; Helm, 2014; Krammer, 2009). Eine Erklärung wird oft darin gesehen, dass die Intensität und Qualität der Umsetzung meist nicht standardisiert erfolgt und sehr stark zwischen den Lehrpersonen variiert (bspw. beim Teamteaching). Gleichzeitig darf die Rolle der Lehrperson und des Unterrichts nicht „kleingeredet“ werden: Nach Hattie (2009) ist bei manchen Unterrichtsmerkmalen (z. B. formative Evaluation des Unterrichts, reziprokes Lehren) der Effekt beachtlich groß (Cohens' $d > .70$).
- (3) In der vorliegenden Studie gilt diese Variabilität in der Umsetzung nicht nur auf Klassenebene, sondern auch auf Schulebene: Wie in Kapitel 12 ist auch hier darauf hinzuweisen, dass Schulen bei der Implementierung innerhalb eines allgemein vorgegebenen Rahmens unterschiedliche Schwerpunkte setzen konnten. Dadurch wird die Intervention „NMS“ verzerrt und die Zuschreibung von Effekten erschwert, wenn bestimmte NMS-Elemente sich gegenseitig kompensieren können (z. B. das Ausmaß an Teamteaching durch qualitativ vollen schülerorientierten Unterricht). Des Weiteren ist nicht immer klar, wie deutlich die NMS-Konzeptmerkmale an der empirischen Unterrichtsforschung ausgerichtet sind. Beispielsweise geht aus den Konzeptbeschreibungen nicht hervor, ob und inwiefern empirisch bestätigte Qualitätsmerkmale der Schülerorientierung und Differenzierung berücksichtigt wurden.
- (4) Maßnahmen der Schülerorientierung und Differenzierung (z. B. durch das Anbieten von Wahlmöglichkeiten im Unterricht) werden oft als Strategie zur Förderung leistungsschwacher Schüler/innen eingesetzt (Helm, 2014; Seifried & Wuttke, 2010). Dadurch können negative Zusammenhänge zwischen theoretisch als positiv einzuschätzenden Unterrichtsmerkmalen und den Schülerleistungen entstehen (umgekehrte Kausalität). Auch egalisieren sich so positive und negative Effekte.
- (5) Weiters muss gefragt werden, ob die Konzeptelemente spezifisch genug operationalisiert wurden: Aufgrund der standardisierten Befragung ist es wohl hauptsächlich gelungen, die Quantität der Implementation der Konzeptmerkmale zu erheben; die Erfassung der Qualität der Umsetzung bringt aber große methodische Herausforderungen mit sich. Zum einen ist diese Qualität von Unterrichtsfach zu Unterrichtsfach neu zu definieren, wie bspw. Klieme und Rakoczy (2008) für das Unterrichtsmerkmal „kognitive Aktivierung“ argumentieren. Was guten Unterricht ausmacht, ist zum anderen aber nicht nur in Abhängigkeit von der Domäne zu bestimmen, sondern auch in Ab-

hängigkeit von den Schülermerkmalen, allen voran der kognitiven Leistungsfähigkeit. Leistungsfähigere Schüler/innen tendieren dazu, das Lehrerverhalten positiver einzuschätzen, obwohl dies objektiv gesehen womöglich gar nicht der Fall ist.

- (6) Viele der NMS-Konzeptelemente sind vermutlich zu „distal“ – d. h., zwischen ihnen und dem Lernerfolg von Schüler/inne/n liegt eine zu lange Wirkungskette –, sodass ihre Effekte nicht oder vielfach (und) unterschiedlich vermittelt bei den Lernenden ankommen bzw. in den Schülerleistungen ersichtlich sind.
- (7) Letztlich fehlt der vorliegenden Evaluationsstudie eine qualitative Tiefenstudie, die Stärken und Schwächen der Umsetzung der NMS-Konzeptmerkmale in der schulischen Praxis und deren Wirkung auf das Schülerlernen differenziert beschreibt. Tatsächlich wäre es nötig und sinnvoll gewesen, von Beginn an einige Schulen im Sinne qualitativer Studien intensiv zu begleiten. Die Ergebnisse solcher Studie hätten für die Hypothesengenerierung und die Variablenauswahl und -operationalisierung der hier vorgelegten summativen Evaluation wichtige Funktionen einnehmen können.

6. Schlussfolgerungen und Weiterentwicklungsempfehlungen

Welche Schlussfolgerungen sind nun aus diesen Ergebnissen zu ziehen?

- (1) Zunächst ist zu konstatieren, dass die neuen Merkmale der Unterrichtsorganisation, insbesondere Teamteaching und die „neue Lernkultur“, die sich v. a. durch Individualisierung und schülerorientierte Förderung ausdrücken soll, zwar in manchen Aspekten von den Schüler/inne/n wahrgenommen werden, sich aber – in der Weise, wie sie in den untersuchten Schulen umgesetzt wurden und wie sie in den von uns zu analysierenden Untersuchungen erhoben wurden – nur bedingt positiv in den Schülerleistungen niederschlagen. Dies kann – wie schon oben angedeutet – im Wesentlichen drei Gründe haben: (a) Es kann in der Sache selbst liegen – in dem Sinne, dass die beobachteten Merkmale tatsächlich untaugliche Mittel für die Förderung von Schülerleistungen sind (aber möglicherweise für andere Ziele sinnvolle Auswirkungen erwarten lassen; vgl. Argument 2 unten). (b) Als zweiter möglicher Grund kann eine unzureichende oder unangemessene Implementation/praktische Realisierung an sich sinnvoller Veränderungsstrategien genannt werden (vgl. Argument 3 unten). (c) Ein dritter Grund kann in der unzureichenden Operationalisierung der zentralen NMS-Merkmale und anderen Mängeln der Art und Weise der empirischen Erhebung liegen (vgl. Argument 5 unten). Mit den verfügbaren Daten kann man nicht entscheiden, was die wahrscheinlichste Erklärung ist (allerdings sind methodische Verbesserungen im Sinne von [c] auf jeden Fall sinnvoll). Man kann aber jeweils unterschiedliche Wahrscheinlichkeiten und angemessene Konsequenzen diskutieren.
- (2) Ist aufgrund des bisherigen Forschungsstands zu erwarten, dass *Teamteaching und Neue Lernkultur* im Sinne von unterrichtlicher Individualisierung und schülerorientierter Förderung überhaupt positive Effekte auf Leistungsergebnisse haben werden? Beide Merkmale sind durchaus plausible Strategien, um mit der Heterogenität von Schüler/inne/n konstruktiv umzugehen; gleichzeitig sind sie offenbar aber keine sehr durchschlagskräftigen Strategien, die unabhängig vom jeweiligen

Kontext erwünschte Ergebnisse erbringen. So berichtet Hattie (2009, S. 219) eine schwache mittlere Effektstärke von $d = .19$ für Teamteaching (keine Erwähnung bei Helmke, 2009) und konstatiert gleichzeitig einen Mangel an belastbaren Untersuchungen zu diesem Thema.

Die Situation schaut für *Individualisierung und Innere Differenzierung* des Unterrichts nicht grundlegend anders aus. Auch hier haben die bisherigen empirischen Untersuchungen sehr unterschiedliche Ergebnisse für die Leistungsentwicklung zutage gefördert, während die Effekte sozialer und motivationaler Variablen häufiger positiv zu sein scheinen (vgl. den Überblick in Altrichter, Trautmann, Wischer, Sommerauer & Doppler, 2009). Die in der Metaanalyse von Hattie (2009, S. 198 ff.) erfassten Programme von „individual instruction“ haben mit einer mittleren Effektstärke von $d = .23$ nur schwache Wirkungen. Auch hier ist zu vermuten, dass das Konzept „Individualisierung“ so weit ist, dass hier sehr unterschiedliche Realisierungsformen – effektivere und weniger effektive – miteinander verglichen werden.

„Die bloße Tatsache, dass Individualisierung stattfindet, ist weder gut noch schlecht – je nachdem, wie sie realisiert wird, kann sie dilettantisch oder brillant sein, Gutes oder Schlechtes (oder gar nichts) bewirken.“ (Helmke, 2009, S. 259)

Man kann also zusammenfassen: Teamteaching und Neue Lernkultur im Sinne von unterrichtlicher Individualisierung und schülerorientierter Förderung sind keine unsinnigen Merkmale einer sinnvoll auf Schülerheterogenität eingehenden und leistungsfördernden Lernkultur, sie stehen aber in Gefahr, sehr unterschiedlich implementiert zu werden, und sie sind für sich nicht die durchschlagskräftigsten Strategien zur Leistungsförderung von Schüler/inne/n. Daraus kann man zwei Schlüsse ziehen, die in den folgenden beiden Abschnitten kurz angesprochen werden.

- (3) Der erste Schluss setzt bei der Begründung (b) an: Man kann wohl bei kurzfristig in großer Breite umgesetzten Innovationen wie der NMS mit einer gewissen Wahrscheinlichkeit vermuten, dass die Konzeptmerkmale Teamteaching und Individualisierung in sehr unterschiedlichem Ausmaß und in sehr unterschiedlicher Qualität implementiert wurden. Weitere Initiativen vonseiten der *Implementierungsunterstützung* können daher nützlich sein: Einesteils könnten wichtige qualitative Aspekte der Neuerungen genauer beschrieben und konkrete, sehr praktikabel aufbereitete Realisierungsbeispiele angeboten werden, um die Kernelemente der Neuerung bewusster zu machen. Andererseits könnten Implementierungs- und Realisierungsschritte erleichtert werden, indem Fortbildungsmaßnahmen sehr direkt bei der Umsetzung ansetzen und diese begleiten. Merkmale solcher Lehrerfortbildung können beispielsweise bei Lipowsky (2004) gefunden werden; auch scheint das „*lesson study*“-Konzept, bei dem Lehrergruppen sukzessive ihre Praxis weiterentwickeln (Fernandez, 2002; Kullmann, 2012), für diese Zwecke geeignet.
- (4) Die zweite Konsequenz könnte sein, die genannten Neuerungen explizit durch *Unterrichtsstrategien zu ergänzen, die in einem relativ gut gesicherten Zusammenhang mit Leistungsergebnissen stehen*. Die Grundideen lassen sich beispielsweise mit den von Klieme und Rakoczy (2008, S. 228) genannten drei Basisdimensionen guten Unterrichts zusammenfassen: strukturierte, klare und störungspräventive Unterrichtsführung, unterstützendes, schülerorientiertes Sozialklima sowie kognitive Aktivierung. Eine solche „unterrichtsbezogene Verbreiterung“ der NMS-Kernmerkmale mag schon in der jetzt laufenden Implementierungsstrategie angelegt sein, in der auf Lerndesigner/

innen und verschiedene didaktische Überlegungen verwiesen wird. In diesem Falle wäre möglicherweise (a) eine stärkere Betonung dieser Merkmale durch ihre Aufnahme in das NMS-Kernkonzept sowie (b) eine Intensivierung der Implementationsbegleitung dieser didaktischen Aspekte im Sinne von Vorschlag (3) sinnvoll.

- (5) Um eine mögliche Wirkungskette von Änderungen der Schulstruktur und der Unterrichtsgestaltung bis zu Schülerleistungen zu rekonstruieren, ist vor allem (und hier wird auf die methodische Begründung [c] eingegangen) eine differenziertere Erfassung der Implementation der Unterrichtsneuerungen in verschiedenen Klassen und deren Nutzung durch Schüler/innen notwendig, die auch sensibel für qualitative Unterschiede ist. Dazu bieten sich zunächst wohl qualitative Studien an, die unterschiedliche Realisierungs- und Nutzungsmodi der zentralen unterrichtsbezogenen NMS-Merkmale durch Unterrichtsbeobachtung und Videoanalysen erfassen und kategoriale Grundlagen für weiterführende quantitative Studien legen.

Literatur

- Altrichter, H., Trautmann, M., Wischer, B., Sommerauer, S. & Doppler, B. (2009). Unterrichten in heterogenen Gruppen: Das Qualitätspotenzial von Individualisierung, Differenzierung und Klassenschülerzahl. In W. Specht (Hrsg.), *Nationaler Bildungsbericht 2009. Band 2. Fokussierte Analysen Bildungspolitischer Schwerpunktthemen* (S. 341–360). Graz: Leykam.
- BMBF (2014). *Neue Mittelschule. Wien: Bundesministerium für Bildung und Frauen*. Verfügbar unter: <http://www.neuemittelschule.at/> [11.08.2014]
- Bradley, R. H. & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53, 371–399.
- Bruneforth, M., Weber, C. & Bacher, J. (2012). Chancengleichheit und garantiertes Bildungsminimum in Österreich. In B. Herzog-Punzenberger (Hrsg.), *Nationaler Bildungsbericht Österreich 2012. Band 2. Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (S. 189–228). Graz: Leykam.
- Ceci, S. J. (1991). How Much Does Schooling Influence General Intelligence and Its Cognitive Components? A Reassessment of the Evidence. *Developmental Psychology*, 27(5), 703–722.
- Enders, C. K. & Tofighi, D. (2007). Centering Predictor Variables in Cross-Sectional Multilevel Models: A New Look at an Old Issue. *Psychological Methods*, 12(2), 121–138.
- Expertenkommission (2007). *Expertinnenkommission Zukunft der Schule. Erster Zwischenbericht*. Wien: BMUKK. Verfügbar unter: https://www.bmbf.gv.at/ek_zwb_01_15690.pdf?4dtiae [11.08.2014]
- Fernandez, C. (2002). Learning from Japanese approaches to professional development: The case of lesson study. *Journal of Teacher Education*, 53(5), 393–405.

- Flach, T. & Sandgren Massih, S. (2011). The effect of education on cognitive ability. *Economic Inquiry*, 49(3), 838–856.
- Giaconia, R. M. & Hedges, L. V. (1982). Identifying Features of Effective Open Education. *Review of Educational Research*, 52(4), 579–602.
- Grube, D. & Hasselhorn, M. (2006). Längsschnittliche Analysen zur Lese-, Rechtschreib- und Mathematikleistung im Grundschulalter: Zur Rolle von Vorwissen, Intelligenz, phonologischem Arbeitsgedächtnis und phonologischer Bewusstheit. In I. Hosenfeld & F.-W. Schrader (Hrsg.), *Schulische Leistung: Grundlagen, Bedingungen, Perspektiven* (S. 87–105). Münster: Waxmann.
- Hattie, J. (2009). *Visible Learning*. New York: Routledge.
- Heller, K. A. & Perleth, Ch. (2000). *Kognitiver Fähigkeits-Test (Rev.) für 5.–12. Klassen (KFT 5–12+ R)*. Göttingen: Beltz-Testgesellschaft.
- Helm, C. (2014). *Lernen in Offenen und Traditionellen UnterrichtSettings (LOTUS)*. Univ. Dissertation. Abteilung für Pädagogik und Psychologie, Johannes Kepler Universität, Linz.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Klett-Kallmeyer.
- Hox, J. J. (2013). Multilevel regression and multilevel structural equation modeling. In T. D. Little (Hrsg.), *The Oxford Handbook of Quantitative Methods* (S. 281–294). Oxford: Oxford University Press.
- Kemethofer, D., & Helm, C. (2015). Operationalisierung der Konzeptmerkmale der Neuen Mittelschule. In F. Eder, H. Altrichter, J. Bacher, F. Hofmann & C. Weber (Hrsg.), *Evaluation der Neuen Mittelschule (NMS). Befunde aus den Anfangskohorten*. Forschungsbericht, Salzburg und Linz, 2015. Online-Dokument. Verfügbar unter: <https://www.bifie.at/node/2736>
- Klieme, E. & Rakoczy, K. (2008) Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237.
- Krammer, K. (2009). *Individuelle Lernunterstützung in Schülerarbeitsphasen*. Münster: Waxmann.
- Kullmann, H. (2012). Lesson Study – eine konsequente Form unterrichtsbezogener Lehrerverkooperation. In S. G. Huber & F. Ahlgrimm (Hrsg.), *Kooperation. Aktuelle Forschung zur Kooperation in und zwischen Schulen sowie mit anderen Partnern* (S. 69–88). Münster: Waxmann.
- Lipowsky, F. (2004). Was macht Fortbildungen für Lehrkräfte erfolgreich? *Die Deutsche Schule*, 4, 462–479.
- Muthén, B. O. & Muthén, L. K. (1998–2014). *Mplus (Version 7.2)* [Computer Software]. Los Angeles, CA.
- Schrader, F.-W. & Helmke, A. (2008). Determinanten der Schulleistung. In M. K. W. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion* (S. 286–302), Wiesbaden: VS.

- Seifried, J. & Wuttke, E. (2010). Potenziale des Lernens aus Fehlern in Abhängigkeit von methodischen Grundentscheidungen. *Zeitschrift für Berufs- und Wirtschaftspädagogik, (Beiheft 23)*, 155–171.
- Weber, C. (2015). Sind Leistungsverbesserungen an NMS-Modellschulen zu beobachten? Technischer Bericht. In F. Eder, H. Altrichter, J. Bacher, F. Hofmann & C. Weber (Hrsg.), *Evaluation der Neuen Mittelschule (NMS). Befunde aus den Anfangskohorten*. Forschungsbericht, Salzburg und Linz, 2015. Online-Dokument. Verfügbar unter: <https://www.bifie.at/node/2736>