

Hartmut Esser

Professor für Soziologie und Wissenschaftslehre (em.) an der Universität Mannheim

Stellungnahme zum Entwurf eines Gesetzes zur Änderung des Schleswig-Holsteinischen Schulgesetzes - Drucksache 19/1107

In dem o.a. genannten Entwurf wird die Frage aufgeworfen, welche Effekte die (externe) Differenzierung nach den kognitiven Fähigkeiten und Leistungen auf das (spätere) Kompetenzniveau und die Ungleichheit im Bildungserfolg nach der sozialen Herkunft haben. Das wird auf einen Vorschlag zur Änderung des Schleswig-Holsteinischen Schulgesetzes angewandt, wonach, so wie ich es verstehe, eine jetzt nicht mehr vorgesehene Möglichkeit zur externen Differenzierung wieder ein geführt werden soll. Dabei beruft sich der Antrag auf eine in der Kölner Zeitschrift für Soziologie und Sozialpsychologie publizierte Studie von mir (und Kerstin Hoenig, LifBi Bamberg), wonach „... über die kognitive Homogenisierung der Schüler und Schulklassen die Effizienz des Kompetenzerwerbs steigt. Zudem würden ‘die Ungleichheiten nach der sozialen Herkunft in den nach Leistung differenzierenden Systemen nicht nur nicht größer, sondern eher geringer‘“ (Auszug Drucksache 19/1107; s. unten). Dazu wäre das Folgende zu sagen.

Die angesprochene Studie bezieht sich auf die Bildungsbeteiligung im Übergang von der Grundschule auf die Sekundarstufe 1 und beruht auf einem Vergleich aller 16 deutschen Bundesländer über die Daten der „National Educational Panel Study“ (NEPS), einer bundesweit durchgeführten (und Bund-Länder-finanzierten) breit angelegten Wiederholungsuntersuchung u.a. zu den Bildungsverläufen im Übergang nach der Grundschule und dem weiteren Verlauf in der Sekundarstufe. Dabei wurde der Effekt der *Verbindlichkeit* der Empfehlungen auf den Übergang in das Gymnasium untersucht. Das wichtigste Ergebnis war, dass sich in den Ländern mit Verbindlichkeit der Empfehlungen die Effekte der sozialen Herkunft nicht verstärkten, sondern eher abschwächten, vor allem aber, dass die Kopplung des Übergangs an die Leistungen vorher in der Grundschule stärker wurde, sich also die Leistungsgerechtigkeit des Gymnasialübergangs erhöhte. Besonders interessant war daneben noch, dass Eltern aus den unteren sozialen Schichten öfter auch ihre begabten Kinder, nämlich die mit Empfehlungen, nicht auf das Gymnasium ließen. Das ist bei den bestehenden Regelungen auch bei Verbindlichkeit möglich, weil Abweichungen von den Empfehlungen nach *unten* immer möglich sind. Es war dabei vermutet worden, dass eine entsprechende Erweiterung der Verbindlichkeit eine weitere Maßnahme sein könnte, die Effekte der sozialen Herkunft weiter zu begrenzen bzw. zu senken – was freilich einen entsprechend weiteren Eingriff in den „Elternwillen“ bedeuten würde.

Zu der Entwicklung der Leistungen dann in der Sekundarstufe war in diesem Beitrag (noch) nichts gesagt worden, außer zum Schluss mit einigen Hinweise aus anderen Studien, die darauf hindeuten, dass bei einer strikt nach Leistung vorgenommenen Verteilung auch die Leistungen danach steigen dürften. Allerdings war in den beiden Beiträgen in der FAZ/FAS (Heike Schmoll und Gerald Wagner) zu der Studie, auf die sich die Passagen in dem Entwurf offenbar auch beziehen, auch schon hingewiesen worden. Der Hintergrund ist eine bisher noch nicht veröffentlichte weitere Untersuchung (die der FAZ bekannt war), wieder für die 16 Bundesländer und mit den Daten des NEPS, gerade speziell zu den Leistungen und den Effekten der verschiedenen Regelungen der Implementation der Differenzierung, darunter neben der Verbindlichkeit der Aufteilung nun auch Aspekte der Kontrolle und der Anreize für

das Lehrpersonal, wie die Standardisierung des Stoffs, regelmäßige Evaluationen und die Zentralisierung der Prüfungen. Das Ergebnis war auf der gleichen Linie: Mit der stärkeren Stringenz der Regeln (Verbindlichkeit, Kontrolle, Standardisierung, Zentralisierung) steigt das Leistungsniveau in den betreffenden Ländern. Der Grund ist die Verbesserung der Lern-Effizienz (bei ansonsten gleichen Bedingungen) über die dann stärkere kognitive Homogenität in den Schulen und Schulklassen. Dabei nehmen die Einflüsse der sozialen Herkunft, wie vorher beim Übergang, nicht nur nicht zu, sondern tendenziell auch eher ab, und die leistungsschwächeren Schülerinnen und Schüler scheinen auch eher sogar davon zu profitieren, mindestens aber sind die Vorteile nicht auf die Besseren in den Gymnasien begrenzt. Die Heterogenität bringt, wenn überhaupt, allenfalls den talentierten Schülerinnen und Schüler in Schulklassen mit hohem kognitiven Niveau etwas, also gerade nicht denjenigen, die es nötig hätten. Ich füge die beiden Beiträge für die näheren Informationen bei bzw. zur raschen Übersicht die Titel und die Abstracts (s. unten).

Ansonsten möchte ich mich zu den bildungspolitischen Aspekten hier nicht äußern, insbesondere weil ich nicht weiß, um welche Regelungen es sich in Schleswig-Holstein im Einzelnen handelt. Hinzuzufügen wäre auch noch, dass beide Studien zwar Schleswig-Holstein einbeschlossen haben, es aber keine publizierten Ergebnisse zu einzelnen Ländern gibt (und auch im Regelfall nicht publiziert werden dürfen). Zudem wäre zu beachten, dass sich der Untersuchungszeitraum auf die Jahre 2010 bis 2012 bezog, also noch vor den letzten Änderungen des Schulgesetzes offenbar. Die Befunde lassen sich auch nur auf die abstrakt-theoretischen Effekte der jeweiligen Regelungen beziehen und für die betreffenden Gruppierungen von Ländern. Mit Ergebnissen aus anderen vergleichenden Berichten (wie die regelmäßigen Bildungsberichte oder die IQB-Berichte) sind die Befunde auch nicht zu vergleichen, dort vor allem, weil – anders als hier – nicht alle wichtigen Einflussgrößen auf die Bildungsbeteiligung und die Leistungen erfasst oder nicht analysiert worden sind. Das erklärt u.a. die Abweichung der Befunde hier von den üblichen Berichten, auch denen der OECD, nicht zuletzt in manchen „Sonderauswertungen“ der deutschen Dependance. In den beiden Beiträgen wird darauf ausdrücklich hingewiesen.

Gleichwohl möchte ich meinen persönlichen Eindruck angesichts der zunächst auch für mich einigermaßen erstaunlichen Befunde nicht verschweigen. Es ist, so kann man schon sagen, fast das genaue Gegenteil zu dem, was seit PISA als gängige Meinung gilt. Wenn man aber schon theoretisch nur etwas genauer als üblich hinsieht, ist es nicht mehr sehr erstaunlich, dass bei zu den jeweiligen Fähigkeiten und Interessen besser passenden Curricula und einem darauf eher fokussierbaren Unterricht die Lern-Effizienz steigt bzw. dass vorher beim Übergang mit der stärkeren Orientierung auf die kognitiven Fähigkeiten die Effekte der Klassen- und Standeszugehörigkeiten, eher verringert und die meritokratische Leistungsgerechtigkeit eher verstärkt werden. Dass sich bei den empirischen Befunden diese, lern-psychologisch wie modernisierungstheoretisch gesehen, nicht überraschenden Zusammenhänge dann auch so konsistent zeigen, hätte ich freilich nicht gedacht. Mir ist klar, dass wissenschaftliche Ergebnisse das Eine, die politische Bewertung das Andere sind. Meine Aufgabe sehe ich als Wissenschaftler in der wertneutralen und möglichst objektivierbaren Analyse. Bisher hat es dazu jedoch die geeigneten Daten, man muss es auch fast 20 Jahre nach PISA so sagen, nicht gegeben. Für die nötigen internationalen Vergleiche gibt es sie immer noch nicht. Aber für die deutschen Bundesländer und ihrer doch schon beträchtlichen Varianz in den bildungspolitischen Regelungen haben wir mit der NEPS (endlich) geeignete Daten. Sie stehen jederzeit zu Replikationen zur Verfügung. Auch das sollte beachtet werden, wenn es darum geht, konkrete politische Folgerungen zu ziehen.

Leistungsgerechtigkeit und Bildungsungleichheit.

Effekte der Verbindlichkeit der Grundschulempfehlungen beim Übergang auf das Gymnasium. Ein Vergleich der deutschen Bundesländer mit den Daten der „National Educational Panel Study“ (NEPS)

(erschieden in „Kölner Zeitschrift für Soziologie und Sozialpsychologie, 70, 21018, Heft 3, S.

Zusammenfassung

Der Beitrag untersucht die Frage, ob die Einschränkung der Wahlfreiheit beim Übergang in die Sekundarstufe durch die Verbindlichkeit der Empfehlungen der Schule die Einflüsse der vorherigen Leistungen einerseits und der sozialen Herkunft andererseits verändert. Das Ziel ist die Systematisierung verschiedener Ansätze und Befunde durch ein übergreifendes theoretisches Modell der Effekte der Differenzierung der Bildungswege nach den Leistungen und der empirische Test dieses Modells mit Daten der „National Educational Panel Study“ (NEPS) für die deutschen Bundesländer. Der zentrale Befund lautet: Die Verbindlichkeit verändert die Effekte der sozialen Herkunft auf den Übergang nicht, es verstärkt sich jedoch die Leistungsgerechtigkeit, verstanden als Kopplung des Übergangs an die Leistungen der Kinder in der Grundschule bzw. die kognitiven Fähigkeiten. Für das mit der Differenzierung verbundene Ziel einer Verstärkung der Effizienz des schulischen Kompetenzerwerbs über die kognitive Homogenisierung der Schulen und Schulklassen wären danach nicht die weitere Öffnung und Liberalisierung, sondern die *Stärkung* der Selektivität nach den kognitiven Fähigkeiten und Leistungen der Kinder angeraten, auch um Abweichungen des Übergangs von den Gymnasialempfehlungen nach *unten* zu verhindern – und das gerade für die talentierteren Kinder aus den *unteren* Schichten, die eine solche Empfehlung nach oben tatsächlich erhalten haben.

Kognitive Homogenisierung, Leistungsniveau und Bildungsungleichheit

Theoretische Modellierung und empirische Analyse der Effekte der Leistungsdifferenzierung und des Einflusses der sozialen Herkunft auf die Leistungen in der Sekundarstufe in den deutschen Bundesländern mit den Daten der „National Educational Panel Study“ (NEPS)

(unv. Manuskript, Universität Mannheim 11/2018)

Zusammenfassung

Die Differenzierung der Bildungswege nach den kognitiven Fähigkeiten wird damit begründet, dass die Aufteilung nach den Leistungen eine für das Lernen aller Kinder förderliche Homogenisierung der schulischen Lernumgebung ermögliche. Dagegen wird angeführt, dass die Differenzierung keine besonderen Effekte auf die Leistungen in der Sekundarstufe habe, sondern eher nur die Effekte der sozialen Herkunft verstärke. Diese Auffassung gilt seit Langem als Standardposition. Der Beitrag untersucht auf der Grundlage eines allgemeinen theoretischen Modells die Effekte einer unterschiedlich strikt nach den kognitiven Fähigkeiten vorgenommenen Differenzierung auf die Leistungen in der Sekundarstufe. Grundlage sind Daten der „National Educational Panel Study“ (NEPS) für die 16 deutschen Bundesländer, weil (bisher nur) dafür die theoretisch erforderlichen Konstrukte zur Verfügung stehen. Sechs Befunde lassen sich benennen. Mit der strikten Differenzierung steigt die Leistungsgerechtigkeit beim Übergang, ohne dass sich der Einfluss der sozialen Herkunft ändert. Sie erhöht die kognitive Homogenität der Schulklassen und steigert das Leistungsniveau in der Sekundarstufe über die Verstärkung des Effekts des kognitiven Niveaus in den Schulklassen. Am stärksten ist die Leistungssteigerung für kognitiv homogene Schulklassen bei strikter Differenzierung. Davon profitieren vor allem die leistungsschwächeren Kinder. Die Befunde entsprechen nahezu ausnahmslos den Hypothesen des theoretischen Modells und widersprechen der Standardposition. Das lässt sich damit erklären, dass die theoretisch erforderlichen Daten bisher nicht zur Verfügung standen, insbesondere nicht in den meist verwendeten internationalen Vergleichsstudien.

Gesetz zur Änderung des Schleswig-Holsteinischen Schulgesetzes vom

Der Landtag hat das folgende Gesetz beschlossen:

Artikel 1

Das Schleswig-Holsteinische Schulgesetz in der Fassung vom 04.02.2014 (GVOBl. 2014, S. 21), zuletzt geändert durch Gesetz vom 02.05.2018 (GVOBl. 2018, S. 162) wird wie folgt geändert:

§ 43 Abs. 1 wird wie folgt geändert:

1. a) Satz 2 erhält folgende Fassung:

„Den unterschiedlichen Leistungsmöglichkeiten der Schülerinnen und Schüler kann sowohl durch Unterricht in binnendifferenzierter Form als auch durch Unterricht in nach Leistungsfähigkeit und Neigung der Schülerinnen und Schüler differenzierten Lerngruppen sowie in abschlussbezogenen Klassenverbänden entsprochen werden.“

2. b) Satz 3 wird gestrichen.

Artikel 2

Dieses Gesetz tritt am 31. Juli 2019 in Kraft.

2

Schleswig-Holsteinischer Landtag - 19. Wahlperiode **Drucksache 19/1107 Begründung:**

Mit dem Schulgesetz des Jahres 2007 erfolgte in Schleswig-Holstein die Einführung von Regional- und Gemeinschaftsschulen. Für die Gemeinschaftsschulen war in der ursprünglichen Fassung des § 43 Abs. 1 SchulG dabei vorgesehen, dass ein Unterricht für alle Schülerinnen und Schüler grundsätzlich gemeinsam und vor allem in binnendifferenzierter Form stattfindet. Mit der Novelle des Schulgesetzes vom 28.01.2011 wurde ergänzend festgelegt, dass der Unterricht an Gemeinschaftsschulen sowohl in binnendifferenzierter Form als auch durch Unterricht in nach Leistungsfähigkeit und Neigung der Schülerinnen und Schüler differenzierten Lerngruppen sowie in abschlussbezogenen Klassenverbänden erfolgen kann.

Mit der erneuten Ergänzung des Schulgesetzes vom 22.02.2013 wurde die Möglichkeit, an Gemeinschaftsschulen abschlussbezogene Klassen zu bilden, nur noch für eine Übergangszeit bis zum 31. 07. 2014 eingeräumt. Die umfassende Novelle des Schulgesetzes vom 04.02.2014 bestätigte diesen Ausschluss der Bildung von abschlussbezogenen Klassen. Grundsätzlich ist seitdem an Gemeinschaftsschulen vorrangig ein Unterricht in binnendifferenzierter Form vorgesehen. Als einzige Ausnahme davon gestattet § 43 Abs. 1 S. 3 SchulG die Bildung differenzierter Lerngruppen in einzelnen Fächern ab Jahrgangsstufe 7.

Die mit dem Ausschluss der Bildung von abschlussbezogenen Klassen an Gemeinschaftsschulen verbundenen Erwartungen haben sich seit dem In-Kraft-Treten der neuen Rechtslage am 31.07.2014 nicht erfüllt. Der Mannheimer Soziologe Hartmut Esser stellt fest, dass über die kognitive Homogenisierung der Schüler und Schulklassen die Effizienz des

Kompetenzerwerbs steigt. Zudem würden „die Ungleichheiten nach der sozialen Herkunft in den nach Leistung differenzierenden Systemen nicht nur nicht größer, sondern eher geringer“ (Professor Hartmut Esser, Leistungs- gerechtigkeit und Bildungsungleichheit: Effekte der Verbindlichkeit der Grundschul- empfehlung beim Übergang auf das Gymnasium. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, S. 443).

Esser kommt zu dem Schluss, dass eine Stärkung der Differenzierung nach kogniti- ven Fähigkeiten und Leistungen angeraten sei, wenn man die Effizienz des schuli- schen Kompetenzerwerbs verstärken wolle. Dies gelte auch bei den leistungschwä- cheren Kindern und Kindern aus bildungsfernen Schichten.

Es ist deshalb angezeigt, auch an Gemeinschaftsschulen die Bildung abschlussbe- zogener Klassen in Zukunft wieder zu ermöglichen. Hierzu sieht der Gesetzentwurf eine erneute Änderung des § 43 Abs. 1 SchulG nach Maßgabe der Gesetzesnovelle vom 28.01.2011 vor.

Der Wiedereinführung der Möglichkeit eines Unterrichts in abschlussbezogenen Klassenverbänden soll mit Beginn des Schuljahres 2018/2019 und deshalb zum 31.07.2019 in Kraft treten.

Dr. Frank Brodehl und Fraktion



Leistungsgerechtigkeit und Bildungsungleichheit

Effekte der Verbindlichkeit der Grundschulempfehlungen beim Übergang auf das Gymnasium. Ein Vergleich der deutschen Bundesländer mit den Daten der „National Educational Panel Study“ (NEPS)

Hartmut Esser · Kerstin Hoenig

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2018

Zusammenfassung Der Beitrag untersucht die Frage, ob die Einschränkung der Wahlfreiheit beim Übergang in die Sekundarstufe durch die Verbindlichkeit der Empfehlungen der Schule die Einflüsse der vorherigen Leistungen einerseits und der sozialen Herkunft andererseits verändert. Das Ziel ist die Systematisierung verschiedener Ansätze und Befunde durch ein übergreifendes theoretisches Modell der Effekte der Differenzierung der Bildungswege nach den Leistungen und der empirische Test dieses Modells mit Daten der „National Educational Panel Study“ (NEPS) für die deutschen Bundesländer. Der zentrale Befund lautet: Die Verbindlichkeit verändert die Effekte der sozialen Herkunft auf den Übergang nicht, es verstärkt sich jedoch die Leistungsgerechtigkeit, verstanden als Kopplung des Übergangs an die Leistungen der Kinder in der Grundschule oder die kognitiven Fähigkeiten. Für das mit der Differenzierung verbundene Ziel einer Verstärkung der Effizienz des schulischen Kompetenzerwerbs über die kognitive Homogenisierung der Schulen und Schulklassen wären danach nicht die weitere Öffnung und Liberalisierung, sondern die *Stärkung* der Selektivität nach den kognitiven Fähigkeiten und Leistungen der Kinder angeraten, auch um Abweichungen des Übergangs von den Gymnasialempfehlungen nach *unten* zu verhindern – und das gerade für die talentierteren Kinder aus den *unteren* Schichten, die eine solche Empfehlung nach oben tatsächlich erhalten haben.

H. Esser (✉)
MZES, Universität Mannheim
68131 Mannheim, Deutschland
E-Mail: hartmut-esser@t-online.de

K. Hoenig
Leibniz-Institut für Bildungsverläufe (LifBi)
Wilhelmsplatz 3, 96047 Bamberg, Deutschland
E-Mail: kerstin.hoenig@lifbi.de

Schlüsselwörter Bildungssysteme · Leistungsdifferenzierung · Bildungsungleichheit · Verbindlichkeit · Deutsche Bundesländer

Performance Equity and Educational Inequality

Effects of Binding Recommendations at the Transition to Gymnasium. A Comparison of German Country States by Data of the “National Educational Panel Study” (NEPS)

Abstract The question of the contribution is whether and how the limitation of freedom of choice by binding recommendations for the transition to higher educational tracks after elementary school changes the effects of prior performance on the one hand and of social background on the other. Aim is the systematization of different approaches and findings by a comprehensive theoretical model for effects of ability tracking and its empirical test using data of the of the “National Educational Panel Study” (NEPS) for the different regulations in the German country states. The central finding is that effects of social background are not changed with binding recommendations, but that performance equity, understood as degree of the coupling of attainment with performances in elementary school resp. cognitive abilities of the children, increases. A system that aims to improve educational competencies through cognitive homogenization of schools and school classes should therefore strive to *strengthen* the selectivity by cognitive abilities and achievements instead of trying to open and liberalize regulations, especially in order to prevent deviations in transitions from recommendations for the higher track to a *lower* one—especially for the talented children from *lower* classes with such a recommendation for the higher track.

Keywords Educational systems · Ability tracking · Educational inequality · Binding of recommendations · German country states

1 Einleitung

Die (frühe) Differenzierung der Bildungswege wird damit begründet, dass die Aufteilung nach den Leistungen am Ende der Grundschule eine für das Lernen aller Kinder förderliche Homogenisierung der schulischen Lernumgebung ermögliche (vgl. etwa Sörensen 1970, S. 356 ff.; Betts 2011, S. 343). Dagegen wird angeführt, dass es die erhofften Leistungssteigerungen nicht gebe, und eher (nur) die soziale Ungleichheit im Bildungserfolg verstärkt oder gar erst geschaffen würde (vgl. etwa Oakes 2005 [1985], S. 40 ff.; Betts 2011, S. 343 f.). Tatsächlich weisen die empirischen Befunde der meisten Beiträge dazu in diese Richtung (vgl. für die Bildungsbeteiligung u. a. Pfeffer 2008, S. 555 ff.; Becker und Solga 2012, S. 19 ff.; für die Leistungen u. a.: Wößmann et al. 2009, Tab. 7 A.1; Van de Werfhorst und Mijs 2010, S. 415 ff.). Man könnte es als die – weithin als selbstverständlich geltende – Standardposition bezeichnen (vgl. etwa: Allmendinger 2012, Kap. 10; oder neuerdings noch: Holtappels 2017, S. 16 ff.). Allerdings gab es schon von Beginn an eine Reihe von Hinweisen auf Ausnahmen, Inkonsistenzen und offene Fragen, die

Zweifel hätten wecken können (vgl. Jackson und Jonsson 2013, S. 323 ff.; Marks 2014, Kap. 5; verschiedene Beiträge und Abschnitte bei Blossfeld et al. 2016 oder Hadjar und Gross 2016).

Der Beitrag greift die Frage nach den Folgen der Differenzierung mit Sicht auf eine spezielle Regelung auf: Effekte der *Verbindlichkeit* der Empfehlungen der Schule auf den Übergang in die Sekundarstufe. Zwei Aspekte stehen im Mittelpunkt. Erstens: Der Einfluss der kognitiven Fähigkeiten oder der Leistungen am Ende der Grundschule auf den Übergang, die *Leistungsgerechtigkeit* der Sortierung also. Zweitens: Die *Bildungsungleichheit* als die Stärke des Zusammenhangs zwischen der sozialen Herkunft und dem Übergang. Der Bezug zur allgemeinen Frage liegt auf der Hand: Mit der Verbindlichkeit würden, weil die Eltern nicht mehr selbst über die Schulform entscheiden könnten und das auch unabhängig von den Leistungen auf der Grundschule und den daran in der Regel doch stärker orientierten Noten und Empfehlungen, die Leistungsgerechtigkeit der Sortierung stärker und die Bildungsungleichheit auf das Maß begrenzt, das es überall, also auch schon in integrierten Systemen, gibt, womöglich sogar gesenkt. Geklärt ist auch das bisher nicht: Für alle drei denkbaren Effekte gibt es empirische Hinweise – keine Änderung, Zunahme und Abnahme der sozialen Bildungsungleichheit gegenüber der freien Wahl des Bildungsweges.

Ausgangspunkt sind Befunde aus sechs aktuelleren, auch aufeinander bezogenen Studien für die deutschen Bundesländer (Gresch et al. 2009; Neugebauer 2010; Dollmann 2011; Jähnen und Helbig 2015; Roth und Siegert 2015, 2016; Büchler 2016). Die Anlage und die Kernergebnisse der sechs Beiträge werden in Abschn. 4 beschrieben. Sie sind sehr heterogen und z. T. auch widersprüchlich. Daher werden zunächst in Abschn. 2 ein allgemeiner theoretischer Rahmen vorgestellt, der eine Einordnung der Ansätze und Befunde erlauben soll, und in Abschn. 3 die daraus folgenden Hypothesen an empirischen Erwartungen für den Prozess der Sortierung und den Übergang in die Sekundarstufe zusammengefasst, um die es hier insbesondere geht. Für den empirischen Test des Modells und für den Vergleich mit den sechs Studien werden Analysen mit den Daten der „National Educational Panel Study“ (NEPS) vorgenommen. Die technischen Einzelheiten werden in Abschn. 5 beschrieben. In Abschn. 6 finden sich die empirischen Ergebnisse und in Abschn. 7 gibt es eine kurze zusammenfassende Bewertung.

2 Das Modell der Leistungsdifferenzierung

Ausgangspunkt ist ein übergreifendes theoretisches Modell, das Modell der Leistungsdifferenzierung für die Erklärung der Effekte von Bildungssystemen allgemein (Abb. 1). Das Modell (abgekürzt: MoAbiT¹; vgl. ausführlicher Esser 2016a, 2016b) bezieht sich auf den Verlauf eines typischen Bildungsweges von der kognitiven Entwicklung in der Familie über die Leistungen in der Grundschule, die Noten und die Empfehlungen, den Übergang in die Sekundarstufe oder die Bildungsbeileiligung und schließlich die Leistungen dort. Hier wird das Modell *speziell* mit

¹ Akronym aus dem englischsprachigen Begriff „Model of Ability Tracking“.

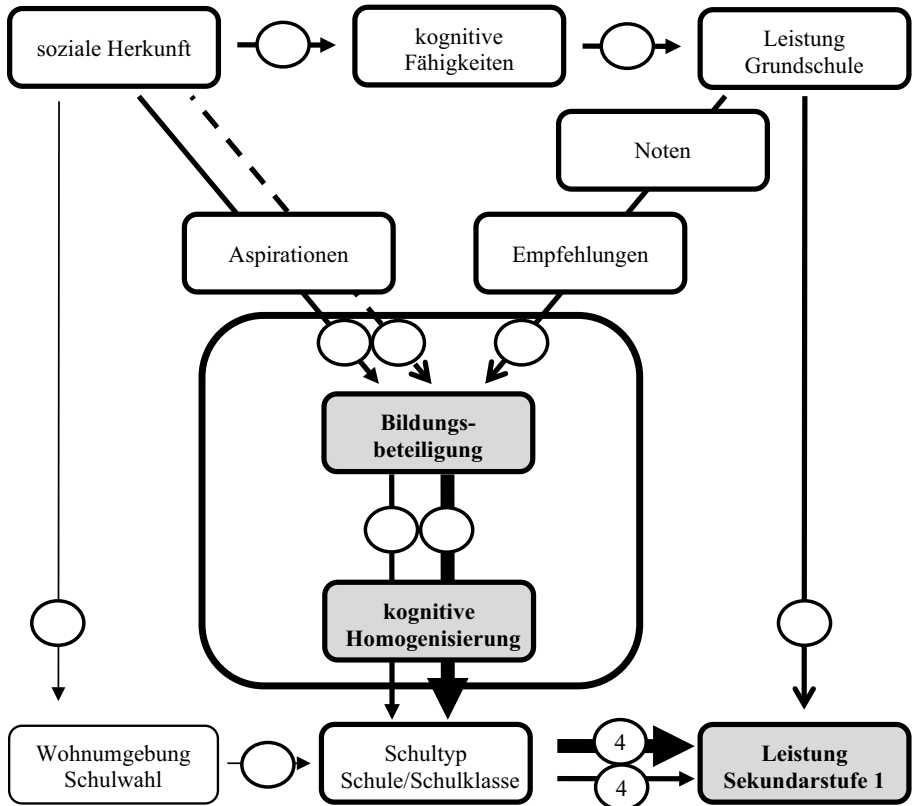


Abb. 1 Das Modell der Leistungsdifferenzierung (MoAbiT) und die Effekte der Verbindlichkeit

Blick auf Effekte der *Verbindlichkeit* für den Einfluss der sozialen Herkunft und der Empfehlungen auf die Bildungsbeteiligung und die daran hängenden Folgen für die kognitive Homogenisierung der Schulen und Schulklassen dargestellt.

Es geht um die Erklärung von Unterschieden in den *Leistungen* in der Sekundarstufe. Der Grundvorgang dafür ist der Kompetenzerwerb durch Lernen. Die Erklärung bezieht sich auf einen gut belegten allgemeinen Zusammenhang: Leistungen sind das Produkt eines bestimmten Ausmaßes an Anregungen („Exposure“) mit der Effizienz, in der diese in ein Lernergebnis umgesetzt werden (Sörensen 1970, S. 357 ff.; Sörensen und Hallinan 1977, S. 275 ff.). Die Anregungen bestehen aus den mit dem jeweiligen Curriculum festgelegten Lerninhalten, die Effizienz hängt hauptsächlich von den kognitiven Fähigkeiten sowie den bis dahin erreichten Leistungen der Kinder ab. Für ein gegebenes Maß an Anregungen steigt danach das Lernergebnis direkt mit den Leistungen zuvor (Pfad 1 in Abb. 1). Die Leistungen in der Grundschule wiederum hängen unmittelbar von den kognitiven Fähigkeiten der Kinder (und anderen Einflüssen in der Grundschule) ab (Pfad 2). Die kognitiven Fähigkeiten sind teilweise genetisch vererbt, teilweise durch die Umstände in den Familien geprägt, speziell über die soziale Herkunft (Pfad 3). Auf die Leistungen in der Sekundarstufe können die kognitiven Fähigkeiten und die soziale Herkunft

neben ihren indirekten Effekten über die Pfade 2 oder 3 auch direkte Effekte haben (in der Abbildung nicht weiter bezeichnet). Die Einflüsse von sozialer Herkunft, den kognitiven Fähigkeiten und den Grundschulleistungen bilden damit die *individuellen Effekte* auf die Leistungen in der Sekundarstufe.

Inhalt und Ausmaß der Anregungen werden für alle Schulen des gleichen Typs zunächst als gleich angenommen. Aber es gibt darüber hinaus Unterschiede zwischen den Schulen, die als *Schuleffekte* einen eigenen Einfluss auf die Effizienz des Lernens und darüber auf die Leistungen haben können (Pfad 4; vgl. u. a. Sörensen und Hallinan 1977). Vier Merkmale sind bei Schuleffekten bedeutsam (in der Abbildung nicht weiter bezeichnet): Das *Curriculum* des jeweiligen Schultyps oder der internen Kurse, die *Qualität* der Schulen in Ausstattung, Unterrichtsstil, Qualifikation und Einsatz des Lehrpersonals, die *Zusammensetzung* der Schulen und Schulklassen nach den *kognitiven Fähigkeiten* und die nach der *sozialen Herkunft* der Kinder.

Die individuellen Zusammenhänge und die Schuleffekte gelten zunächst *allgemein*, auch also in integrierten Bildungssystemen. Es gibt *überall* Effekte der kognitiven Fähigkeiten, der in der Grundschule erworbenen Kompetenzen und der sozialen Herkunft auf das Lernen und auf die Leistungen in der Sekundarstufe. Und es gibt *überall* selektive Verteilungen auf Schulen und Schulklassen nach der sozialen Herkunft und den kognitiven Fähigkeiten, und darüber entstehen auch *überall* unterschiedliche Lernumgebungen nach der sozialen und kognitiven Zusammensetzung sowie nach der Qualität der Schulen und der Qualifikation des Lehrpersonals, etwa als Folge sozial-ökologischer Segregationen bei den administrativen Einzugsgebieten oder einer selektiven Schulwahl der Familien (Pfade a und b links in Abb. 1).

Mit der Differenzierung der Bildungswege kommen Bedingungen und Vorgänge einer auch *institutionell* gezielt organisierten Sortierung in unterschiedliche Schultypen hinzu, die die allgemeinen Zusammenhänge, die individuellen Effekte wie die Schuleffekte, verändern können. Derartige *Änderungen* der allgemeinen Zusammenhänge bilden die *System-Effekte* der Differenzierung auf die verschiedenen Vorgänge und Einflüsse – im Unterschied zur Integration ohne eine solche institutionelle Sortierung. Die System-Effekte würden sich statistisch als *Interaktionseffekte* der allgemeinen Bedingungen mit dem jeweiligen Systemmerkmal zeigen.

Die Differenzierung der Bildungswege nach der Grundschule ist im Kern über zwei Vorgänge bestimmt: Die institutionelle Sortierung durch die *Schule* nach einer *Bewertung* der Leistungen in der Grundschule in Noten und Empfehlungen (Pfad 5) und die – davon gegebenenfalls auch abweichenden – *Entscheidungen* der *Eltern* über den weiteren Bildungsweg (Pfad 6). Beides zusammen ergibt den Übergang in die Sekundarstufe, die *Bildungsbeteiligung*. Zwei System-Effekte werden erwartet, wenn Übergang und Bildungsbeteiligung allein nach den Leistungen erfolgen. Erstens: Eine *strukturell* erzeugte *Homogenisierung* der Schulen und Schulklassen nach den *kognitiven Fähigkeiten* (Pfad 7). Zweitens: Eine durch diese Homogenisierung bewirkte *Verstärkung* des allgemeinen Schuleffekts der *kognitiven Zusammensetzung* der Schulen oder Schulklassen (Pfad 4) gegenüber dem Schuleffekt aus der bloßen sozial-ökologischen Verteilung (Pfad b).

Die Frage ist dann, von welchen weiteren institutionellen Bedingungen und Vorgängen es abhängen könnte, dass bei der Sortierung oder bei der Bildungsbeteiligung die angestrebte kognitive Homogenisierung der Schulen und Schulklassen auch wirklich erreicht wird und nicht andere Einflüsse das überlagern oder auch ganz ersetzen, insbesondere eine leistungsfremde und/oder sozial ungerechte Sortierung nach Status, Einkommen und Bildung der Eltern, die von der Kritik an der Differenzierung so betont wird. Im Anschluss an die bekannte Typologie von Boudon lassen sich drei typisch unterschiedliche Effekte der sozialen Herkunft auf die institutionelle Sortierung und die Bildungsbeteiligung unterscheiden. Die *primären* Effekte bezeichnen den Einfluss der sozialen Herkunft über den Zusammenhang mit den kognitiven Fähigkeiten (und den Grundsulleistungen) aus der vorschulischen Entwicklung in der Familie. Die *sekundären* Effekte beziehen sich auf die von den Fähigkeiten und Leistungen der Kinder auch unabhängigen Bildungsentscheidungen (und andere Interventionen) der Eltern. Die hier sogenannten *tertiären* Effekte verweisen auf die nach der sozialen Herkunft der Kinder gegebenenfalls systematisch verzerrten Bewertungen der Leistungen in Noten und Empfehlungen durch das Lehrpersonal, soziale Diskriminierungen durch die Schule also.

Alle drei Effekte wirken beim faktischen Übergang und der Bildungsbeteiligung zusammen: Die primären Effekte indirekt über die kognitiven Fähigkeiten und die Leistungen in der Grundschule über die Pfade 3, 2 und 5, die sekundären Effekte direkt über die Aspirationen in Pfad 6 und die tertiären Effekte noch einmal über die Vermittlung des Einflusses der Grundsulleistungen über die Noten und die Empfehlungen in Pfad 5.² Der Idealfall einer Bildungsbeteiligung rein nach Leistung wäre dann, dass es, bei gegebenen primären Effekten, keinerlei sekundäre und tertiäre Effekte gäbe: Die Noten transportieren die Informationen über die Grundsulleistungen ohne soziale Verzerrung auf die Bewertungen und die Bildungsentscheidungen der Eltern weichen davon nicht ab. Die nähere Ausgestaltung der institutionellen Regeln der Differenzierung könnte diese beiden Vorgänge entsprechend steuern. Eine stärkere Objektivität der Bewertungen etwa durch eine stärkere Kontrolle und Standardisierung der schulischen Abläufe, auch höhere Anreize und die Verfügung über „objektive“ Informationen über die Potenziale an Fähigkeiten und Leistungen der Kinder einerseits, die Eingrenzung oder gar komplette Unterbindung von Abweichungen der elterlichen Bildungsentscheidungen von den schulischen Empfehlungen andererseits, die Ausschaltung der tertiären und der sekundären Effekte also.

² Soweit entspricht das Modell dem class-performance-outcome-Schema bei Erikson et al. (2005, Abb. 2a, S. 9732) zur Erklärung von sozialen Ungleichheiten in der *Bildungsbeteiligung*. Es fehlen dort allerdings die tertiären Effekte. Daher werden bei den üblich gewordenen Aufteilungen in die Anteile der primären zu den sekundären Effekten die sekundären tendenziell überschätzt (vgl. auch die Kritik von Morgan et al. 2013, S. 296 ff. an dem Erikson et al.-Modell). In einem ähnlich angelegten Konzept bei Maaz und Nagy (2009, Abb. 2, S. 162) werden die tertiären Effekte berücksichtigt, nicht aber, wie auch bei Erikson et al. und schon bei Boudon, die Verbindung zu den strukturellen Änderungen der Schulen und Schulklassen, zu den Schuleffekten und schließlich zu den *Leistungen* in der Sekundarstufe, um die es bei der Differenzierung letztlich geht.

Die *Verbindlichkeit* der Empfehlungen würde in diesem Rahmen dann *speziell* jenen Teil von zusätzlichen institutionellen Regeln der Differenzierung bilden, der die sekundären Effekte begrenzt, jedenfalls im Vergleich zu anderen Regelungen wie die Kontrolle und Standardisierung der schulischen Abläufe. In Abb. 1 ist das über den gebrochenen Pfeil für den Pfad 6' skizziert: Die Eltern und ihre Aspirationen können sich mit der Verbindlichkeit der Empfehlungen nur noch begrenzt auf die Bildungsbeteiligung auswirken. Das hätte nach den Annahmen des MoAbiT im Vergleich zur freien Schulwahl der Eltern weniger Fehlplatzierungen nach den Fähigkeiten und Leistungen und darüber eine stärkere kognitive Homogenisierung zur Folge (dickerer Pfeil für Pfad 7'). Das wiederum verstärkt den Schuleffekt der kognitiven Zusammensetzung der Schulen oder Schulklassen und führt darüber zu einer Steigerung des Leistungsniveaus in der Sekundarstufe (dickerer Pfeil für Pfad 4').

Theoretisch wären die speziell mit der Verbindlichkeit zu erwartenden (System-) Effekte also eindeutig. Empirisch käme es darauf an, dass die auch wirklich „verbindlich“ ist und die sekundären Effekte, bei gegebenen primären und tertiären Effekten, möglichst komplett ausschaltet. Das ist aber weder in der faktischen Implementation noch schon in den institutionellen Regelungen so. Es gibt z. B. in den deutschen Bundesländern mit Verbindlichkeit immer noch Möglichkeiten, bei einer ablehnenden Gymnasial-Empfehlung den Weg nach oben zu finden, was besonders den Familien aus den oberen und bildungsnäheren Familien gelingt. Vor allem aber sind Abweichungen von einer Gymnasial-Empfehlung nach *unten* schon institutionell ohne weiteres möglich, was gerade die *talentierteren* Kinder aus den oft übervorsichtigen *unteren* und den *bildungsferneren* Schichten trifft. Mit einer wirklich strikt geregelten Verbindlichkeit gerade auch gegen Abweichungen nach unten wäre das, so die Theorie, zu vermeiden (vgl. dazu noch Abschn. 5.3).

3 Hypothesen

In Tab. 1 sind die aus dem Modell der Leistungsdifferenzierung folgenden Hypothesen für die allgemeinen Zusammenhänge und die System-Effekte der *Verbindlichkeit* für die Sortierung und den Übergang auf das Gymnasium oder die Bildungsbeteiligung zusammengefasst. Effekte anderer Regelungen oder evtl. mit der Verbindlichkeit empirisch verbundene andere Effekte, wie die Antizipation der Sortierung durch verstärkte Anstrengungen oder auch Interventionen der Eltern oder eine vorausseilende Großzügigkeit bei den Noten und Empfehlungen, sind nicht aufgeführt. Es geht um die kausal *direkt* zu erwartenden Effekte, die sich mit der Verbindlichkeit, wenn sie korrekt implementiert und umgesetzt ist, *unmittelbar* und *zwingend* auf die sekundären Effekte beziehen müssten. Die von den Vorgaben der Schule auch unabhängigen Entscheidungen der Eltern über den Bildungsweg, wenn sie die Wahl haben.

Allgemein sind bei der *Sortierung*, die dem Übergang vorausgeht, die *kognitiven Fähigkeiten* (ABL) direkt von der sozialen Herkunft bestimmt (SES mit + in Spalte 1). Die *Grundschulleistungen* (ACE) hängen direkt von den kognitiven Fähigkeiten ab, aber es gebe auch nach Kontrolle der kognitiven Fähigkeiten weiter Herkunftseffekte, etwa aus den Unterstützungen der Eltern oder auch den Erwar-

Tab. 1 Hypothesen zu den allgemeinen und zu den System-Effekten der Verbindlichkeit der Empfehlungen auf die Sortierung und die Bildungsbeteiligung (Gymnasium gegenüber anderen Bildungswegen) nach den Annahmen des Modells der Leistungsdifferenzierung

Hypothesen MoAbiT	Sortierung Kürzel	1				Übergang	
		2	3	4	5		
		ABL	ACE	MRK	REC	Gymnasium	
<i>Allgemeine Effekte</i>							
Soziale Herkunft	SES	+	+	+	+	(+) (+)	
Kognitive Fähigkeiten	ABL		+	(+)	(+)	(+) (+)	
Leistungen Grundschule	ACE			+	(+)	(+) (+)	
Noten	MRK				+	(+) (+)	
Empfehlungen	REC				+	+	
Aspirationen	ASP					+	
<i>System-Effekte</i>							
Verbindlichkeit	V	0	0	0	0	≤0	
SES*Verbindlichkeit	SES*V	0	0	0	0	≤0	
ABL*Verbindlichkeit	ABL*V		0	0	0	+	
ACE*Verbindlichkeit	ACE*V			0	0	+	

Kürzel für die Bezeichnung der Konstrukte in den folgenden Tabellen und dem weiteren Text; zu den Klammern vgl. den Text

tungen des Lehrpersonals in der Grundschule (SES mit+ in Spalte 2). Die Noten (MRK) folgen unmittelbar den Grundsichulleistungen. Es gebe auch hierfür direkte Herkunftseffekte (SES mit+ in Spalte 3). Analoges gilt für die Empfehlungen (REC, Spalte 4), wobei die kognitiven Fähigkeiten auf die Noten nur indirekt wirken, ebenso wie die Leistungen in der Grundschule auf die Empfehlungen (mit+ jeweils in Klammern gesetzt).

System-Effekte der Verbindlichkeit als Änderung dieser allgemeinen Zusammenhänge bei der Sortierung gebe es nicht (alle in den Spalten 1 bis 4 auf die Verbindlichkeit bezogenen Effekte mit 0): Die kognitiven Fähigkeiten und die primären Effekte sind über die Bildungsregimes gleich verteilt und die Bewertungen und die tertiären Effekte von anderen Regelungen des Bildungssystems als der Verbindlichkeit abhängig. Das gilt auch für die Herkunftseffekte: Die Verbindlichkeit der Empfehlungen beeinflusst (unmittelbar) nur die sekundären Effekte, aber davon ist erst der faktische Übergang betroffen, wenn es die Leistungen und die Bewertungen schon gibt.

Für den *Übergang* (auf das Gymnasium; Spalte 5) sehen die *allgemeinen* Effekte – mutatis mutandis – ansonsten ähnlich aus. Einen direkten Einfluss haben die Empfehlungen der Schule und, soweit es institutionell möglich ist, die Aspirationen der Familien. Die Bedingungen und Vorgänge der Sortierung vorher wirken sich nach dem Modell allerdings jeweils nur indirekt aus (Effekte+ jeweils in Klammern), und wenn es doch Effekte darüber hinaus gibt, dann wären zusätzliche Einflüsse anzunehmen (vgl. Fußnote 3). Für die Herkunftseffekte auf den Übergang wird, nun anders als für die Sortierung, angenommen, dass sie mit der statistischen Kontrolle der Sortierung vorher und mit der Berücksichtigung der Aspirationen komplett

erklärt werden, es also nur indirekte Effekte aus der Sortierung vorher gibt (SES mit + in Klammern). Zeigen sich empirisch dennoch Effekte der Herkunft, dann verweist das ebenfalls auf zusätzliche Einflüsse.

Weil mit der Verbindlichkeit gerade die sekundären Effekte begrenzt werden, wird als *System-Effekt* auf den Übergang oder auf die Bildungsbeteiligung ein eher negativer (konditionaler Haupt-) Effekt angenommen (V mit ≤ 0 in Spalte 5): Wenn es bei freien Schulwahl eher Abweichungen von den Empfehlungen nach oben gibt und die mit der Verbindlichkeit unterbunden würden, muss die Rate der Übergänge in den höheren Bildungsweg sinken, mindestens aber sollte sie mit der Verbindlichkeit nicht steigen. Zwei unmittelbar daraus folgende (Interaktions-)Effekte werden erwartet. Erstens: Die Zunahme der *Leistungsgerechtigkeit* beim Übergang auf das Gymnasium oder bei der Bildungsbeteiligung über die nun engere Kopplung an die kognitiven Fähigkeiten oder an die Leistungen in der Grundschule (ABL*V bzw. ACE mit V). Zweitens: Keine Zunahme (oder sogar eher eine Verringerung) der sozialen *Bildungsungleichheit* (SES*V: ≤ 0). Es sind die beiden Grundhypothesen für die Effekte der Verbindlichkeit.

4 Sechs Studien

Zu den Effekten der Verbindlichkeit im Vergleich der deutschen Bundesländer gibt es schon länger eine Reihe von vereinzelt und auch kontroversen Befunden, was u. a. daran lag, dass es die geeigneten Daten für eine Identifikation der betreffenden Effekte kaum gab. Inzwischen liegen einige Untersuchungen vor, bei denen das anders ist. Sie schließen an die älteren Beiträge und auch aneinander an.

Eine der ersten, im Vergleich der Bundesländer breiter angelegten Studien stammt von Gresch et al. (2009). Erfasst werden die soziale Herkunft und die Empfehlung am Ende der Grundschule. Die Studie berücksichtigt alle Bundesländer mit Ausnahme derer, in denen der Übergang erst nach der 6. Klasse erfolgt (BE, BB, MV).³ In der multivariaten Analyse zeigen sich allgemein starke positive Effekte der Empfehlung und der sozialen Herkunft (Gresch et al. 2009, Tab. 3, Modell 5, S. 247). Bei den *System-Effekten* gibt es eine merkliche Absenkung der Übergangswahrscheinlichkeit, aber keine Änderung des Einflusses der sozialen Herkunft und auch keine der Empfehlungen (vgl. auch noch Abschn. 6.3 unten).

Die Datengrundlage der Analyse von Neugebauer (2010) ist das DJI-Kinderpanel über den Zeitraum von 2002 bis 2005. Für die allgemeinen Bedingungen werden die soziale Herkunft und die Noten berücksichtigt. Die Bestimmung der Effekte der sozialen Herkunft folgt dem Ansatz von Erikson et al. (2005, S. 9731f.) im Vergleich der relativen Bedeutung der sekundären Effekte zu den primären Effekten der sozialen Herkunft. Die Noten dienen als Proxy-Variable für die Leistungen in der Grundschule. Kontrollvariablen wurden weiter nicht berücksichtigt. Allgemein kommt ein Anteil der sekundären zu den primären Effekten von 59% heraus

³ BB Brandenburg, BW Baden-Württemberg, BY Bayern, HB Bremen, HE Hessen, HH Hamburg, MV Mecklenburg-Vorpommern, NI Niedersachsen, NW Nordrhein-Westfalen, RP Rheinland-Pfalz, SL Saarland, SN Sachsen, ST Sachsen-Anhalt, SH Schleswig-Holstein, TH Thüringen.

(Neugebauer 2010, Tab. 2, S. 208). Für den *System-Effekt* beträgt der Anteil der sekundären Effekte 61 % bei der freien Wahl, wird also etwas größer, und er sinkt bei Verbindlichkeit auf 54 % (Neugebauer 2010, Tab. 3, S. 209). Allerdings ist der Effekt statistisch nicht signifikant und auch die Übergangsraten insgesamt zwischen den beiden Regimes von Verbindlichkeit und freier Wahl unterscheiden sich nicht weiter.

Die Grundlage der Analyse von Dollmann (2011) bilden zwei Übergangskohorten (2005/2006 und 2006/2007) in Köln. Die (zufällige) politisch bedingte Umstellung von der freien Wahl auf die Verbindlichkeit zwischen den beiden Kohorten ließ sich als natürliches Experiment für einen kausalen Effekt der Verbindlichkeit nutzen. Als allgemeine Bedingungen wurden die kognitiven Fähigkeiten, die Leistungen in der Grundschule, die soziale Herkunft und die Aspirationen der Eltern aufgenommen. Die Empfehlungen sind nicht erhoben worden, wohl aber die Noten, die jedoch in den Analysen nicht weiter berücksichtigt wurden. Allgemein haben die kognitiven Fähigkeiten und die Leistungen in der Grundschule die stärksten Einflüsse auf den Gymnasialübergang und es gibt auch hier nach allen Kontrollen einen direkten Effekt der sozialen Herkunft (Dollmann 2011, Tab. 3, Modell 2, Übergang G vs R, S. 607 ff.). Als *System-Effekt* vermindert sich der Einfluss der sozialen Herkunft deutlich. Der konditionale (Haupt-)Effekt der Verbindlichkeit ist positiv. Die Bedeutung der Aspirationen der Eltern nimmt bei Verbindlichkeit ab (Dollmann 2011, Tab. 7, Anhang Materialien). In einer kombinierten Betrachtung des Einflusses der Leistungen, der sozialen Herkunft und der Verbindlichkeit kommt noch heraus, dass die weniger talentierten Kinder aus den oberen Bildungsschichten durch die Verbindlichkeit beim Gymnasialübergang tatsächlich gebremst, die Kinder aus den unteren Schichten mit überdurchschnittlichen Leistungen zum Übergang hingegen eher ermuntert werden (Dollmann 2011, S. 610 ff. und Abb. 1, S. 611).

Jähnen und Helbig (2015) untersuchen die Effekte verschiedener Reformen zur Einführung und zur Abschaffung der Verbindlichkeit über einen Zeitraum von 1949 bis 2010 im Vergleich zu den jeweiligen allgemeinen Trends. Insgesamt wurden 13 Reformen erfasst. Für die Effekte der Abschaffung der Verbindlichkeit wurden die Reformen in Niedersachsen von 1978/1979 und in Rheinland-Pfalz von 1984/1985 betrachtet, für die Einführung der Verbindlichkeit die in Nordrhein-Westfalen 2006/2007 und in Brandenburg von 2007/2008. Bei den allgemeinen Bedingungen für die Bildungsbeteiligung wurde nur die soziale Herkunft berücksichtigt. Der Effekt ist in allen Fällen positiv. Bei den *System-Effekten* zeigte sich – mit einigen Abweichungen – für die Abschaffung der Verbindlichkeit eine höhere, für die Einführung eine geringere Gymnasialquote (Jähnen und Helbig 2015, Abb. 1, S. 558). Bei den Änderungen des Herkunftseffekts gab es in beiden Fällen der Abschaffung der Verbindlichkeit (Niedersachsen 1978/79 und Rheinland-Pfalz 1984/1985) ein Absinken der Bildungsbeteiligung bei den bildungsnäheren Familien und ein Ansteigen bei den bildungsferneren Familien (Jähnen und Helbig 2015, Tab. 4, S. 560, Tab. 5, S. 561). Die Ergebnisse zur Einführung der Verbindlichkeit sind für die beiden betrachteten Fälle unterschiedlich. In Nordrhein-Westfalen ändert sich nach der Reform von 2006/2007 nichts weiter (Jähnen und Helbig 2015, Tab. 6, S. 562), während in Brandenburg nach 2008/2009 die Bildungsbeteiligung

der Kinder aus den bildungsnäheren Schichten sinkt und die aus den bildungsferneren ansteigt (Jähnen und Helbig 2015, Tab. 7, S. 563).

Die beiden Beiträge von Roth und Siegert (2015, 2016) beziehen sich auf der Grundlage verschiedener Mikrozensen ebenfalls auf längere Zeiträume. Berücksichtigt werden in dem Beitrag von 2015 die sieben (westdeutschen) Bundesländer BW, BY, HE, NI, NW, RP und SH und in dem von 2016 alle (außer SL), jeweils auch hier nicht die mit einer längeren Grundschuldauer oder schulformunabhängigen Orientierungsstufe (BB, BE und MV). Als allgemeine Bedingung wird nur die soziale Herkunft erfasst. Die Vergleiche des Beitrags von 2015 beziehen sich auf Kinder im Alter von 13 und 14 Jahren, also anders als sonst meist in der Mitte der Sekundarstufe 1, in dem Beitrag von 2016 wurden auch Kinder im Alter von 11, 12 und 15 Jahren erfasst. Drei Analysen werden vorgenommen: Ein Vergleich der Herkunftseffekte für die beiden über Zeit verbindlich sortierenden Bundesländer BW und BY gegenüber den durchgängig unverbindlichen HE, RP und SH, für die Abschaffung der Verbindlichkeit in NI im Jahr 1979/1980 und für die Einführung der Verbindlichkeit in NW im Jahr 2006/2007. Die soziale Herkunft hat danach allgemein stets einen deutlichen positiven Effekt. In keinem der drei Vergleiche gab es als *System-Effekt* der Verbindlichkeit eine signifikante Änderung.

Die Untersuchung von Büchler (2016) nutzt das Sozio-oekonomische Panel (SO-EP). Damit konnte ebenfalls ein längerer Zeitraum (von 1984 bis 2010) verfolgt werden. Von den allgemeinen Bedingungen werden die soziale Herkunft und die Empfehlungen berücksichtigt. Dafür finden sich jeweils positive Effekte (Büchler 2016, Tab. 2, Modell 3, S. 76). Als *System-Effekt* wird für die freie Wahl (in beiden Varianten) eine Zunahme der Gymnasialbeteiligung gefunden (ebd). Diese Zunahme ist für die Kinder aus den *unteren* Schichten jedoch stärker (Büchler 2016, Abb. 3, S. 77; nach den Ergebnissen aus Tab. 2, Modell 3, S. 77). Mit der Verbindlichkeit nimmt demnach die soziale Bildungsungleichheit zu.

In Tab. 2 sind die Ansätze und die Befunde aus den sechs Studien untereinander und im Vergleich mit den Hypothesen des MoAbiT zusammengefasst.

Die Wirkungen der allgemeinen Bedingungen weisen bis auf den SES-Effekt alle in den sechs Studien in die vom MoAbiT erwartete Richtung: Kognitive Fähigkeiten, die Leistungen in der Grundschule, die Noten, die Empfehlungen und die Aspirationen unterstützen – unterschiedlich stark – den Übergang auf das Gymnasium. Abweichend von der Hypothese des MoAbiT bleibt in allen sechs Fällen neben den indirekten Effekten aus der Sortierung ein z. T. deutlicher direkter SES-Einfluss auf den Übergang (mit ^a gekennzeichnet). Zu beachten ist dabei allerdings, dass in keiner der sechs Studien alle nach dem MoAbiT relevanten Bedingungen kontrolliert wurden.

Für den *System-Effekt* sind die Ergebnisse dagegen sehr inkonsistent: In drei Beiträgen wird kein Effekt der Verbindlichkeit gefunden (Beiträge 1, 2 und 5), einmal eine Verringerung (Beitrag 3) und einmal eine Zunahme (Beitrag 6). Der Beitrag 4 erbringt alle drei Effekte: negativ, neutral, positiv, teilweise in Widerspruch zu anderen Studien für das gleiche Land. Die *System-Effekte* einer Verstärkung des Herkunftseinflusses stehen in Widerspruch zur MoAbiT-Hypothese (Beiträge 4 und 6; mit ^a gekennzeichnet). *System-Effekte* der Verbindlichkeit auf die *Leistungsgerechtigkeit* (ABL/ACE*V) wurden in keiner der Studien geprüft. In der Studie von Gresch

Tab. 2 Übersicht über die Ansätze und Befunde der sechs Studien im Vergleich mit den Hypothesen des MoAbiT

	MoAbiT	Gresch et al. (2009)	Neuge- bauer (2010)	Doll- mann (2011)	Jähnen und Helbig (2015)	Roth und Siegert (2015, 2016)	Büchler (2016)
		1	2	3	4	5	6
<i>Allgemeine Effekte</i>							
SES	(+)	+ ^a	+ ^a	+ ^a	+ ^a	+ ^a	+ ^a
ABL	(+)			+			
ACE	(+)			+			
MRK	(+)		+				
REC	+	+					+
ASP	+			+			
<i>System-Effekte</i>							
V	≤0	-	...	+	(-)	...	0
SES*V	≤0	0	0	-	-	0	+ ^a
ABL/ACE*V	+	(0)

Zu den Klammern vgl. den Text

SES soziale Herkunft, ABL kognitive Fähigkeiten, ACE Leistungen Grundschule, MRK Noten, REC Empfehlungen, ASP Aspirationen, V System-Effekt Verbindlichkeit

... nicht analysiert

^aAbweichung von Hypothesen des MoAbiT

et al. (2009) wurde allerdings die Interaktion von Empfehlung und Verbindlichkeit berücksichtigt, die sich als proxy für die Interaktion der Leistungen in der Grundschule mit der Verbindlichkeit verstehen lässt. Dafür findet sich jedoch kein Effekt (Tab. 2, Spalte 1, ABL/ACE*V mit 0; in Klammern). Die berichteten konditionalen (Haupt-)Effekte der Verbindlichkeit lassen sich kaum vergleichen, auch weil sie teilweise nicht mit analysiert oder über andere Quellen ermittelt wurden. Sie werden deshalb in den Vergleich mit den Hypothesen des MoAbiT nicht mit einbezogen.

5 Daten, Variablen, Analysen

Grundlage der Analysen zum Test des Modells sind Daten der „National Educational Panel Study“ (NEPS)⁴ für die Startkohorte 3, die im Schuljahr 2010/2011 in Deutschland die fünfte Klasse besuchten, mit Angaben aus der ersten Welle mit einer schriftlichen Befragung und Kompetenztestung im Herbst 2010 und einer telefonischen Befragung jeweils eines Erziehungsberechtigten im Frühjahr 2011.

⁴ <https://doi.org/10.5157/NEPS:SC3:5.0.0> Die Daten des NEPS wurden von 2008 bis 2013 als Teil des Rahmenprogramms zur Förderung der empirischen Bildungsforschung erhoben, welches vom Bundesministerium für Bildung und Forschung (BMBF) finanziert wurde (siehe auch Blossfeld et al. 2011). Seit 2014 wird NEPS vom Leibniz-Institut für Bildungswissenschaften e. V. (LifBi) an der Otto-Friedrich-Universität Bamberg in Kooperation mit einem deutschlandweiten Netzwerk weitergeführt.

Tab. 3 Wertebereich, Mittelwerte (*av*) und Streuungen (*sd*) der verwendeten Konstrukte

Konstrukte	min/max	Gesamt		Freie Wahl		Verbindlich	
		av	sd	av	sd	av	sd
SES	16–90	53,06	16,77	53,43	16,98	52,87	16,66
ABL	0–1	0,59	0,21	0,59	0,21	0,59	0,21
ACE	–2,9–3,3	0,14	1,01	0,03	0,99	0,19	1,01
MRK	1–6	4,68	0,76	4,59	0,77	4,73	0,76
REC	0–1	0,57	0,50	0,55	0,50	0,58	0,49
ASP	0–1	0,74	0,44	0,78	0,42	0,72	0,45
GYM	0–1	0,55	0,50	0,54	0,50	0,56	0,50
V	0–1	0,66	0,47				
FEM	0–1	0,47	0,50	0,51	0,50	0,45	0,50
MHG	0–1	0,31	0,56	0,27	0,44	0,33	0,47
VSB	0–83	39,18	11,05	38,36	10,55	39,61	11,06
N		3032		1032		2000	

SES soziale Herkunft, *ABL* kognitive Fähigkeiten, *ACE* Leistungen Grundschule, *MRK* Noten, *REC* Empfehlungen, *ASP* Aspirationen, *V* System-Effekt Verbindlichkeit, *FEM* Geschlecht, *MHG* Migrationshintergrund, *VSB* Vorschulbesuch

5.1 Die Variablen

Eine Übersicht der verwendeten Variablen und deren Kennzahlen finden sich in Tab. 3. Die Aufstellung folgt der Reihenfolge der Konstrukte des Modells der Leistungsdifferenzierung nach Tab. 1 oben.

Der *soziale Hintergrund* (SES) wird als der höchste Status bei den Eltern, gemessen über den Index ISEI-88 (nach Ganzeboom et al. 1992) erfasst.⁵ Als Maß für die *kognitive Entwicklung* (ABL) wird der für NEPS entwickelte NEPS-MAT verwendet, ein Matrizen-Test des schlussfolgernden Denkens (Lang et al. 2014). Die Leistungen in der Grundschule (ACE) wurden zu Beginn der fünften Klasse über die Kompetenzen in Orthographie, Leseverständnis und Mathematik erfasst. Die *Noten* (MRK) werden über den Mittelwert der von den Eltern berichteten Noten in Deutsch und Mathematik am Ende der vierten Klasse gemessen. Der Notendurchschnitt wurde zur einfacheren Vorzeicheninterpretation von ungenügend (Schulnote 6 = 1) bis sehr gut (Schulnote 1 = 6) umkodiert. Die *Empfehlung* am Ende der Grundschulzeit (REC) wurde auf Basis der Angaben der Eltern als Dummy (Gymnasium mit 1 vs. kein Gymnasium mit 0) gebildet. Die drei spät sortierenden Länder BB, BE und MV wurden von den Analysen ausgeschlossen. Zur Erfassung der familiären *Aspirationen* (ASP) wurden die Angaben der Kinder in der ersten Welle (Klasse 5) für deren idealistische Aspirationen auf das Gymnasium genommen. Weil die Angaben dazu nach dem Übergang gemacht wurden, sind die Effekte womöglich überschätzt. Das wird bei der Interpretation zu berücksichtigen sein. Der Index dient als Proxy

⁵ Zum Vergleich wurden auch Analysen mit dem höchsten Bildungsstand in der Familie nach dem CAS-MIN-Index vorgenommen. In den wesentlichen Strukturen gab es keine Unterschiede, ob die Bildung der Eltern alternativ oder simultan mit dem Status berücksichtigt wurde. Zur Vermeidung von Überkontrollen wurde in den Analysen nur der HISEI-Index verwendet.

auch für die Aspirationen der Eltern, insofern die Einschätzungen zwar nicht perfekt übereinstimmen, aber, wie auch sonst zu finden ist, stark korrelieren. Die Ergebnisse änderten sich bei unterschiedlichen Verwendungen der Angaben so gut wie nicht. Der *Übergang* auf das Gymnasium (GYM) oder die Bildungsbeteiligung wird als Dummy der zu Beginn der fünften Klasse besuchten Schulform (Gymnasium mit 1 vs. kein Gymnasium mit 0) operationalisiert, wobei Kinder, die den Gymnasialzweig einer kooperativen Gesamtschule besuchen, ebenfalls als Gymnasiasten klassifiziert sind. Die *Verbindlichkeit* (V) als System-Merkmal wurde nach der Aufstellung bei Büchler (2016, Abb. 1, S. 64) vorgenommen mit BW, BY, NW, SN, ST und TH als „verbindlich“ (Wert mit 1 gegen 0 bei freier Wahl).

In allen Modellen wird für *Geschlecht* des Kindes (FEM), den *Migrationshintergrund* (MHG, mit 1 gegen 0) und die Dauer des *Vorschulbesuchs* vor der Einschulung in Monaten (VSB) kontrolliert, die Einzelergebnisse dazu jedoch aus Platzgründen nicht eigens ausgewiesen.

Die Verteilungen sind nicht weiter auffällig. Allenfalls ließe sich erwähnen, dass in den Bundesländern mit Verbindlichkeit die Leistungen in der Grundschule höher, die Noten besser, die Empfehlungen großzügiger, die Aspirationen niedriger und die Übergänge häufiger sind als bei der freien Wahl, alles jeweils aber nur leicht. Rückschlüsse auf Kausalitäten wären damit natürlich ohnehin nicht möglich.

5.2 Analysesample und Modellierung

Das Analysesample umfasst die Fälle mit gültigen Angaben für alle Modellvariablen (listwise deletion). Kinder aus den drei später sortierenden Bundesländern (BE, BB, MV) sowie Kinder an Förderschulen sind von den Analysen ausgeschlossen. Für die verbleibenden Bundesländer sind vollständige Angaben für 3032 Kinder verfügbar. Dies entspricht 60 % der Gesamtstichprobe von 5050 Kindern in diesen Bundesländern (ohne Förderschulen). Wichtigster Grund für die hohe Anzahl fehlender Werte ist die Nichtteilnahme der Eltern am freiwilligen Elterninterview. 1569 (78 %) der Fälle mit fehlenden Werten gehen auf diese Ursache zurück. Für diese Kinder ist eine Imputation der fehlenden Informationen aus dem Elterninterview zum sozialen Hintergrund sowie zur Grundschulempfehlung nicht sinnvoll, da diese lediglich auf den Angaben aus den Kinderfragebögen beruhen würde, die keine geeigneten Informationen zur Imputation dieser Variablen enthalten. Die verbleibenden 449 Fälle mit fehlenden Werten ergeben sich aus Missings auf den verbleibenden Variablen. Verschiedene andere Verfahren der Kontrolle der Robustheit der Ergebnisse, fallweiser Ausschluss oder die Mittelwertersetzung der fehlenden Werte, erbrachten keine substanziellen Unterschiede in den Befunden, auch nicht zu Analysen mit einem vergleichbaren Datensatz (dem BiKS-Projekt) für Unterschiede zwischen Bayern und Hessen, bei dem multiple Imputationen möglich waren und zur Kontrolle vorgenommen wurden (Esser und Relikowski 2015, S. 17).

Für die Bestimmung der Effekte wurden bei metrischen Variablen OLS-Regressionen, bei kategorialen Variablen logistische Regressionen berechnet. Zur Interpretation der Interaktionseffekte mit der Verbindlichkeit in den logistischen Regressionen wird auf die Berechnung konditionaler durchschnittlicher Marginaleffekte (Average Marginal Effects, AME) zurückgegriffen und die Ergebnisse in Grafiken

der vorhergesagten Wahrscheinlichkeiten sowie der AME präsentiert. Zum Vergleich der Koeffizienten für genestete Logit-Modelle zur Erklärung des Gymnasialbesuchs wird das KHB-Verfahren (Karlson et al. 2012; Kohler et al. 2011) angewendet. Zum Vergleich von separat geschätzten Modellen für Länder mit und ohne Verbindlichkeit werden die AME als Maß der Effektstärken verwendet.

6 Ergebnisse

Die empirischen Analysen beziehen sich auf die Hypothesen des MoAbiT zur Erklärung der Bildungsbeteiligung (Tab. 1 in Abschn. 3) und auf die Befunde aus den sechs Studien zu den Effekten der Verbindlichkeit (Tab. 2 in Abschn. 5). Zunächst wird der Vorgang der Sortierung von der kognitiven Entwicklung in der Familie bis zu den Empfehlungen am Ende der Grundschule dokumentiert (Abschn. 6.1), dann der Übergang in die Sekundarstufe (Abschn. 6.2). Anschließend werden die

Tab. 4 Kognitive Entwicklung, Leistungen in der Grundschule, Noten und Empfehlungen (Spalten 1–3: OLS-Koeffizienten; Spalte 4: Logit-Koeffizienten; Geschlecht, Migrationshintergrund und Vorschulbesuch kontrolliert; Standardfehler unter Berücksichtigung der Clusterung nach Schulklassen)

	Kognitive Fähigkeiten		Leistungen Grundschule		Noten Grundschule		Empfehlung Gymnasium	
	1		2		3		4	
	H	B	H	B	H	B	H	B
<i>Allgemeine Effekte</i>								
SES	+	0,002***	+	0,014***	+	0,007***,a	+	0,023***,a
ABL			+	2,288***	(+)	0,323***	(+)	1,032***
ACE					+	0,411***	(+)	1,609***
MRK							+	... ^b
<i>System-Effekte</i>								
V	0	0,006	0	0,130	0	0,173	0	-0,116
SES*V	0	0,000	0	0,001	0	-0,002	0	0,003
ABL*V			0		0	... ^b	0	... ^b
ACE*V					0	0,034 ^a	0	-0,250 ^a
C		0,432***		-2,124***		3,913***		-1,876***
(Psd-)R ²		0,05		0,35		0,46		0,30
N	3032		3032		3032		3032	

Für die Kürzel bei den Effekten vgl. Tab. 1 oben

H Hypothese MoAbiT, B Befund NEPS, SES soziale Herkunft, ABL kognitive Fähigkeiten, ACE Leistungen Grundschule, MRK Noten, REC Empfehlungen, ASP Aspirationen, V System-Effekt Verbindlichkeit
* $p \leq 0,05$, ** $p \leq 0,01$, *** $p \leq 0,001$

^aAbweichung von Hypothesen des MoAbiT

^bWeil die Empfehlungen in einigen Bundesländern schon institutionell mit den Noten verbunden sind, werden zur Vermeidung von Überkontrollen die Noten bei den Analysen der System-Effekte auf die Empfehlungen nicht mit aufgenommen. Analysen mit Berücksichtigung der Noten zeigen, wie zu erwarten, dass ein Teil der Effekte der Leistungen über die Noten vermittelt wird, ergaben darüber hinaus aber keine substanziellen Änderungen. Bei den System-Effekten für die Leistungsgerechtigkeit werden entsprechend nur die Leistungen in der Grundschule und nicht gleichzeitig auch die kognitiven Fähigkeiten berücksichtigt.

Befunde zur Replikation der sechs Studien über die Daten des NEPS zusammengefasst und einigen verstreuten Hinweisen in verschiedenen Beiträgen nachgegangen, dass besonders die Kinder aus den unteren Schichten durch die Verbindlichkeit beim Übergang benachteiligt würden (Abschn. 6.3).

6.1 Sortierung

Tabelle 4 enthält die Ergebnisse zur Erklärung der Vorgänge hin zur *Sortierung* auf die verschiedenen Schulformen nach der Grundschule: Die kognitive Entwicklung, die Leistungen in der Grundschule, die Noten und die Empfehlungen.

Bei den allgemeinen Bedingungen bestätigt sich zunächst, dass die frühe *kognitive Entwicklung* deutlich vom sozialen Hintergrund bestimmt ist (Tab. 4, Spalte 1). Das ändert sich unter Kontrolle der Verbindlichkeit nicht. Bei den *Leistungen in der Grundschule* ist es ähnlich (Spalte 2): Einen starken Einfluss haben die kognitiven Fähigkeiten, aber es verbleibt auch ein direkter Effekt der sozialen Herkunft. Mit der Verbindlichkeit ändert sich praktisch nichts. Die *Noten* sind stark von den Leistungen in der Grundschule abhängig und sowohl indirekt wie auch noch einmal unmittelbar von den kognitiven Fähigkeiten (Spalte 3). Es finden sich deutliche direkte Herkunftseffekte. Mit der Verbindlichkeit ändert sich als *System-Effekt* daran wieder nichts, aber auch nicht an der Kopplung der Noten an die Leistungen in der Grundschule, was man nach dem MoAbiT hätte erwarten können (ACE*V mit 0,034, n. s., mit ^a gekennzeichnet). Die *Empfehlungen* schließlich folgen eng den Leistungen in der Grundschule, sind aber auch direkt noch von den kognitiven Fähigkeiten und den Grundschulleistungen abhängig (Spalte 4). Wiederum haben die Kinder aus den oberen Schichten bei gleichen Leistungen Vorteile und das auch, wenn alles andere statistisch kontrolliert ist. Im Vergleich der *System-Effekte* gibt es, auch über die AMEs, keine signifikanten Änderungen, jedoch fällt die Kopplung der Empfehlungen bei Verbindlichkeit an die Grundschulleistungen gegen die Hypothese des MoAbiT geringer aus (ACE*V mit -0,250, n. s., mit ^a gekennzeichnet), wengleich nicht signifikant.

Insgesamt lassen sich damit für die Sortierung vier Befunde festhalten. Erstens: Die *allgemeinen* Herkunftseffekte werden *nicht* komplett mit der jeweiligen Vorgeschichte der Sortierung erklärt. Zweitens: Die Herkunftseffekte haben im Vergleich zu den leistungsbezogenen Einflüssen (nach den in den Ergebnissen oben nicht weiter dokumentierten Anteilen an jeweils erklärter Varianz) eine erheblich geringere Bedeutung. Drittens: Es gibt mit der Verbindlichkeit so gut wie *keine* Änderungen der allgemeinen Zusammenhänge. Viertens: Mit der Verbindlichkeit sind die Noten nicht enger und die Empfehlungen sogar *weniger* an die Leistungen gekoppelt, wengleich das nicht signifikant.

6.2 Übergang

Die Tab. 5 und 6 geben die Befunde zum faktisch vollzogenen Übergang auf das Gymnasium wieder: Tab. 5 die jeweiligen logistischen Koeffizienten, korrigiert nach der KHB-Methode, Tab. 6 die AMEs für die Interaktionseffekte.

Wie zu erwarten gibt es allgemein einen starken Brutto-Effekt der sozialen Herkunft auf den Gymnasialübergang (SES mit 0,094, Modell 5a in Tab. 5). Der verringert sich erheblich allein schon mit der Kontrolle der kognitiven Fähigkeiten und der Leistungen in der Grundschule (Modell 5b). In den Modellen 5c bis 5e werden der *System-Effekt* der Verbindlichkeit sowie dessen Interaktion mit der sozialen Herkunft und den Leistungen hinzugefügt, in 5d zusätzlich unter Kontrolle der Empfehlungen und in 5e auch der Aspirationen. Zur Interpretation der Effekte sei zusätzlich auf Tab. 6 sowie Abb. 2 verwiesen. Tabelle 6 zeigt den AME für die soziale Herkunft und die Grundschulleistungen in Abhängigkeit von der Verbindlichkeit sowie jeweils die Differenz. Diese kann als Maß für die Stärke und Signifikanz der Interaktionen der sozialen bzw. der Grundschulleistungen mit der Verbindlichkeit verwendet werden.

Der Effekt der Verbindlichkeit ist in Modell 5c wie erwartet negativ, jedoch nicht signifikant. Ebenfalls nicht signifikant ist die Interaktion von sozialer Herkunft und Verbindlichkeit ($\text{AME}(\text{SES}|\text{V}) - \text{AME}(\text{SES}|\bar{\text{V}}) = 0,001$, n. s.). Das positive Vorzeichen des Effekts entspricht jedoch nicht den theoretischen Erwartungen. Der Einfluss der Grundschulleistungen ist dagegen, wie theoretisch erwartet, nach Modell 5c im Durchschnitt in Ländern mit Verbindlichkeit stärker, die Differenz der AME erreicht jedoch ohne weitere Kontrollen ebenfalls nicht das 5%-Signifikanzniveau ($\text{AME}(\text{ABL}|\text{V}) - \text{AME}(\text{ABL}|\bar{\text{V}}) = 0,027$, $p = 0,08$). Mit der Kontrolle der Empfehlungen in Modell 5d, die selbst *allgemein* einen erheblichen Einfluss haben, ändert sich als *System-Effekt* an dem Herkunftseffekt mit der Verbindlichkeit nichts weiter ($\text{SES}^*\text{V} = 0,011$, n. s.; $\text{AME}(\text{SES}|\text{V}) - \text{AME}(\text{SES}|\bar{\text{V}}) = 0,001$ n. s.), aber der Leistungseinfluss nimmt weiter zu auf nunmehr ($\text{ACE}^*\text{V} = 0,833$, $p < 0,001$; $\text{AME}(\text{ACE}|\text{V}) - \text{AME}(\text{ACE}|\bar{\text{V}}) = 0,051$, $p < 0,001$). In die gleiche Richtung verweisen die Befunde mit der Kontrolle schließlich auch der Aspirationen (Modell 5e): Sie haben allgemein einen erheblichen direkten Einfluss und das Pseudo- R^2 steigt noch einmal deutlich an.⁶ Nun geht der Herkunftseffekt noch einmal zurück (Effekt SES mit 0,016), er ist aber – wie zuvor bei den Bewertungen und erneut *entgegen* der Hypothese des MoAbiT – noch (knapp) signifikant. Ein direkter Effekt der kognitiven Fähigkeiten ist nun (mit 0,606) auch noch vorhanden, aber nicht mehr signifikant. Als *System-Effekt* vermindert sich nun der Einfluss der sozialen Herkunft noch etwas ($\text{SES}^*\text{V} = 0,009$, n. s.; $\text{AME}(\text{SES}|\text{V}) - \text{AME}(\text{SES}|\bar{\text{V}}) = 0,000$, n. s.) und der Unterschied im Leistungseffekt bleibt gleich ($\text{ACE}^*\text{V} = 0,775$, $p < 0,001$; $\text{AME}(\text{ACE}|\text{V}) - \text{AME}(\text{ACE}|\bar{\text{V}}) = 0,051$, $p < 0,001$).⁷

Allgemein sind demnach die Effekte der sozialen Herkunft und der kognitiven Fähigkeiten auf den Übergang weitgehend über die Leistungen in der Grundschule und die daran orientierten Noten und Empfehlungen vermittelt. Die für die soziale Herkunft auch nach allen Kontrollen verbleibenden und nach dem MoAbiT so nicht erwarteten direkten Effekte (Spalte 5e in Tab. 5; mit ^a gekennzeichnet) verweisen

⁶ Die Effekte der Aspirationen sind womöglich überschätzt, weil die Angaben erst nach dem Übergang erhoben wurden; vgl. dazu Abschn. 5.2.

⁷ In einer ähnlich angelegten Analyse, auch mit den Daten der dritten Startkohorte des NEPS für die fünfte Klasse, finden Buchholz et al. (2017) eine Verstärkung der Herkunftseffekte mit der Verbindlichkeit, und zwar nach Kontrolle allein der Noten und des Geschlechts (ebd., Tab. 5.1, Modell 2, S. 84). Die AMEs zeigen allerdings keinen Unterschied, ebenso wenig wie schon die Logit-Koeffizienten in einer Re-Analyse des gleichen Modells.

Tab. 5 Der Übergang auf das Gymnasium (gegenüber allen anderen Schultypen); Logit-Koeffizienten korrigiert nach KHB-Methode; Geschlecht, Migrationshintergrund und Vorschulbesuch kontrolliert; Standardfehler unter Berücksichtigung der Clustering nach Schulklassen

	MoAbiT	NEPS				
		5a	5b	5c	5d	5e
<i>Allgemeine Effekte</i>						
SES	(+)	0,094 ^{***,a}	0,048 ^{***,a}	0,040 ^{***,a}	0,025 ^{**a}	0,016 ^{*a}
ABL	(+)		1,586 ^{***}	1,586 ^{***}	0,909 [*]	0,606
ACE	(+)		2,167 ^{***}	1,666 ^{***}	0,792 ^{***}	0,550 ^{***}
MRK	(+)				, ⁶	, ⁶
REC	+				3,592 ^{***}	2,461 ^{***}
ASP	+					3,419 ^{***}
<i>System-Effekte</i>						
V	≤0			-0,928	-0,857	-0,438
SES*V	≤0			0,012	0,011	0,009
ABL*V	+			, ^b	, ^b	, ^b
ACE*V	+			0,763 ^{***}	0,833 ^{***}	0,775 ^{***}
Psd-R ²		0,10	0,34	0,34	0,50	0,59

Für die Kürzel bei den Effekten vgl. Tab. 1 oben

H Hypothese MoAbiT, *B* Befund NEPS, *SES* soziale Herkunft, *ABL* kognitive Fähigkeiten, *ACE* Leistungen Grundschule, *MRK* Noten, *REC* Empfehlungen, *ASP* Aspirationen, *V* System-Effekt Verbindlichkeit
* $p \leq 0,05$, ** $p \leq 0,01$, *** $p \leq 0,001$

^aAbweichung von Hypothesen des MoAbiT

^bWeil die Empfehlungen in einigen Bundesländern schon institutionell mit den Noten verbunden sind, werden zur Vermeidung von Überkontrollen die Noten bei den Analysen der *System-Effekte* auf die Empfehlungen nicht mit aufgenommen. Analysen mit Berücksichtigung der Noten zeigen, wie zu erwarten, dass ein Teil der Effekte der Leistungen über die Noten vermittelt wird, ergaben darüber hinaus aber keine substantziellen Änderungen. Bei den *System-Effekten* für die Leistungsgerechtigkeit werden entsprechend nur die Leistungen in der Grundschule und nicht gleichzeitig auch die kognitiven Fähigkeiten berücksichtigt.

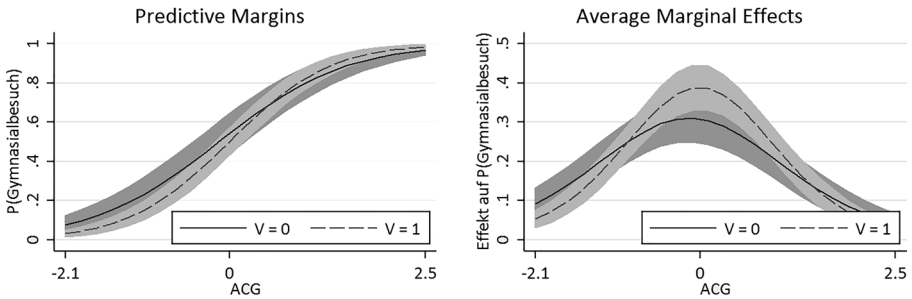
Tab. 6 Average Marginal Effects (AME) für die soziale Herkunft (SES) und die Leistungen in der Grundschule (ACE) auf den Gymnasialübergang in Abhängigkeit der Verbindlichkeit der Empfehlungen (V) für die Modelle 5c bis 5e in Tab. 5

Effekt	MoAbiT	Übergang GYM		
		Ohne weitere Kontrollen 5c	Plus Kontrolle Empfehlung 5d	Plus Kontrolle Aspiration 5e
<i>SES</i>				
V = 0		0,004 ^{***}	0,002 ^{**}	0,002 [*]
V = 1		0,005 ^{***}	0,003 ^{***}	0,002 ^{***}
Differenz	0	0,001	0,001	0,000
<i>ACE</i>				
V = 0		0,217 ^{***}	0,083 ^{***}	0,054 ^{***}
V = 1		0,244 ^{***}	0,134 ^{***}	0,105 ^{***}
Differenz	+	0,027	0,051 ^{***}	0,051 ^{***}

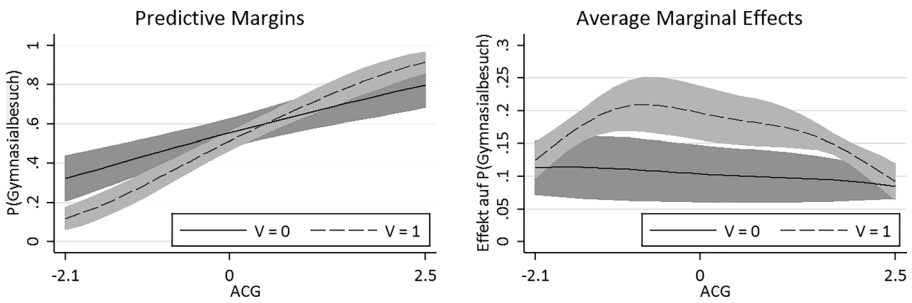
SES soziale Herkunft, *ACE* Leistungen Grundschule, *V* System-Effekt Verbindlichkeit

* $p \leq 0,05$, ** $p \leq 0,01$, *** $p \leq 0,001$

Modell 5c



Modell 5d



Modell 5e

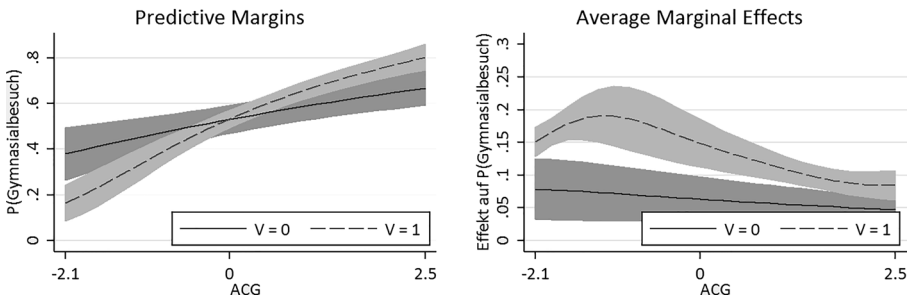


Abb. 2 Grafische Darstellung der Effekte der Verbindlichkeit auf den Gymnasialübergang und die Effekte der Leistungen in der Grundschule (basierend auf den Modellen 5c, 5d und 5e in Tab. 5)

darauf, dass es noch andere Vorgänge gibt, wie frühzeitige Interventionen der Familien und deren Bedienung durch das Lehrpersonal (vgl. dazu auch bereits Fußnote 4 oben). Am Einfluss der sozialen Herkunft ändert sich als *System-Effekt* der Verbindlichkeit nichts, es steigt jedoch der Effekt der Leistungen deutlich und wird mit den Kontrollen für die Empfehlungen und die Aspirationen eher noch stärker.

In Abb. 2 sind die Effekte der Verbindlichkeit auf die Kopplung des Übergangs an die Grundschulleistungen für die drei Modelle mit dem System-Vergleich über die vorhergesagten Werte grafisch für den kompletten Leistungsbereich abgebildet.

Aus der Darstellung wird der nicht signifikante Haupteffekt der Verbindlichkeit in den Grafiken in der linken Spalte gut erkennbar. Das Niveau der beiden Kurven unterscheidet sich im Durchschnitt jeweils kaum, wohl aber die Steigung, die den Interaktionseffekt der Grundschnulleistungen mit der Verbindlichkeit repräsentiert. Diese Steigung ist in der rechten Spalte in Form des AME nochmals abgetragen. In Modell 5c ohne weitere Kontrolle der Empfehlungen oder der Aspirationen gibt es in der Mitte der Verteilung ein deutliches Anwachsen des Verbindlichkeits-Effektes, aber die Differenzen überlappen sich noch stark und an den Rändern der Leistungsverteile besteht kein Unterschied. Mit der Kontrolle der Empfehlungen (Modell 5d) zeigt sich in einem breiten Bereich zwischen den Extremen eine signifikant stärkere Kopplung der Sortierung an die Leistungen, die sich in Modell 5e mit der Kontrolle auch der Aspirationen noch einmal vergrößert, und das besonders bei den leistungsschwächeren Schülern.

Die Verbindlichkeit alleine sorgt demnach noch nicht für eine signifikant stärkere Leistungsgerechtigkeit. Es kommt nach den Modellen offenbar auf die konkrete Implementation der Verbindlichkeit an: Die Empfehlungen müssen *wirklich* den Leistungen entsprechen und es darf *tatsächlich* keine sonderlichen Abweichungen davon geben. Beides ist jedoch hier nicht wirklich erfüllt: Die Noten sind in den verbindlich sortierenden Ländern nicht objektiver und die daran geknüpften Empfehlungen folgen den Leistungen in der Grundschule sogar weniger, wenngleich nicht signifikant (nach Tab. 4, Spalten 3 und 4, mit ^a gekennzeichnet). Und es gibt mit der „Verbindlichkeit“ immer noch nennenswerte Abweichungen von den Empfehlungen, sowohl nach oben noch wie auch nach unten, was nach der aktuellen Regelung ohne weiteres möglich ist (vgl. dazu Abschn. 6.3 unten). Das überlagert und verdünnt das Ausmaß der mit einer wirklich *strikt* umgesetzten Verbindlichkeit erreichbaren Leistungsgerechtigkeit. Mit den entsprechenden Kontrollen wird das in den Grafiken gut sichtbar.⁸

6.3 Wie verbindlich ist die „Verbindlichkeit“?

In Replikationen der oben vorgestellten sechs Studien nach den jeweiligen Ansätzen mit den NEPS-Daten zeigten sich, hier im Einzelnen nicht weiter dokumentiert, große Übereinstimmungen – untereinander und nach den Hypothesen und oben berichteten Befunden nach dem Ansatz des MoAbiT. Das gilt ausnahmslos für die allgemeinen Bedingungen des Übergangs, gleichgültig wie die Modellspezifikationen im Einzelnen aussehen. Für die Replikationen der *System-Effekte* sieht es anders aus. Bei den Befunden von Dollmann (2011), Jähnen und Helbig (2015) und Büchler (2016) gibt es jeweils eine Abweichung vom Originalbeitrag: Bei Dollmann findet sich statt einer Verringerung des SES-Effektes im Original bei der Replikation jetzt

⁸ Die Änderung mit der Berücksichtigung der Aspirationen gegenüber Modell 5d nur mit den Empfehlungen ist nicht sonderlich gravierend, sodass das in Fußnote 8 erwähnte Problem einer möglichen Überschätzung für die Bewertung des Gesamtbefundes kaum ins Gewicht fällt.

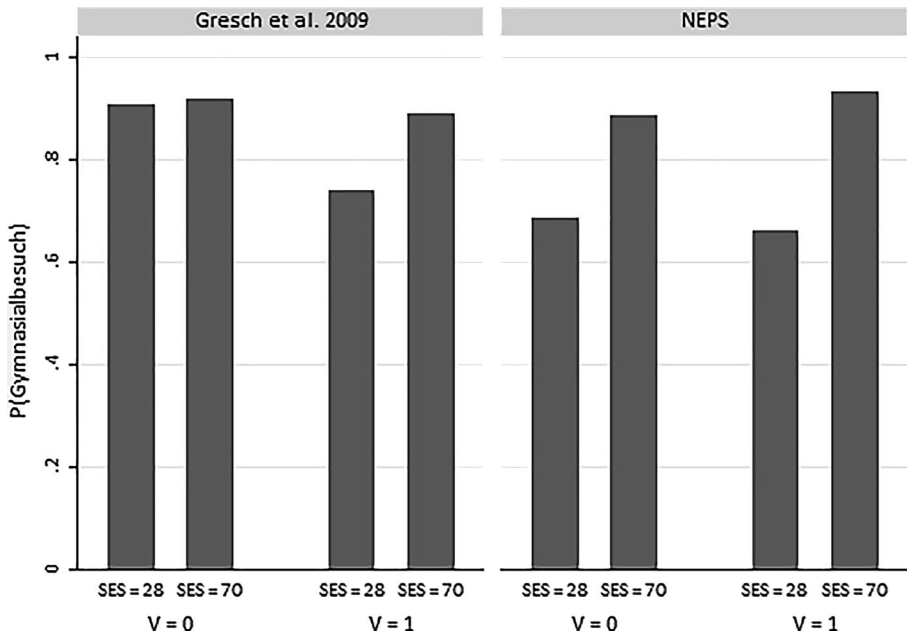


Abb. 3 Übergangswahrscheinlichkeit auf das Gymnasium für Kinder mit einer Empfehlung nach sozialer Herkunft und Verbindlichkeit (*links*: nach Gresch et al. 2009, Abb. 1, S. 251, Werte näherungsweise aus der Grafik dort übernommen; *rechts*: NEPS-Daten der Replikation des Modells von Gresch et al. 2009)

keine Änderung, bei Jähnen und Helbig und bei Büchler statt einer Verstärkung nun ebenfalls keinen Herkunftseffekt.⁹

In den Replikationen lösen sich also alle Abweichungen für den System-Effekt zur sozialen Herkunft aus den Originalbeiträgen auf: Die *soziale Bildungsungleichheit* im Gymnasialübergang ändert sich mit der Verbindlichkeit im Vergleich zur freien Wahl nicht. Das könnte man nunmehr, auch angesichts der Heterogenität der Ansätze, als eine schon wirklich gut abgesicherte Befundlage nehmen. Nicht also, dass sich mit der Verbindlichkeit etwas bei den Herkunftseffekten ändern würde oder dass dieser oder jener der Beiträge in dieser Hinsicht mit dem MoAbiT in Widerspruch stünde, bilden das zentrale Ergebnis der (Re-)Analysen. Es ist vielmehr der Befund, dass sich mit der Verbindlichkeit die *Leistungsgerechtigkeit* der Verteilung auf die Schulformen in der Sekundarstufe offenbar tatsächlich erhöht. Und dass dann darüber gegebenenfalls die mit der Differenzierung angezielte kognitive Homogenisierung der Schulen und Schulklassen einsetzt, von der aus dann sich der Effekt der Komposition der Schulen und Schulen und Schulklassen verstärkt, was das Niveau der Leistungen anheben könnte.

⁹ In einer (hier auch nicht weiter dokumentierten) grafischen Darstellung der Replikation der Befunde bei Büchler (2016, Abb. 3, S. 77; nach den Ergebnissen aus Tab. 2, Modell 3, S. 77 dort) zeigt sich zwar das gleiche Muster, wonach es mit der Verbindlichkeit zu relativen Nachteilen der Kinder aus unteren sozialen Schichten komme. Die Effekte sind in der Replikation nicht signifikant.

Tab. 7 Effekte der strikten Verbindlichkeit auf Leistungsgerechtigkeit und Bildungsgleichheit (logistische Regression, AME; Geschlecht, Migrationshintergrund und Vorschulbesuch kontrolliert)

	Freie Wahl				Verbindlichkeit				Strikt nach Empfehlung			
	1				2				3			
	a	b	c	d	a	b	c	d	a	b	c	d
SES	0,009 ^{***}	0,005 ^{***}	0,004 ^{***}	0,003 ^{***}	0,011 ^{***}	0,005 ^{***}	0,004 ^{***}	0,002 ^{***}	0,009 ^{***}	0,004 ^{***}	0,003 ^{***}	0,002 ^{***}
ABL	0,161 [*]	0,128 [*]	0,128 [*]	0,086 [*]	0,160 ^{***}	0,160 ^{***}	0,093 ^{***}	0,052 ^{***}	0,191 ^{***}	0,191 ^{***}	0,119 [*]	0,091 [*]
ACE	0,224 ^{***}	0,164 ^{***}	0,164 ^{***}	0,119 ^{***}	0,239 ^{***}	0,239 ^{***}	0,134 ^{***}	0,094 ^{***}	0,212 ^{***}	0,212 ^{***}	0,102 ^{***}	0,074 ^{***}
MRK		0,141 ^{***}	0,141 ^{***}	0,083 ^{***}	0,488 ^{***}	0,488 ^{***}	0,223 ^{***}	0,147 ^{***}	0,457 ^{***}	0,457 ^{***}	0,237 ^{***}	0,187 ^{***}
ASP		0,28 ^{***}	0,31 ^{***}	0,39 ^{***}	0,12 ^{***}	0,38 ^{***}	0,47 ^{***}	0,60 ^{***}	0,08 ^{***}	0,29 ^{***}	0,37 ^{***}	0,42 ^{***}
N	1032				2000				2000			

SES soziale Herkunft, ABL kognitive Fähigkeiten, ACE Leistungen Grundschule, MRK Noten, REC Empfehlungen, ASP Aspirationen
^{*} $p \leq 0,05$, ^{**} $p \leq 0,01$, ^{***} $p \leq 0,001$

Bei Gresch et al. (2009, Abb. 1, S. 251) fand sich für diese Zusammenhänge ein besonders interessantes Puzzle: Mit der Verbindlichkeit wird die Übergangsrate für die Kinder *ohne* Empfehlungen generell gesenkt, besonders aber für die aus den oberen Schichten. Für die Kinder *mit* Empfehlungen war es jedoch anders: Die Verbindlichkeit senkt den Übergang auf das Gymnasium für die Kinder *mit* einer Empfehlung aus den unteren Schichten, bei den oberen Schichten hat sie keinen Effekt. Der Effekt entspricht dem Befund bei Büchler (2016, Abb. 3, S. 77), wonach es bei der freien Wahl keinen Abstand zwischen den Schichten gibt, wohl aber bei der Verbindlichkeit (vgl. Abschn. 4 oben). Es gäbe also zwei gegenläufige Effekte der Verbindlichkeit auf den Zusammenhang von sozialer Herkunft und Übergang. Aber warum?

Für die Lösung des Puzzles gäbe es eine einfache Erklärung: „Verbindlichkeit“ sollte bedeuten, dass der Übergang wirklich *strikt* den Empfehlungen folgt und es weder Abweichungen nach oben noch nach unten gibt. Das ist aber so nicht geregelt: Abweichungen nach oben werden institutionell zwar (weitgehend) ausgeschlossen, aber nach unten sind sie immer möglich (Füssel et al. 2010, S. 93 f.) Dieser Option folgen offenbar besonders die Eltern der Kinder mit den Empfehlungen aus den unteren Schichten (vgl. etwa Lohmann und Groh-Samberg 2010, S. 479 ff.). Das zeigt sich in den NEPS-Daten so nicht: Die Kinder aus den unteren Schichten *mit* einer Empfehlung werden *unabhängig* von der Verbindlichkeit von ihren Eltern weniger auf das Gymnasium gelassen als die aus den oberen. Abbildung 3 zeigt das anhand der vorhergesagten Wahrscheinlichkeit eines Gymnasialbesuchs in Abhängigkeit von sozialer Herkunft und Verbindlichkeit, links nach Gresch et al. 2009, rechts nach entsprechenden Analysen mit den NEPS-Daten.

Somit stellt sich die Frage: Was wäre mit dem Effekt der sozialen Herkunft (und der Leistungsgerechtigkeit), wenn die Abweichungen nach unten auch unterbunden würden. Mit den vorliegenden Daten lässt sie sich nicht wirklich beantworten. Es würde ein (Feld-)Experiment voraussetzen, das kaum zu implementieren wäre. Der Fall lässt sich nur näherungsweise simulieren. Tabelle 7 zeigt, was geschähe, wenn in den Bundesländern mit der „Verbindlichkeit“ alle Abweichungen ausgeschlossen sind. Dazu werden drei getrennte Analysen vorgenommen: Die freie Wahl (Spalte 1), die aktuelle Regelung der Verbindlichkeit mit der Möglichkeit für eine Abweichung nach unten für den u.U. von den Empfehlungen abweichenden Übergang auf das Gymnasium (Spalte 2) und die (fiktive) Regelung, dass die Empfehlungen den Übergang direkt und ohne weitere Entscheidungen der Eltern bestimmen (Spalte 3). Die Vergleiche beziehen sich auf jeweils auf den (Brutto-)Effekt der sozialen Herkunft alleine (Spalten a), unter Kontrolle der primären Effekte über die kognitiven Fähigkeiten und die Grundschulleistungen (Spalten b), der tertiären Effekte über die Noten (Spalten c) und schließlich auch der sekundären Effekte über die Aspirationen (Spalten d).¹⁰

¹⁰ Im getrennten Vergleich der Effekte besteht das Problem der Überkontrolle bei den Interaktionseffekten, anders als über die Modellierung über Interaktionseffekte wie in Tab. 5, nicht, sodass hier beide Aspekte der Leistungsgerechtigkeit, die Änderung des Effekts der Leistungen in der Grundschule (ACE) und die der kognitiven Fähigkeiten (ABL), gleichzeitig betrachtet werden können. Zur Kontrolle der tertiären Effekte und zum Vergleich mit dem Eingang 3 (mit den Empfehlungen als „Übergang“) werden die Noten genommen.

Es gibt nur geringe Unterschiede in den Herkunftseffekten zwischen den drei Systemen, aber eine deutliche Tendenz wird sichtbar: Der Effekt der sozialen Herkunft nimmt in allen Modellen (a–d) mit der Verbindlichkeit nach der aktuellen Regelung (Eingang 2) im Vergleich zur freien Wahl (Eingang 1) etwas zu, sinkt dann aber bei einer strikten Regelung (Eingang 3), und zwar stets unter die Größe bei der freien Wahl.

Gewiss nicht überraschend, aber in den Einzelheiten bemerkenswert, sind die Änderungen auch der anderen Effekte: Der Einfluss der *Schule* nimmt über die Noten mit der Stringenz der Verbindlichkeit zu (Modelle c und d für die Eingänge 1 bis 3) und der Einfluss der *Familie* über die Aspirationen ganz erheblich ab (Modell d für die Eingänge 1 bis 3). Das muss nicht heißen, dass damit die Leistungsgerechtigkeit in gleichem Maße ansteigen würde, aber zumindest zu einem erheblichen Anteil: Die Noten sind (wie die Empfehlungen) nach den Befunden in Abschn. 6.1 (Tab. 4, Modelle 3 und 4) zum überwiegenden Teil über die kognitiven Fähigkeiten und Grundschulleistungen bestimmt und bilden damit immer auch eigene Anteile der intellektuellen Fähigkeiten ab (vgl. die Befunde zu den Noten in Abschn. 6.1).

Ähnlich ist es für die beiden unmittelbar *leistungsbezogenen* Aspekte – die kognitiven Fähigkeiten und die Grundschulleistungen. Zusammen gesehen kommt darüber eine nochmalige Zunahme an Leistungsgerechtigkeit (mit gewissen Verschiebungen im Anteil jeweils) zustande. Bemerkenswert ist, dass bei der strikten Verbindlichkeit die kognitiven Fähigkeiten sogar direkt wieder mehr an Bedeutung gewinnen, wenngleich nicht signifikant. Es ist das, was die Leistungsdifferenzierung eigentlich soll: Die Sortierung nach den oft im Hintergrund bleibenden und auch manchmal in der Grundschule noch verborgenen generellen intellektuellen Fähigkeiten – und nicht so sehr die Leistungen nach den oft mühsamen Anstrengungen zum Schluss der Grundschule und den dabei evtl. nur mit einigem familiären Nachhelfen erzielten, aber unverdienten Noten und Empfehlungen.

Wichtiger fast noch aber als diese Aufklärung und Zerlegung der einzelnen kausalen Prozesse sind für die *praktische* Beurteilung die für den Schulalltag relevanten *Brutto-Zusammenhänge*: Die Einflüsse der sozialen Herkunft und die der Fähigkeiten und Leistungen ohne die Kontrolle der Noten und Aspirationen. So kommen die Kinder *faktisch* auf die Schulen! Der Vergleich betrifft jeweils die Modelle b in den drei Eingängen. Und hier zeigt sich das generelle Muster besonders deutlich. Im Vergleich ist der Einfluss *der sozialen Herkunft* bei der *strikten* Verbindlichkeit am *geringsten* und der Einfluss gerade der *kognitiven* Fähigkeiten am *stärksten*. Genau *das* ist es aber, worum es bei der Differenzierung letztlich geht!

Die Effekte der strikten Bindung des Übergangs an die Empfehlungen sind für die Minderung der Bildungsungleichheit deutlich und konsistent, für die Leistungsgerechtigkeit dagegen weniger als man hätte erwarten können. Dabei ist jedoch zu beachten, dass nach den Befunden in Tab. 4 in Abschn. 6.1 in den verbindlich sortierenden Ländern die Noten ähnlich und die Empfehlungen weniger an die Leistungen gekoppelt sind – warum auch immer. Wenn das geändert würde, wäre der Effekt der Leistungsgerechtigkeit also noch ausgeprägter. Es verweist darauf, dass es mit einer einzelnen Regel nicht getan ist, sondern andere hinzukommen müssen, wie die Kontrolle und Standardisierung der Bewertungen, auch Anreize für die Vermei-

derung von Fehlplatzierungen, und auch darauf zu achten ist, dass das organisatorisch umgesetzt wird.

7 Zusammenfassung und Bewertung

Auf die Frage, ob die Differenzierung der Bildungswege nach der Grundschule wirklich (nur) zur Verschärfung der sozialen Bildungsungleichheiten führt und die Vorgabe der Leistungsgerechtigkeit das, wie es oft auch heißt, bloß ideologisch bemäntelt, geben die Ergebnisse für die Effekte der Verbindlichkeit der Empfehlungen und die damit verbundene Beschränkung des „Elternwillens“ eine einfache und eindeutige Antwort: Es gibt bei Verbindlichkeit der Empfehlungen die Verstärkung der sozialen Bildungsungleichheit bei der Bildungsbeteiligung gegenüber der freien Elternwahl *nicht*, wohl aber *steigt* – *ceteris paribus* bei den Bewertungen – die Leistungsgerechtigkeit bei der Aufteilung.

Davon sind die beiden anderen Arten von Herkunftseffekten unberührt: Die primären Effekte der nach sozialer Schicht und Bildungsnähe der Eltern unterschiedlichen kognitiven Entwicklung und die tertiären Effekte der nach der sozialen Herkunft verzerrten Bewertungen in den Noten und Empfehlungen. Es gibt das überall und über die Bildungssysteme in ähnlicher Weise, auch bei (voller) Integration, spätestens wenn auch die späteren Verläufe und die bei Integration immer auch vorhandenen, die internen und eher nur verdeckten Differenzierungen in die Betrachtung einbezogen werden (vgl. Lucas 1999, Kap. 8; Jackson und Jonsson 2013, S. 318 ff.; Blossfeld et al. 2016, S. 396 f.). Um diese Einflüsse zu begrenzen, wären andere institutionelle Maßnahmen möglich: Eine frühe und auch schon leistungsbezogene vorschulische Bildung bei den primären Effekten, stärkere Anreize, wirksamere Kontrollen und bessere Informationen für das Lehrpersonal zur Förderung der Genauigkeit und Objektivität bei den Bewertungen und den tertiären Effekten.

An beidem wäre gewiss auch etwas zu tun, aber die primären und die tertiären Effekte sind, nach allem, was sich zeigt, nicht der Kern des Problems der sozialen Bildungsungleichheiten. Es sind vielmehr die *sekundären* Effekte, die die Vorgänge im Wesentlichen tragen und bei denen die besser gestellten Familien ihre Privilegien an einer entscheidenden Stelle der Biografie ihrer Kinder ausspielen können und die weniger gut positionierten aus nachvollziehbaren Gründen eher vorsichtiger sind. Sie bilden die „dark side of choice“ (Jackson und Jonsson 2013, S. 329). Und nur, wenn die von den Empfehlungen abweichenden Entscheidungen der Eltern wirklich begrenzt werden, wenn also den Eltern mit oder ohne Empfehlung *keine* Wahl bleibt und auch – vorher wie nachher – kein „room for maneuver“ (Jackson und Jonsson 2013, S. 329) etwa „to lobby teachers for more advantageous tracks placements“ (Jackson und Jonsson 2013, S. 329) würden die Ungleichheiten nach der sozialen Herkunft in den nach Leistung differenzierenden Systemen nicht nur nicht größer, sondern eher sogar geringer. Was denn anders? Und weil es dann auch weniger an Fehlplatzierungen nach den kognitiven Fähigkeiten gibt, nehmen, so darf man ergänzen, auch die Leistungsgerechtigkeit und die Effizienz des Kompetenzerwerbs noch einmal zu. Was denn sonst?

Für das Ziel einer Erhöhung der Leistungsgerechtigkeit, der Begrenzung von Bildungsungleichheit und -ungerechtigkeit *und* der Verbesserung der Effizienz des schulischen Kompetenzerwerbs wären also nicht die weitere Öffnung und Liberalisierung, sondern die Stärkung der Selektivität nach den Fähigkeiten und Leistungen der Kinder angeraten, einschließlich der Unterbindung auch von Abweichungen nach unten. Das wird gewiss nicht überall gerne gehört. Aber das Ziel wäre schon ein ganz besonderes: „No *talent* left behind!“ – egal woher sie herkommen, die Talente. Mehr noch: Es sind ja gerade die talentierten Kinder aus den unteren Schichten, denen – auch aus nicht unbegründeter Vorsicht der Eltern – die von der Schule eröffnete Chance eher verwehrt wird, wenn man ihnen die Wahl überlässt.

Ob sich diese weitere Einschränkung des „Elternwillens“ durchsetzen ließe, ist wohl mehr als zweifelhaft. Es wäre gewiss auch von der für das *Ability Tracking* letztlich entscheidenden Frage abhängig: Zeigen sich die förderlichen Effekte der (strikten) Differenzierung wirklich in den Leistungen in der Sekundarstufe danach? Auch bei den leistungsschwächeren Kindern? Bei denen aus den unteren Schichten? Und bei den Kindern mit Migrationshintergrund? Dem Modell der Leistungsdifferenzierung nach sollte es so sein. Bisher weiß man auch darüber nur wenig, weil es die für die Analysen der Leistungen in der Sekundarstufe in den internationalen Vergleichsstudien nötigen Informationen über die Sortierung vorher nicht gibt – wie bei PISA insbesondere. Daher das NEPS und der Vergleich der deutschen Bundesländer mit ihren unterschiedlichen Regelungen. Einige aktuellere Befunde weisen jedenfalls darauf hin, dass es wohl so ist: In einem Vergleich von Hessen und Bayern über die Daten der BIKS-Studie zeigten sich keine Unterschiede in den Herkunftseffekten auf die Leistungen, wohl aber eine Verstärkung der förderlichen Effekte der Leistungskomposition in den Schulklassen für Bayern mit seinem strikteren Regime (Esser und Relikowski 2015, S. 20 ff.). Die Berliner Strukturreform der Zusammenlegung der nicht-gymnasialen Schulformen zu stärkerer Heterogenität, nach den Leistungen wie nach der sozialen Herkunft, hatte einen Rückgang im Leistungsniveau zur Folge, und das besonders bei den Gymnasien (Becker et al. 2017, S. 166 ff.). Und dass in Baden-Württemberg nach dem 2. IQB-Bericht für 2015 die Leistungen nach der Abschaffung der Verbindlichkeit nach der Landtagswahl 2011 als einzigem Land im Vergleich zum 1. IQB-Bericht für 2009 deutlich gesunken sind, passt auch in dieses Bild: „Durchgehend *ungünstige* Veränderungen zeigen sich nur in *einem* Land (Baden-Württemberg)“ (Böhme und Hoffmann 2016, S. 257; Hervorhebungen nicht im Original). Wie es nach den NEPS-Daten mit den System-Effekten bei den Leistungen in den 16 Bundesländern insgesamt dann aussieht, wird abzuwarten sein.

Danksagung Die Verfasser danken den Teilnehmerinnen und Teilnehmern einer Reihe von Konferenzen und sonstigen Zusammenkünften, den Herausgebern der Kölner Zeitschrift für Soziologie und Sozialpsychologie sowie einem anonymen Gutachter für ihre Anregungen und Kritik und den Autorinnen und Autoren der sechs im Beitrag speziell besprochenen Studien für ihre Kooperation in der Zusammenfassung und Bewertung ihrer Befunde, besonders aber Walter Müller für seine besonders engagierte Diskussion kritischer Fragen. Gedankt sei auch der Leitung und dem Team der National Educational Panel Study (NEPS) in Bamberg für ihre Unterstützung. Hans-Peter Blossfeld und Jan Skopek haben die ersten theoretischen Überlegungen im Rahmen des Projekts EduLife am European University Institute in Florenz konzeptionell begleitet, ebenso wie Gary N. Marks von der Australian Catholic University, Melbourne, und Jaap Dron-

kers von der Universität Maastricht. Jaap Dronkers fehlt dem noch nicht abgeschlossenen Projekt nach seinem überraschenden Tod sehr.

Literatur

- Allmendinger, Jutta. 2012. *Schulaufgaben. Wie wir das Bildungssystem verändern müssen, um unseren Kindern gerecht zu werden*. München: Pantheon.
- Becker, Rolf, und Heike Solga. 2012. Soziologische Bildungsforschung – eine kritische Bestandsaufnahme. In *Soziologische Bildungsforschung*. Sonderheft 52 der Kölner Zeitschrift für Soziologie und Sozialpsychologie, Hrsg. Rolf Becker und Heike Solga, 7–43. Wiesbaden: Springer.
- Becker, Michael, Marko Neumann, Susanne Radmann, Malte Jansen, Gabriel Nagy, Christoph Borzikowsky, Michael Leucht, Kai Maaz, Olaf Köller und Jürgen Baumert. 2017. Schulleistungen vor und nach der Berliner Schulstrukturreform. In *Zweigliedrigkeit im deutschen Schulsystem. Potenziale und Herausforderungen in Berlin*, Hrsg. Marko Neumann, Michael Becker, Jürgen Baumert, Kai Maaz und Olaf Köller, 155–188. Münster: Waxmann.
- Betts, Julian R. 2011. The economics of tracking in education. In *Handbook of the economics of education*, Band 3, Hrsg. Eric A. Hanushek, Stephen Machin und Ludger Woessmann, 341–381. Amsterdam: North Holland.
- Blossfeld, Hans-Peter, Hans-Günter Roßbach und Jutta von Maurice. 2011. *Education as a lifelong process—The German National Educational Panel Study (NEPS)*. Wiesbaden: Springer VS.
- Blossfeld, Hans-Peter, Moris Triventi, Jan Skopek, Nevena Kulic und Sandra Buchholz. 2016. Varieties of secondary education models and social inequality—Conclusions from a large scale international comparison. In *Models of secondary education and social inequality*, Hrsg. Hans-Peter Blossfeld, Sandra Buchholz, Jan Skopek und Moris Triventi, 377–400. Cheltenham, UK: Edward Elger Publishing.
- Böhme, Katrin, und Lars Hoffmann. 2016. Mittelwerte und Streuungen der im Fach Deutsch erreichten Kompetenzen. In *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich*, Hrsg. Petra Stanat, Katrin Böhme, Stefan Schipolowski und Nicole Haag, 335–358. Münster: Waxmann.
- Buchholz, Sandra, Jan Skopek, Markus Zielonka, Hartmut Ditton, Florian Wohlkinger und Antonia Schier. 2017. Secondary school differentiation and inequality of educational opportunity in Germany, In *Models of secondary education and social inequality*, Hrsg. Hans-Peter Blossfeld, Sandra Buchholz, Jan Skopek und Moris Triventi, 79–92. Cheltenham, UK: Edward Elger Publishing.
- Büchler, Theresa. 2016. Schulstruktur und Bildungsungleichheit: Die Bedeutung von bundeslandspezifischen Unterschieden beim Übergang in die Sekundarstufe I für den Bildungserfolg. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 68:53–87.
- Dollmann, Jörg. 2011. Verbindliche und unverbindliche Grundschulempfehlungen und soziale Ungleichheiten am ersten Bildungsübergang. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 63:595–621.
- Erikson, Robert, John H. Goldthorpe, Michelle Jackson, Meir Yaish und David R. Cox. 2005. On class differentials in educational attainment. *Proceedings of the National Academy of Sciences of the United States of America* 102:9730–9733.
- Esser, Hartmut 2016a. Bildungssysteme und ethnische Bildungsungleichheit. In *Ethnische Ungleichheiten im Bildungsverlauf. Mechanismen, Befunde, Debatten*, Hrsg. Claudia Diehl, Christian Hunkler und Cornelia Kristen, 331–396. Wiesbaden: Springer VS.
- Esser, Hartmut. 2016b. Educational systems and educational inequality. The model ability tracking and empirical findings. In *Models of secondary education and social inequality*, Hrsg. Hans-Peter Blossfeld, Sandra Buchholz, Jan Skopek und Moris Triventi, 25–44. Cheltenham, UK: Edward Elger Publishing.
- Esser, Hartmut, und Ilona Relikowski. 2015. *Is ability tracking (really) responsible for educational inequalities in achievement? A comparison between the country states Bavaria and Hessen in Germany*. IZA-Discussion Paper Nr. 9082. Bonn: Forschungsinstitut Zukunft der Arbeit.
- Füssell, Hans-Peter, Cornelia Gresch, Jürgen Baumert und Kai Maaz. 2010. Der institutionelle Kontext von Übergangsentscheidungen: Rechtliche Regelungen und die Schulformwahl am Ende der Grundschulzeit. In *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten*, Hrsg. Kai Maaz, Jürgen Baumert, Cornelia Gresch und Nele McElvany, 87–106. Bonn: BMBF.
- Ganzeboom, Harry B., Paul M. de Graaf und Donald J. Treiman. 1992. A standard international socioeconomic index of occupational status. *Social Science Research* 21:1–56.

- Gresch, Cornelia, Jürgen Baumert und Kai Maaz 2009. Empfehlungsstatus, Übergangsempfehlung und der Wechsel in die Sekundarstufe I: Bildungsentscheidungen und soziale Ungleichheit. *Zeitschrift für Erziehungswissenschaften* 12:230–256.
- Hadjar, Andreas, und Christiane Gross. Hrsg. *Education systems and inequalities. International comparisons*. Bristol: Policy Press.
- Holtappels, Heinz Günter. 2017. *Entwicklung und Qualität des Schulsystems. Neue empirische Befunde und Entwicklungstendenzen*. Münster: Waxmann.
- Jackson, Michelle, und Jan O. Jonsson. 2013. Why does inequality of educational opportunity vary across countries? Primary and secondary effects in comparative context. In *Determined to succeed? Performance versus choice in educational attainment*, Hrsg. Michelle Jackson, 306–337. Stanford, CA: Stanford University Press
- Jähnen, Stefanie, und Marcel Helbig. 2015. Der Einfluss schulrechtlicher Reformen auf Bildungsungleichheiten zwischen den deutschen Bundesländern. Eine quasiexperimentelle Untersuchung am Beispiel der Verbindlichkeit von Übergangsempfehlungen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 67:539–571.
- Karlson, Kristian B., Anders Holm und Richard Breen. 2012. Comparing regression coefficients between same-sample nested models using logit and probit: A new method. *Sociological Methodology* 42:286–313.
- Kohler, Ulrich, Kristian B. Karlson und Anders Holm. 2011. Comparing coefficients of nested nonlinear probability models. *Stata Journal* 11:420–438.
- Lang, Frieder R., Stefan Kamin, Margund Rohr, Conrad Stünkel und Bettina Williger. 2014. *Erfassung der fluiden kognitiven Leistungsfähigkeit über die Lebensspanne im Rahmen des Nationalen Bildungspanels: Abschlussbericht zu einer NEPS-Ergänzungsstudie* NEPS Working Paper No. 43. Bamberg: Leibniz-Institut für Bildungsverläufe.
- Lohmann, Henning, und Olaf Groh-Samberg. 2010. Akzeptanz von Grundschulempfehlungen und Auswirkungen auf den weiteren Bildungsverlauf. *Zeitschrift für Soziologie* 39:470–492.
- Lucas, Samuel R. 1999. *Tracking inequality: Stratification and mobility in American high schools*. New York: Teachers College Press.
- Maaz, Kai, und Gabriel Nagy. 2009. Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarsystems: Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte. *Zeitschrift für Erziehungswissenschaften* 12:153–182.
- Marks, Gary N. 2014. *Education, social background and cognitive ability. The decline of the social*. London: Routledge.
- Morgan, Stephen L., Michael W. Spiller und Jennifer J. Todd. 2013. Class origins and education in the United States. In *Determined to succeed? Performance versus choice in educational attainment*, Hrsg. Michelle Jackson, 279–305. Stanford, CA: Stanford University Press.
- Neugebauer, Martin. 2010. Bildungsungleichheit und Grundschulempfehlung beim Übergang auf das Gymnasium: Eine Dekomposition primärer und sekundärer Herkunftseffekte. *Zeitschrift für Soziologie* 39:202–214.
- Oakes, Jeannie. 2005. *How schools structure inequality*, 2. Aufl. New Haven: Yale University Press.
- Pfeffer, Fabian, 2008. Persistent inequality in educational attainment and its institutional context. *European Sociological Review* 24:543–565
- Roth, Tobias, und Manuel Siegert 2015. Freiheit versus Gleichheit? Der Einfluss der Verbindlichkeit der Übergangsempfehlung auf die soziale Ungleichheit in der Sekundarstufe. *Zeitschrift für Soziologie* 44:118–136.
- Roth, Tobias, und Manuel Siegert. 2016. Does the selectivity of an educational system affect social inequality in educational attainment? Empirical findings for the transition from primary to secondary level in Germany. *European Sociological Review* 32:779–791.
- Sörensen, Aage B. 1970. Organizational differentiation of students and educational opportunity? *Sociology of Education* 43:355–376.
- Sörensen, Aage B., und Maureen T. Hallinan. 1977. A reconceptualization of school effects. *Sociology of Education* 50:273–289.
- Van de Werfhorst, Herman G., und Jonathan B. Mijs. 2010. Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology* 36:407–428.
- Wößmann, Ludger, Elke Lüdemann, Gabriela. Schütz und Martin R. West. 2009. *School accountability, autonomy and choice around the world*. Cелtenham: Edward Elgar Publishing.

Hartmut Esser Prof. (em.) Dr., Fakultät für Sozialwissenschaften Universität Mannheim. Forschungsgebiete: Allgemeine und spezielle Soziologie; Sozialwissenschaftliche Methodologie, Migration, Soziologische Handlungstheorie, Integration und ethnische Konflikte, Bildungssoziologie. Letzte Veröffentlichung: Sanktionen, Reziprozität und die symbolische Konstruktion einer Kooperations-„Gemeinschaft“. Ein theoretischer Vergleich und empirischer Test von Rational-Choice-Theorie und dem Modell der Frame-Selektion anhand von Befunden und Daten aus der experimentellen Spieltheorie zur Erklärung der Bereitstellung von Kollektivgütern. *Zeitschrift für Soziologie* 47, 2018.

Kerstin Hoenig Dr., Leibniz-Institut für Bildungsverläufe (LifBi), Arbeitsbereichsleitung des Arbeitsbereichs „Soziale Ungleichheit und Bildungsentscheidungen im Lebenslauf“. Forschungsgebiete: Bildungsungleichheit, soziales Kapital, Bildungsaspirationen, Bezugsgruppeneffekte. Letzte Veröffentlichung: Social capital, participation in adult education, and labor market success: Constructing a new instrument (mit R. Pollak, B. Schulz und V. Stocké). In: *Methodological issues of longitudinal surveys. The example of the national educational panel study* (Hrsg. H.P. Blossfeld, J. von Maurice, M. Bayer und J. Skopek). Frankfurt a. M. 2016.

1 **ERRATUM**

2
3
4 **Erratum zu: Leistungsgerechtigkeit und**
5 **Bildungsungleichheit**

6 **Effekte der Verbindlichkeit der Grundschulempfehlungen beim**
7 **Übergang auf das Gymnasium. Ein Vergleich der deutschen**
8 **Bundesländer mit den Daten der „National Educational Panel**
9 **Study“ (NEPS)**

10
11
12 **Hartmut Esser · Kerstin Hoenig**

13
14
15 © Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2018

16
17
18
19 **Erratum zu:**
20 **Köln Z Soziol 2018**
21 <https://doi.org/10.1007/s11577-018-0558-2>

22
23 Leider wurde beim Druck versehentlich die falsche Abb. 1 verwendet.

24
25
26
27
28
29
30
31
32
33
34
35
36
37
38 Die Online-Version des Originalartikels ist unter <https://doi.org/10.1007/s11577-018-0558-2> zu
39 finden.

40 H. Esser (✉)
41 MZES, Universität Mannheim
42 68131 Mannheim, Deutschland
43 E-Mail: hartmut-esser@t-online.de

44 K. Hoenig
45 Leibniz-Institut für Bildungsverläufe (LifBi)
46 Wilhelmsplatz 3, 96047 Bamberg, Deutschland
E-Mail: kerstin.hoenig@lifbi.de

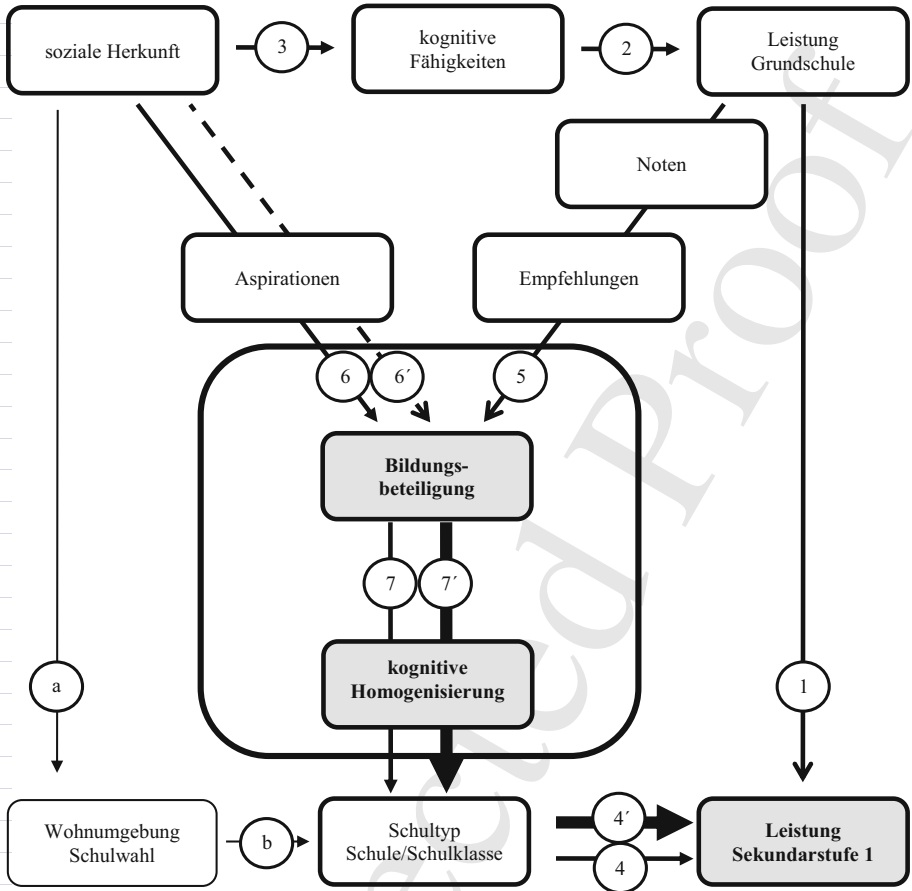


Abb. 1 Das Modell der Leistungsdifferenzierung (MoAbiT) und die Effekte der Verbindlichkeit

Hartmut Esser

Kognitive Homogenisierung, Leistungsniveau und Bildungsungleichheit

Theoretische Modellierung und empirische Analyse der Effekte der Leistungsdifferenzierung und des Einflusses der sozialen Herkunft auf die Leistungen in der Sekundarstufe in den deutschen Bundesländern mit den Daten der „National Educational Panel Study“ (NEPS)

(Stand 25112018)

(80500 (- 1500 Zusammenfassung))

(do not cite)

Zusammenfassung

Die Differenzierung der Bildungswege nach den kognitiven Fähigkeiten wird damit begründet, dass die Aufteilung nach den Leistungen eine für das Lernen aller Kinder förderliche Homogenisierung der schulischen Lernumgebung ermögliche. Dagegen wird angeführt, dass die Differenzierung keine besonderen Effekte auf die Leistungen in der Sekundarstufe habe, sondern eher nur die Effekte der sozialen Herkunft verstärke. Diese Auffassung gilt seit Langem als Standardposition. Der Beitrag untersucht auf der Grundlage eines allgemeinen theoretischen Modells die Effekte einer unterschiedlich strikt nach den kognitiven Fähigkeiten vorgenommenen Differenzierung auf die Leistungen in der Sekundarstufe. Grundlage sind Daten der „National Educational Panel Study“ (NEPS) für die 16 deutschen Bundesländer, weil (bisher nur) dafür die theoretisch erforderlichen Konstrukte zur Verfügung stehen. Sechs Befunde lassen sich benennen. Mit der strikten Differenzierung steigt die Leistungsgerechtigkeit beim Übergang, ohne dass sich der Einfluss der sozialen Herkunft ändert. Sie erhöht die kognitive Homogenität der Schulklassen und steigert das Leistungsniveau in der Sekundarstufe über die Verstärkung des Effekts des kognitiven Niveaus in den Schulklassen. Am stärksten ist die Leistungssteigerung für kognitiv homogene Schulklassen bei strikter Differenzierung. Davon profitieren vor allem die leistungsschwächeren Kinder. Die Befunde entsprechen nahezu ausnahmslos den Hypothesen des theoretischen Modells und widersprechen der Standardposition. Das lässt sich damit erklären, dass die theoretisch erforderlichen Daten bisher nicht zur Verfügung standen, insbesondere nicht in den meist verwendeten internationalen Vergleichsstudien.

Die Differenzierung nach den kognitiven Fähigkeiten, das sog. „*Ability Tracking*“, bezeichnet die institutionell gezielte Sortierung nach Begabungen und Talenten in unterschiedliche Schultypen und Bildungswege in der Sekundarstufe. Die wichtigste Begründung: Damit ließen sich den verschiedenen Fähigkeiten und Interessen der Kinder besser angepasste Curricula einrichten und der jeweilige Unterricht darauf dann auch stärker fokussieren, so dass die Leistungen für *alle* besser werden könnten, ohne dass sich andere Einflüsse wie die der sozialen Herkunft verstärkten, weil die Sortierung nach den *Fähigkeiten* das eher begrenzen sollte (Sörensen 1970: 356ff.; Gamoran und Mare 1989: 1148ff.; Betts 2011: 343). Dagegen wird angeführt, dass es die erhofften Leistungssteigerungen nicht gebe, und eher (nur) die soziale Ungleichheit im Bildungserfolg verstärkt oder gar erst geschaffen werde (Oakes 2005 (1985), S. 40ff.; Gamoran 2009: 4ff.; Betts 2011, S. 343f.). Tatsächlich weisen die empirischen Befunde der meisten Beiträge dazu in diese Richtung (vgl. für die Bildungsbeteiligung u.a. Pfeffer 2008, S. 555ff.; Becker und Solga 2012, S. 19ff.; für die Leistungen u.a.: Wößmann et. al. 2009, Tabelle 7 A.1; van der Werfhorst und Mijs 2010, S. 415ff.). Man könnte es als die – weithin als selbstverständlich geltende – Standardposition bezeichnen (vgl. etwa Allmendinger 2012, Kapitel 10, oder neuerdings wieder Holtappels 2017, S. 16ff.). Allerdings gab es schon von Beginn an eine Reihe von Hinweisen auf Ausnahmen, Inkonsistenzen und offene Fragen, die Zweifel hätten wecken können (vgl. Jackson und Jonsson 2013, S. 323ff.; Marks 2014, Kapitel 5; auch: verschiedene Beiträge und Abschnitte bei Blossfeld et al. 2016 oder Hadjar und Gross 2016).

Geklärt ist die Sache bisher nicht. Die Diskussion bezieht sich sowohl auf die Bildungsbeteiligung wie auf die Leistungen. Zur *Bildungsbeteiligung* zeigt sich inzwischen aus Analysen mit geeigneten Daten, dass eine striktere Leistungsdifferenzierung, etwa über die Verbindlichkeit der Empfehlungen der Schule, nicht nur die Einflüsse der sozialen Herkunft eher dämpft, sondern die Leistungsgerechtigkeit bei der Aufteilung in die verschiedenen Bildungswege stärkt, was ja gerade die Absicht bei der Differenzierung ist (Esser und Hoenig 2018). In dem Beitrag hier geht es im Anschluss daran um die schulischen *Leistungen* in der Sekundarstufe nach dem Übergang. Die Grundhypothese lautet: Eine *strikte* Differenzierung nach den Fähigkeiten und Leistungen vorher erhöht nicht nur die Leistungsgerechtigkeit bei der Sortierung, sondern steigert *dadurch* die kognitive Homogenität der Schulklassen, wodurch sich die leistungssteigenden Effekte des kognitiven Niveaus über die bessere Anpassung von Curriculum und Unterricht noch einmal *erhöhen* – *ohne* dass sich der Einfluss der sozialen Herkunft verstärkt und auch ohne dass es Nachteile

bei den leistungsschwächeren Kindern und denen aus den unteren bzw. bildungsferneren Schichten gibt.

Weil die bisherigen empirischen Befunde auch dazu auf Studien mit sehr verschiedenen Annahmen, Designs und Daten beruhen und (daher) oft kaum vergleichbar sind, wird in Abschnitt 1 zunächst ein übergreifendes theoretisches Modell vorgestellt, das es erlaubt, die Effekte der Differenzierung zu systematisieren. Abschnitt 2 fasst die aus dem Modell ableitbaren Hypothesen zusammen. Daran anschließend werden in Abschnitt 3 die wichtigsten bisherigen Befunde für und gegen die Standardposition berichtet und auf die Hypothesen bzw. das theoretische Modell bezogen. Abschnitt 4 beschreibt die zum empirischen Test des Modells verwendeten Daten und die empirisch-statistische Umsetzung. Es handelt sich um Daten der „National Educational Panel Study“ (NEPS) für die 16 deutschen Bundesländer, die sich in ihren Regelungen der Differenzierung ganz erheblich unterscheiden, speziell in Hinsicht auf die Stringenz der Sortierung nach den Fähigkeiten und den Leistungen vorher. Der Vergleich der deutschen Bundesländer und die Daten des NEPS wurden benutzt, weil so gut wie alle anderen Quellen, darunter die internationalen Vergleichsstudien, die zur Erfassung der Grundvorgänge des theoretischen Modells nötigen Informationen nicht enthalten, darunter insbesondere die zu den kognitiven Fähigkeiten der Kinder und der entsprechenden Angaben über die Zusammensetzung der Schulen und Schulklassen nach Niveau und Homogenität. Abschnitt 5 berichtet über die Befunde, und in Abschnitt 6 gibt es eine kurze, auch bewertende Zusammenfassung.

1. Das Modell der Leistungsdifferenzierung

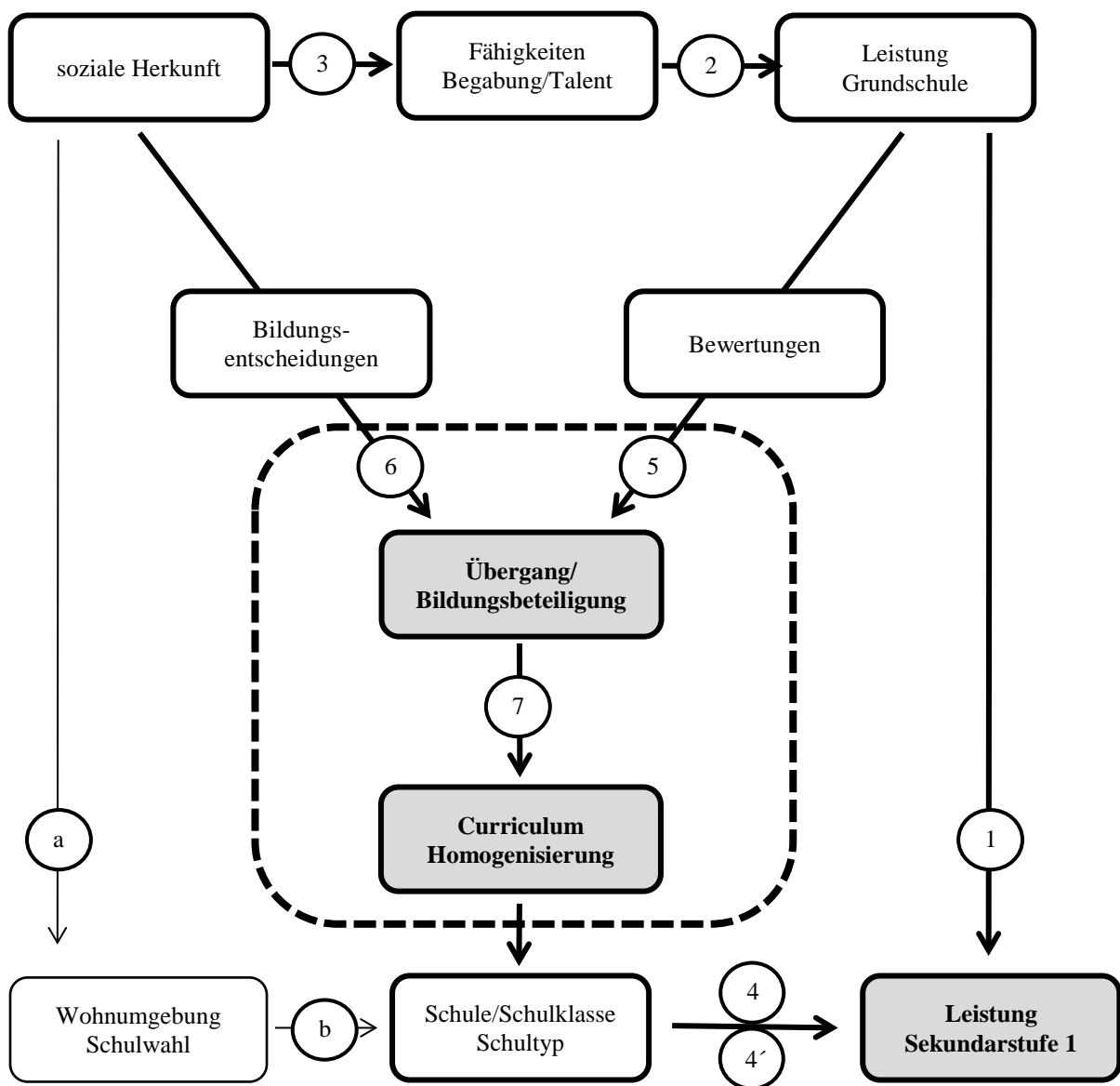
Ausgangspunkt ist ein übergreifendes theoretisches Modell für die Erklärung der Effekte von Bildungssystemen allgemein: Das Modell der Leistungsdifferenzierung (im Text gelegentlich abgekürzt mit: MoAbiT¹; vgl. ausführlicher Esser 2016a,b; Esser und Hoenig 2018).

¹ Akronym aus dem englischsprachigen Begriff „Model of Ability Tracking“.

Das Grundmodell

Das Modell bezieht sich auf den Verlauf eines typischen Bildungsweges von der kognitiven Entwicklung in der Familie über die Leistungen in der Grundschule, die Bewertungen in der Schule, den Übergang in die Sekundarstufe bzw. die Bildungsbeteiligung und schließlich die Leistungen dort (Abbildung 1).

Abbildung 1: Das Modell der Leistungsdifferenzierung (MoAbiT)



Das zentrale Explanandum sind die Leistungen in der Sekundarstufe. Der Grundvorgang für die Erklärung von Unterschieden darin ist das Lernen. Er bezieht sich auf einen gut belegten

allgemeinen Zusammenhang: Leistungen sind das Produkt der *Exposition* zu einem bestimmten Lerninhalt („Exposure“) mit der *Effizienz*, in das in ein Lernergebnis umgesetzt wird (Sörensen 1970, S. 357ff.; Sörensen und Hallinan 1977, S. 275ff.). Die Exposition besteht formal aus den mit dem jeweiligen Curriculum festgelegten Lerninhalten, faktisch aus dem jeweiligen Unterricht in dessen Rahmen. Die Effizienz hängt hauptsächlich von der jeweiligen *Fähigkeit* („Ability“) der Kinder für eine bestimmte Kompetenz ab. Bei den Fähigkeiten sind inhaltlich unterschiedliche *Begabungen*, etwa praktisch-spezifische und abstrahierend-generalisierende, und dafür jeweils differierende Grade an *Talent* zu unterscheiden: Es gibt für jede inhaltlich unterschiedliche Begabung Unterschiede und Begrenzungen im Talent, sie jeweils auszufüllen und zu entwickeln. Für die Effizienz des schulischen Lernens kommt es daher darauf an, dass die jeweilige Exposition inhaltlich über das Curriculum zu den Begabungen *und* im Grad der Schwierigkeit und des Niveaus des Unterrichts zu den unterschiedlichen Talenten passt, wenigstens in gewissen, auch durchaus weiten Grenzen der Erhöhung der Anforderungen. Es ist der Kern der Argumente für die Differenzierung – wie der für jeden „individualisierten“ Unterricht, der ja nichts Anderes wäre als eine „Differenzierung“, im Extremfall möglichst genau passend auf jeden einzelnen Fall.

Für eine gegebene Exposition steigt danach das schulische Lernergebnis zunächst direkt mit den Leistungen zuvor (Pfad 1 in Abbildung 1). Die Leistungen in der Grundschule hängen unmittelbar von den Fähigkeiten der Kinder (und anderen Einflüssen in der Grundschule) ab (Pfad 2). Die Fähigkeiten, Begabung wie Talent, sind teilweise genetisch vererbt, teilweise durch die Umstände in den Familien und den vorschulischen Einrichtungen geprägt, speziell über die soziale Herkunft (Pfad 3). Auf die Leistungen in der Sekundarstufe können die Fähigkeiten und die soziale Herkunft neben ihren indirekten Effekten über die Pfade 2 bzw. 3 auch direkte Effekte haben (in der Abbildung nicht weiter gekennzeichnet).

Schuleffekte

Die Einflüsse der sozialen Herkunft, der Fähigkeiten und der Grundschulleistungen bilden den Teil der *individuellen* Effekte auf die Leistungen in der Sekundarstufe. Hinzu treten *kontextuelle* Einflüsse der schulischen Lernumgebung in der Sekundarstufe: *Schul-Effekte* (Pfad 4; vgl. u.a. Sörensen und Hallinan 1977; Scheerens und Bosker 1997, Kapitel 2; Sacerdote 2011; Baumert, Watermann und Stanat 2006, Abschnitt 4.8; Huang 2009, Tabellen

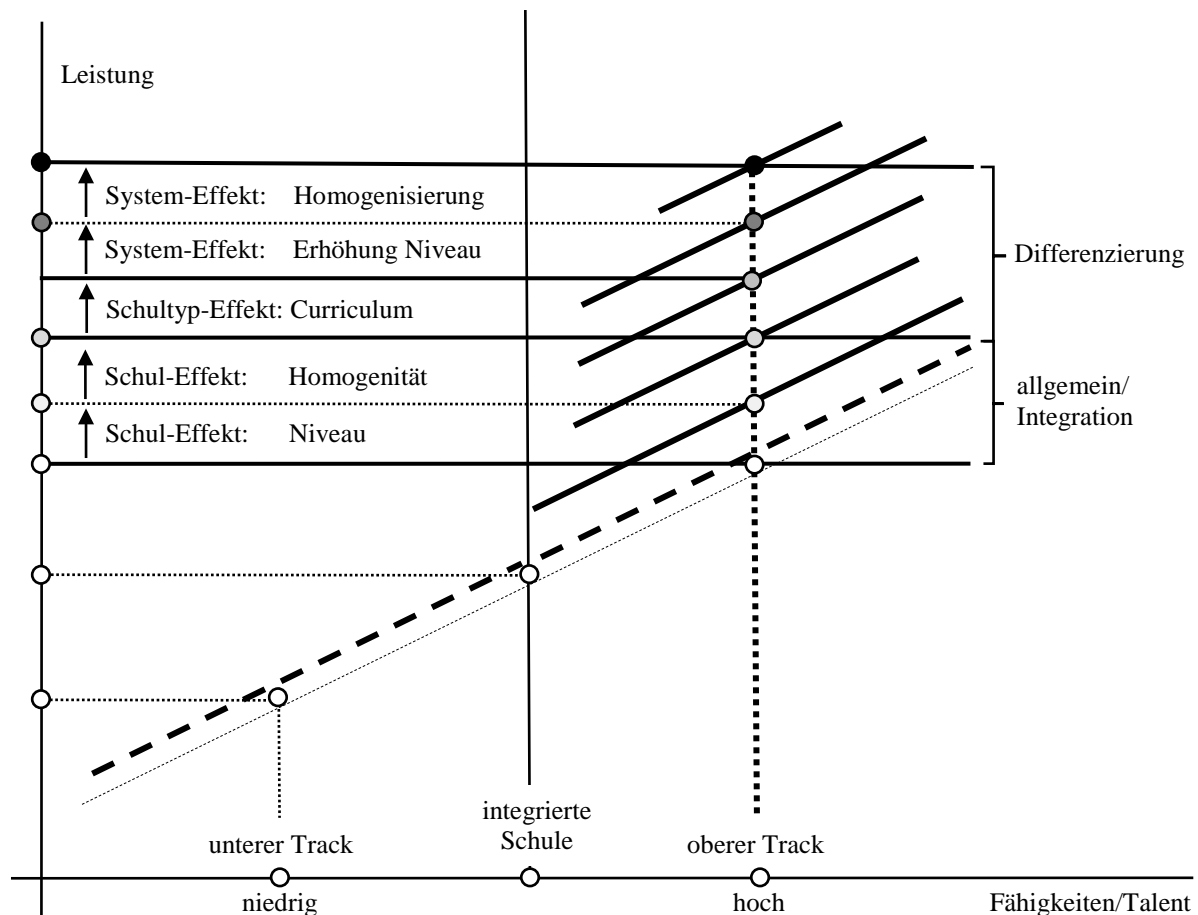
1 und 2: 786f.; Adoso und Agbayewa 2011: 50ff.; Becker, Lüdtke, Trautwein, Köller und Baumert 2012, S. 695f.; Schallock 2016, Abschnitt 4.3 und Tabelle 3: 326f). Sie beziehen sich im Wesentlichen auf vier Merkmale (in Abbildung 1 bei Pfad 4 nicht weiter gekennzeichnet): Die *Qualität* nach Ausstattung, Qualifikation des Lehrpersonals und sozialem Umfeld, die *Lerninhalte* nach dem jeweiligen Curriculum und der Fokussierung des Unterrichts darauf sowie die *Zusammensetzung* der Schulen und Schulklassen nach den jeweiligen *Fähigkeiten* und nach der *sozialen Herkunft*.

Qualität und soziale Zusammensetzung der Schulen und Schulklassen bilden eine vertikale, Curriculum, Fokussierung des Unterrichts und die Zusammensetzung der Schulen und Schulklassen nach den Fähigkeiten eine horizontale Dimension der Schuleffekte. Für die *vertikale* Dimension sind die theoretischen Erwartungen eindeutig: Mit der Qualität der Schulen und der Qualifikation des Lehrpersonals steigen die Effizienz und darüber das Lernergebnis generell, auch mit der Zusammensetzung nach der sozialen Herkunft und einem dadurch ggf. günstigeren ungünstigeren sozialen Umfeld. Für die *horizontale* Dimension der Differenzierung ist es differenzierter. Hier kommt es auf die *Kombination* von vier Aspekten an: Die inhaltliche Passung des *Curriculums* zu den spezifischen *Begabungen* der Kinder, etwa praktisch-spezifisch vs. abstrakt-generalisierend, das *Niveau* des jeweiligen Unterrichts im Schwierigkeitsgrad des Stoffs, dessen *Fokussierung* auf die jeweils vorgefundenen Umstände in der Verteilung der *Talente* und die *Passung* der Fokussierung auf die unterschiedlichen Fähigkeiten der Kinder nach *Begabung und Talent*.

Alles hängt demnach an der jeweiligen Verteilung der Kinder und ihrer relationalen Positionierung in den Schulklassen, und alles lässt sich daher immer nur in der jeweiligen Gewichtung ändern. Für eine jeweils gegebene Konstellation von Qualität und sozialer Zusammensetzung sowie der Fähigkeiten der Kinder in einer Schulklasse würde demnach ein höheres Niveau des Unterrichts zu einer Steigerung der Leistungen führen, ebenso wie eine stärkere Fokussierung des Unterrichts auf dieses Niveau und die Passung zu den Fähigkeiten. Damit werden für die Effizienz des Lernens zwei statistische Kennziffern der Zusammensetzung der Schulklassen bedeutsam: Der Mittelwert bzw. das *Niveau* der Verteilung der jeweiligen Talente als nächster Anhaltspunkt für das bei Unterricht anzusetzende Niveau und die Streuung um diesen Mittelwert bzw. die *Homogenität* für den Grad der möglichen Fokussierung des Unterrichts und der Passung an die Talente der Schülerinnen und Schüler. Mit Niveau und Homogenität der Zusammensetzung der

Schulklassen nach den Fähigkeiten sollten daher die Leistungen steigen. In Abbildung 2 ist das im unteren Teil der Graphik im Eintrag „Schuleffekte allgemein“ skizziert.

Abbildung 2: Schul- und System-Effekte der Leistungsdifferenzierung



Die unterste Linie bezeichnet den Zusammenhang zwischen Fähigkeiten bzw. Begabung und Talent und bestimmten Leistungen für eine Population von Schülerinnen und Schülern insgesamt und ohne jede Differenzierung wie in einer integrierten Schule. Es gebe nun eine Einteilung in zwei Gruppen unterschiedlicher Talente für eine gegebene Begabung (unterer/oberer Track; zwei parallele gebrochene Linien oberhalb). Das ändert an der Funktion noch nichts: Die Unterschiede in den Leistungen zwischen den Tracks sind allein die Folge der Unterschiede in den Fähigkeiten bereits. Erst dazu *zusätzliche* Unterschiede in den Leistungen würden als Schuleffekte verstanden. Für die beiden (positiven) allgemeinen Schuleffekte von Niveau und Homogenität der Zusammensetzung der Schulen und Schulklassen nach den Fähigkeiten ist das in der Abbildung über entsprechende Verschiebungen nach oben und einer Verringerung der Streuung skizziert. Die Effekte auf die

Leistungen sind über unterschiedliche Grautöne der Punkte auf der linken Achse hervorgehoben. Statistisch würde sich der betreffende Zusatzeinfluss der schulischen Umgebung in Resteffekten (von Niveau und Homogenität) nach Kontrolle der jeweiligen individuellen Beziehungen zeigen.

Differenzierung

Die individuellen wie die Schul-Effekte gelten zunächst *allgemein*: Es gibt *überall* Effekte der aus der Familie mitgebrachten Fähigkeiten der Kinder, der in der Grundschule erworbenen Kompetenzen und der sozialen Herkunft auf das Lernen und die Leistungen in der Sekundarstufe – auch also in integrierten Bildungssystemen. Und es gibt *überall* selektive Verteilungen auf Schulen und Schulklassen nach Qualität, Curriculum und speziellen Kursen, Unterricht, sozialer und kognitiver Zusammensetzung, etwa als Folge sozial-ökologischer Segregationen bei den administrativen Einzugsgebieten oder einer selektiven Schulwahl der Familien (Pfade a und b links in Abbildung 1). Mit der *Differenzierung* der Bildungswege kommen Bedingungen und Vorgänge einer auch *institutionell* gezielt organisierten Sortierung in unterschiedliche Bildungswege und Schultypen hinzu. Die Verteilung auf die Bildungswege ist dabei über zwei Vorgänge bestimmt: Die *Bewertung* der Leistungen in der Grundschule in Noten und Empfehlungen durch die *Schule* (Pfad 5) und die – davon ggf. auch abweichenden – *Bildungsentscheidungen* der *Eltern* (Pfad 6). Beides zusammen ergibt den Übergang in die Sekundarstufe, die *Bildungsbeteiligung*.

Zwei *strukturelle* Folgen für die Schulen und Schulklassen werden mit der Differenzierung erwartet: Die Einrichtung besonderer Schultypen mit inhaltlich unterschiedlich strukturierten *Curricula* in den Bildungswegen für die verschiedenen Begabungen, etwa in einen eher berufsbezogen-spezifischen und einen eher akademisch-generalisierenden Zweig und die verstärkte *Homogenisierung* der Schulklassen nach den Begabungen und Talenten (Pfad 7). Entsprechend der jeweils stärkeren Passung mit den jeweiligen Fähigkeiten sollten sich die Leistungen bessern: Mit der inhaltlichen Differenzierung der Curricula würden die unterschiedlichen Begabungen eher anzusprechen sein und den allgemeinen Effekt des Niveaus verstärken, und mit der zusätzlichen Homogenisierung wären eine noch stärkere Fokussierung und Passung des Unterrichts an die unterschiedlichen Talente möglich. Die Effekte daraus würden zu den entsprechenden allgemeinen Schuleffekten von Niveau und Homogenität *hinzutreten*. Das ist in Abbildung 1 über den Pfad 4' gekennzeichnet.

Schultypen

Die Effekte der Differenzierung würden sich entsprechend in Verbesserungen der Leistungen in den Schulklassen der dann eingerichteten *Schultypen* (wie Hauptschule, Realschule und Gymnasium oder bestimmte Fachschulen) zeigen (Abbildung 2, Eintrag „Effekt Curriculum: Schultyp“). Eine besonders umstrittene Frage ist nun, ob das für alle Tracks und Gruppierungen nach Fähigkeiten bzw. Leistungen gilt – oder ob es unterschiedlich strakte Gewinne oder Trade Offs derart gibt, dass die Verbesserungen der einen auf Kosten der anderen gehen. Für die oberen Tracks mit höheren kognitiven Anforderungen im Talent liegt es nach dem Modell auf der Hand: Es gibt dort ein höheres (akademisches) Niveau und eine stärkere Homogenität dafür. In Abbildung 2 ist das im Eintrag „Effekt Schultyp: Curriculum“ und über einen mittelgrauen Punkt gegenüber der Integration für den oberen Track dargestellt. Für die unteren Tracks ist das so nicht ohne Weiteres zu erwarten: Das (akademische) Niveau in den Curricula ist niedriger, es gäbe jedoch ggf. auch Spezialisierungsgewinne durch eine bessere Ausnutzung der jeweils anderen Begabungen, etwa praktisch-technischer Art. Eine entsprechende Homogenisierung würde das verstärken: über die stärkere Fokussierung und Passung des Unterrichts *darauf*.²

Theoretisch lässt sich der (Netto-)Effekt für die unteren Tracks also nicht eindeutig bestimmen (zu evtl. Aspekten der Stratifikation vgl. den Abschnitt über die Schuleffekte oben bereits). Es wird mit dem MoAbiT jedoch davon ausgegangen, dass die Effekte pareto-optimal sind: Gewinne der einen gehen *nicht* zu Lasten anderer, es gibt allenfalls relativ geringere Zuwächse für bestimmte Konstellationen, die aber ganz ausblieben, wenn es keine Differenzierung gäbe.

System-Effekte

Alle Annahmen beruhen darauf, dass es bei der Sortierung *allein* auf die Fähigkeiten bzw. die Leistungen vorher ankommt: *Ability Tracking*. Das ist für die Differenzierung nicht unbeschrieben gesichert. Abgesehen von Fragen der praktischen Implementation, auch in der Interaktion von Familien und Schulen, sind insbesondere zwei Regelungen dafür relevant: Die *Verbindlichkeit* der Empfehlungen und die *Kontrolle* der schulischen Abläufe. Die

² Zu den Vorgängen der Fokussierung und Passung des Unterrichts können peer- und Vergleichseffekte kommen: Kontakte und Vergleiche mit Kindern ähnlicher oder anderer Fähigkeiten und Schulklassen unterschiedlich in Niveaus und Homogenität (vgl. die Übersicht bei Schallock 2016).

Verbindlichkeit soll die sog. *sekundären Effekte* bei den Bildungsentscheidungen der Eltern begrenzen, indem es dann keine eigenen Bildungsentscheidungen der Eltern gäbe, die von den Empfehlungen der Schule abweichen, die *Kontrolle* betrifft die sog. *tertiären Effekte* bei den Noten und besonders bei den Empfehlungen durch das Lehrpersonal, weil dann besondere Anreize und zusätzliche Informationen zu möglichst leistungsgerechten Bewertungen führen sollten. Ohne diese Regelungen könnte es zu Fehlplatzierungen nach Begabung und Talent kommen, die die theoretisch zu erwartenden Effekte der Leistungsdifferenzierung verdünnen oder ggf. ganz überlagern, etwa in der Bevorzugung von Kindern aus den oberen oder aus bildungsnahen Schichten – unabhängig von dem was können oder interessiert.

Verbindlichkeit und Kontrolle wären damit zwei zentrale *institutionelle* Bedingungen dafür, dass die leistungsfördernden Effekte der Leistungsdifferenzierung tatsächlich eintreten. Jede Möglichkeit der Abweichung davon würde die Effekte abschwächen.³

Auf die Effekte *dieser* Regeln beziehen sich die spezifischen *System-Effekte* der Differenzierung als *Ability Tracking*. Um *sie* dreht es sich bei der Debatte vor allem. Es wird erwartet, dass sich mit den spezifischen Regeln einer *strikten* Differenzierung nach den Fähigkeiten, hier: die *Kombination* von *Verbindlichkeit* und *Kontrolle*, die allgemeinen Schulleffekte ändern und mit ihnen die Leistungen. Diese Kombination sei als *Stringenz* bezeichnet. Statistisch würde sich das in entsprechenden Interaktionseffekten zwischen den allgemeinen Bedingungen und den jeweiligen Regeln bzw. der Stringenz zeigen. In Abbildung 2 ist das im oberen Teil im Eintrag „System-Effekt“ wieder für die *Veränderung* der Effekte von Niveau und Homogenität dargestellt (dunkelgrauer und schwarzer Punkt).

2. Hypothesen

In Tabelle 1 sind die aus dem Modell der Leistungsdifferenzierung folgenden Hypothesen für die allgemeinen Zusammenhänge: die individuellen Effekte und die Schuleffekte, und für die dann hier speziell interessierenden System-Effekte, zusammengefasst.

³ Für differenzierende Systeme gibt es noch eine dritte Regelung: Das Alter des ersten Übergangs, etwa mit 10 Jahren oder später bzw. auch in der Pflichtschulzeit gar nicht. Theoretisch gibt es nach dem MoAbiT keine eindeutige Erwartung. Einerseits könnte es bei einer frühen Sortierung eher zu Fehlplatzierungen kommen, weil weder Lehrer noch Eltern so früh wissen könnten, wie die Entwicklung noch verläuft, andererseits würden die Vorteile der Differenzierung früher wirksam werden und über den Verlauf dann auch länger andauern.

Tabelle 1: Hypothesen zu den allgemeinen und zu den System-Effekten des Ability-Tracking auf die Leistungen in der Sekundarstufe nach dem MoAbiT (Kürzel zur Bezeichnung der Konstrukte in Text und Tabellen auch später)

	Kürzel Konstrukte	Hypothesen MoAbiT
allgemeine Effekte		
individuelle Effekte		
soziale Herkunft	SES	+
Fähigkeiten	ABL	+
Schul-/Schultyp-Effekte		
Niveau soziale Herkunft (Schulklasse)	NSES	+
Niveau Fähigkeiten (Schulklasse)	NABL	+
Homogenität Fähigkeiten (Schulklasse)	HABL	+
Curriculum/Schultyp	GYM	+
System-Effekte		
Differenzierung (Stringenz)	T	≥ 0
soziale Herkunft*Stringenz	SES *T	≤ 0
Niveau soziale Herkunft*Stringenz	NSES*T	≤ 0
Fähigkeiten*Stringenz	ABL *T	0
Niveau Fähigkeiten*Stringenz	NABL*T	+
Homogenität Fähigkeiten*Stringenz	HABL*T	+

Allgemein haben die soziale Herkunft (SES) und die Fähigkeiten (ABL) einen positiven Einfluss auf die Leistungen in der Sekundarstufe, ebenso wie bei den Schuleffekten das Niveau der sozialen Herkunft (NSES; über die dann insgesamt besseren materiellen und sozialen Bedingungen) und das Niveau und die Homogenität der Zusammensetzung der Schulklassen nach den Fähigkeiten (NABL, HABL). Die mit der Differenzierung verbundenen Schultypen (GYM) haben dazu *zusätzliche* positive Effekte auf die Leistungen, die sich mit der besseren Anpassung schon der Curricula für die unterschiedlichen Begabungen einstellen.

Die *System-Effekte* der Differenzierung beziehen sich auf die speziellen Regeln der Stringenz der Sortierung über Verbindlichkeit und Kontrolle. Hierfür sind generell als Haupteffekt (T) mindestens positive Effekte zu erwarten. Für Veränderungen der allgemeinen Bedingungen,

ausgedrückt über die entsprechenden Interaktionseffekte, ist *keine* Verstärkung der stratifizierenden Einflüsse, soziale Herkunft bzw. soziale Zusammensetzung der Schulen und Schulklassen, zu erwarten, eher Abschwächungen (SES*T, NSES*T). Die Effekte der individuellen Fähigkeiten (ABL*T) sollten sich nicht weiter verändern – sie sind über die Regelungen der Differenzierung gleich verteilt und insgesamt ohnehin kontextuell relativ stabil. Allerdings sollte es positive System-Effekte für das Niveau und die Homogenisierung der Zusammensetzung der Schulklassen nach den Fähigkeiten geben (NABL*T, HABL*T). Es wäre der Kern der Implikationen des Modells der Leistungsdifferenzierung für die System-Effekte der Differenzierung nach den Fähigkeiten. Unterschiede in den Effekten zu Ungunsten von Kindern aus schwierigeren sozialen Verhältnissen oder schwächeren Leistungen werden nicht erwartet.

3. Befunde und (offene) Fragen

Zu den System-Effekten der Differenzierung auf die Leistungen gibt es Befunde aus sehr unterschiedlichen Arten von Beiträgen: Unmittelbare Vergleiche von Schulen, internationalen Vergleiche speziell auf der Basis der internationalen Vergleichsstudien wie PISA oder TIMMS, regionale Vergleiche wie in der Schweiz für die Kantone und in Deutschland für die Bundesländer und Ergebnisse aus Beiträgen zu Reformen und (Feld-)Experimenten und zu den Effekten homogener und heterogener Schulklassen.

Die *unmittelbaren Vergleiche* beziehen sich überwiegend auf die Frage eines Trade Offs zwischen Gewinnen der Kinder nach der Einteilung in die oberen Tracks und Verlusten mit der Sortierung nach unten. Die Ergebnisse der Beiträge aus den 80er und 90er Jahren vorwiegend aus den USA sind gemischt: Einerseits werden für die Differenzierung Schereneffekte beobachtet, nach denen die Sortierung nach oben im Vergleich deutliche Leistungssteigerungen erbringe. Die integrierten Schulen liegen im Niveau dazwischen. Dabei wurden jedoch für die integrierten Schulen oft keine Schuleffekte berücksichtigt. In den Studien, die das taten, zeigten sich andererseits Vorteile für die höheren Tracks ohne Nachteile bei niedrigeren Tracks und mit Gewinnen der Differenzierung sogar (vgl. die Verweise auf die frühen Beiträge von Kerckhoff, Hoffer, Argys et al., Figlio und Page sowie Betts und Sholnick in der Übersicht bei Betts 2011, Abschnitt 4.1). In einer neueren Studie

von Horn (2013) für Ungarn zeigt sich ein positiver Effekt der Differenzierung gegenüber der Integration und eine Steigerung des Effekts des Fähigkeits-Niveaus in den Schulklassen.

Bei den *internationalen Vergleichsstudien* ergab sich für die Länder mit starker und früher Differenzierung ein durchaus verheerendes Bild (Baumert und Schümer 2001): Ein deutlich geringeres Leistungsniveau, eine starke soziale Undurchlässigkeit und – insbesondere – die Benachteiligung von leistungsschwächeren Kindern oder solchen aus schwierigeren sozialen Verhältnissen. Dieser Eindruck bestätigte sich in einer ganzen Serie nachfolgender, methodisch sorgfältig angelegter multivariater Analysen (Ammermüller 2005, Hanushek und Wößmann 2006, Horn 2008, Schütz et al. 2008, Wößmann et al. 2009, Le Donné 2014). Aber es gab auch von Beginn an Inkonsistenzen und widersprechende Ergebnisse. Waldinger (2007) stellte über den Einbezug von Daten für die Grundschule (über PIRLS) fest, dass die Ungleichheiten nach der sozialen Herkunft in den differenzierenden Ländern schon *vor* der Sortierung bestanden. Mit dem Einbezug der Schulen und Schulklassen zeigte sich sogar eine Verringerung des Herkunftseffekts innerhalb der Schulen (Dunne 2010, Dronkers et al. 2011, 2012). Weil auch hier, wie bei den anderen Beiträgen, die kognitiven Fähigkeiten und die entsprechende Selektivität nicht erfasst worden waren, blieb jedoch unklar, was überwiegt und ob es überhaupt einen Herkunftseffekt gibt. Ähnliches gilt für den Beitrag von Bol et al. 2014, wo einmal ohne und einmal mit Einbezug der Schuleffekte gerechnet wurde und sich die Effekte der sozialen Herkunft jeweils in der Tat drehten. Korthals und Dronkers (2016) schließlich fanden einen positiven Effekt der Differenzierung, wenn bei der Sortierung die von den Schulleitungen angegebene Selektivität kontrolliert wurde.

Bei den *regionalen Vergleichen* bestätigte sich für *Deutschland* zunächst in den PISA-Zusatzstudien die Standardposition: Im Vergleich (sehr) geringe Leistungen und eine hohe soziale Undurchlässigkeit für die besseren Leistungen (vgl. etwa Artelt et al. 2002). Ab 2009 traten an die Stelle der PISA-Zusatzauswertungen die von der KMK organisierten IQB-Berichte für 2009 (Böhme und Weirich 2010) und 2015 (Schipolowski und Böhme 2016). Daraus ging hervor, dass die strikter differenzierenden Bundesländer (Bayern, Baden-Württemberg, Sachsen) in den Leistungen deutlich besser waren, gerade auch bei den leistungsschwächeren und den Migrantenkindern, als die eher liberalen Regimes (in Berlin, Bremen und Hessen etwa). Die (bisher einzige) auch multivariate Analyse zu den Effekten der Differenzierung über PISA-Daten für alle 16 Bundesländer von (Wößmann 2010) kam zu dem gleichen Ergebnis für die Effekte einer frühen Sortierung. Eine multivariate Analyse

unter Einschluss aller nach dem MoAbiT bedeutsamen Bedingungen, insbesondere der kognitiven Fähigkeiten und der Schuleffekte ergab jedoch (mit den Daten der BiKS-Studie) für Bayern mit seinem strikten System im Vergleich zu Hessen einen positiven Effekt (Esser und Relikowski 2015). Für die *Schweiz* scheint es nach Stadelmann-Steffen (2012) keinerlei Effekt der Differenzierung zu geben. Das zeigt sich bei Felouzis und Charmillon (2013, 2017) anders: Es gebe erhebliche Vorteile für die Kinder in den anspruchsvollen Schultypen und eher sogar noch weiter verstärkter Nachteile in den Basis-Zweigen – ganz so wie in den frühen Studien bei Kerckhoff, Hoffer oder Argys et al. auch gefunden wurde.

Die meisten Beiträge zu den Effekten von *Reformen* beziehen sich nicht auf die Sekundarstufe, sondern auf spätere Folgen, insbesondere auf dem Arbeitsmarkt. In einem unmittelbaren Vergleich erhöht nach Galienda-Rueda und Vignoles (2004) die Differenzierung das Leistungsniveau, ohne dass es zu Nachteilen für die Kinder aus den unteren Leistungsbereichen käme. Manning und Pischke (2006) kritisieren dieses – nicht nur damals – ungewöhnliche Ergebnis dahingehend, dass die Frage nach Selektivitätseffekten nicht hinreichend geklärt wäre. Der Beitrag von Kerr et al. (2010/2012) bezieht sich auf die Umstellung des Bildungssystems in Finnland auf die Integration. Das Ergebnis unterstützt die Standardposition, wenngleich nicht eindeutig. Es zeigt sich ein positiver, aber auch nur schwacher Effekt der Reform auf die Entwicklung der mathematischen Fertigkeiten, nicht aber bei den sprachlichen, und die Umstellung verbessert in der Tat die Fertigkeiten der Kinder aus den unteren Schichten. Alle derartige (Feld-)Experimente zu Reformen haben das Problem, dass sich mit den Reformen u.U. konkomitant andere Umstände verändern. Das Experiment von Duflo et al. (2011) vermeidet das über Randomisierung. Diese Befunde sprechen deutlich gegen die Standardposition: Die Differenzierung erhöht das Leistungsniveau, wovon alle profitieren, auch die Kinder in den unteren Leistungsbereichen und solche, die bei der Sortierung nur knapp eine höhere (oder niedrige) Einteilung verpasst haben. Schließlich gibt es noch einen Beitrag zur Vorverlegung der Sortierung von 12 Jahren auf 10 in Bayern im Jahr 2000 (Piopionuk 2014). Danach hat sich das Leistungsniveau mit der Umstellung auf die frühe Differenzierung *verringert*, allerdings immer noch auf einem auch dann noch sehr hohen Gesamtniveau. Als Erklärung werden Veränderungen in den Motiven des Lehrpersonals vermutet.

Zu den Wirkungen der *kognitiven Homogenität* bzw. der Homogenisierung durch die Differenzierung gibt es nur wenige Analysen. Der Beitrag von Zimmer (2003) belegt positive

Effekte der Homogenität wie der Homogenisierung, allerdings ohne Berücksichtigung des jeweiligen kognitiven Niveaus in den Schulklassen. Nach Lichewski und Kutscher (1996) gibt es eine Verbesserung der Leistungen mit der Zuteilung zu leistungsstärkeren Gruppierungen, das aber zulasten der Kinder mit den geringeren oder mittleren Leistungen. Das entspräche der Standardposition. Allerdings ist die Operationalisierung der „Homogenität“ ungewöhnlich: Das seien Klassen mit Unterricht als „multivocal interactions“, also gerade das Gegenteil einer Fokussierung wie sie mit der Homogenität eigentlich angenommen wird. Wenn das so wäre, würde es eine Widerlegung der Standardposition bedeuten. Bei Huang (2009), der die Effekte der Homogenität für unterschiedliche Niveaus untersucht, kommt ein Trade Off heraus: Es gibt Homogenitätsgewinne, aber die beschränken sich auf Kinder mit höheren kognitiven Fähigkeiten und Schulklassen mit höherem kognitivem Niveau.

Es ist kaum möglich, aus den vorliegenden Befunden ohne eine theoretische Richtschnur einigermaßen eindeutige Schlüsse zu ziehen – wohl ein Hauptgrund dafür, dass die Debatte kontrovers geblieben ist. Auffällig ist die erhebliche Unterschiedlichkeit der Ansätze, der berücksichtigten Bedingungen, der verwendeten Daten und der Formen der Analyse, insbesondere aber, dass gerade die so wichtigen Interaktionseffekte zwischen Differenzierung und den allgemeinen Bedingungen nur selten erfasst wurden. Das gilt speziell für die PISA-basierten Beiträge. Hier stellt sich bis auf drei Ausnahmen stets der Standardbefund ein. Die Ausnahmen davon sind Waldinger (2007), der über PIRLS kontrolliert, welchen Einfluss die soziale Herkunft schon in der Grundschule hatte, Dunne bzw. die DVD-Beiträge und Bol et al. (2014) mit dem Einbezug der Schulebene und Korthals und Dronkers (2016) mit der Berücksichtigung der Selektivität nach Leistungen. Stets sind es andere Daten als PISA oder die Analysen gehen anders vor, etwa mit dem Einbezug der Schulebene oder der Selektivität. In allen diesen Fällen entsprechen die Ergebnisse, wie für das randomisierte Experiment bei Duflo et al. (2011), dem Beitrag von Horn (2013) für Ungarn oder dem Vergleich von Hessen und Bayern bei Esser und Relikowski (2015) den Hypothesen des MoAbiT – und widersprechen damit der Standardposition.

4. Daten, Operationalisierungen und Analysen

Grundlage der Analysen zum Test des Modells sind Daten der „National Educational Panel Study“ (NEPS).⁴ Die hier betrachtete Kohorte verfolgt Schülerinnen und Schüler, die im Schuljahr 2010/2011 in Deutschland die fünfte Klasse besuchten, in jährlichen Befragungen und Kompetenztests. Die verwendeten Variablen entstammen der ersten bis dritten Welle einer schriftlichen Befragung und Kompetenztestung der Schülerinnen und Schüler im Herbst 2010 und einer telefonischen Befragung jeweils eines Erziehungsberechtigten im Frühjahr 2012.

4.1 Variablen und Verteilungen

Eine Übersicht über alle verwendeten Variablen und wichtige Kennzahlen findet sich in Tabelle 2. Die Aufstellung folgt der Reihenfolge der Konstrukte des Modells der Leistungsdifferenzierung nach Abbildung 1 oben.

Tabelle 2: Wertebereich, Mittelwerte und Streuungen der verwendeten Konstrukte (Erläuterung der Kürzel im Text)

Konstrukte	min/max	Welle 1 Analysen Übergang		Welle 3 Analysen Leistungen Sekundarstufe	
		Mittelwert	Streuung	Mittelwert	Streuung
SES	16-90	53.1	16.8	53.7	16.7
ABL	0-1	0.59	0.21	0.60	0.21
ACE	-2.9-3.3	0.14	1.00		
MRK	1-6	0.74	0.15		
REC	0-1	0.57	16.8		
ASP	0-1	0.74	0.44		
GYM	0-1	0.55	0.48	0.58	0.49
ACH	-3.2-4.4			0.69	1.08
NSES	16-75			51.4	8.9
NABL	0.17-0.83			0.60	0.11
T	0-1	5.04	1.36	5.06	1.38

⁴ Startkohorte 3, Klasse 5-7, doi:10.5157/NEPS:SC3:6.0.0. Die Daten des NEPS wurden von 2008 bis 2013 als Teil des Rahmenprogramms zur Förderung der empirischen Bildungsforschung erhoben, welches vom Bundesministerium für Bildung und Forschung (BMBF) finanziert wurde (siehe auch Blossfeld et al. 2011). Seit 2014 wird NEPS vom Leibniz-Institut für Bildungsverläufe e.V. (LIfBi) an der Otto-Friedrich-Universität Bamberg in Kooperation mit einem deutschlandweiten Netzwerk weitergeführt.

FEM	0-1	0.47	0.50	0.48	0.50
MHG	0-1	0.31	0.46	0.29	0.46
VSB	0-83	39.2	11.0	39.5	11.0
N		3032		2717	

Der *soziale Hintergrund* (SES) wird als der höchste Status bei den Eltern, gemessen über den Index ISEI-88 (nach Ganzeboom et al. 1992) erfasst. Als Maß für die *kognitive Entwicklung* (ABL) wird der für NEPS entwickelte NEPS-MAT verwendet, ein Matrizen-Test des schlussfolgernden Denkens (Lang et al., 2014). Die Leistungen in der Grundschule wurden zu Beginn der fünften Klasse über die Kompetenzen in Orthographie, Leseverständnis und Mathematik erfasst. Die WLE-Schätzer dieser drei Kompetenzen bilden die Grundlage für den Index der Leistungen am Ende der Grundschule (ACE), in dem die beiden sprachlichen Kompetenzen (Orthographie und Leseverständnis) jeweils mit 25% und die mathematische Kompetenz mit 50% gewichtet sind. Die *Noten* (MRK) werden über den Mittelwert der von den Eltern berichteten Noten in Deutsch und Mathematik am Ende der vierten Klasse gemessen. Der Notendurchschnitt wurde zur einfacheren Vorzeicheninterpretation von ungenügend (Schulnote 6 = 1) bis sehr gut (Schulnote 1 = 6) rekodiert. Im Falle fehlender Elternangaben wurden die Angaben der Kinder verwendet. Die *Empfehlung* am Ende der Grundschulzeit (REC) wurde auf Basis der Angaben der Eltern als Dummy (Gymnasium mit 1 vs. kein Gymnasium mit 0) gebildet. Die drei spät sortierenden Länder Berlin, Brandenburg und Mecklenburg-Vorpommern wurden von den Analysen ausgeschlossen. Zur Erfassung der familiären *Aspirationen* (ASP) wurden die Angaben der Kinder in der ersten Welle (Klasse 5) für deren idealistische Aspirationen auf das Gymnasium genommen. Der Index dient als Proxy auch für die Aspirationen der Eltern insofern die Einschätzungen zwar nicht perfekt übereinstimmen, aber, wie auch sonst zu finden ist, stark korrelieren. Die Elternangaben wurden nicht verwendet, da sie erst in der zweiten Welle (Klasse 6) erhoben wurden, also über ein Jahr nach dem Übergang. Die Ergebnisse änderten sich bei unterschiedlichen Verwendungen der Angaben so gut wie nicht. Der *Übergang* auf das Gymnasium (GYM) bzw. die Bildungsbeteiligung wird als Dummy der zu Beginn der fünften Klasse besuchten Schulform (Gymnasium mit 1 vs. kein Gymnasium mit 0) operationalisiert, wobei Kinder, die den Gymnasialzweig einer kooperativen Gesamtschule besuchen, ebenfalls als Gymnasiasten klassifiziert sind. Die *Leistungen in der Sekundarstufe* (ACH) werden über einen Index aus Lesen und Mathematik analog zu dem für die Leistungen in der Elementarschule (ACE oben) erfasst. Die Mittelwerte von sozialer Herkunft und der kognitiven Fähigkeiten bilden das

sozialen und das kognitive Niveau der Schulklassen (SSES bzw. SABL) in der Sekundarstufe, die entsprechenden Standardabweichungen die Heterogenität bzw. Homogenität (HSES bzw. HABL).

Den Kern der Analysen zu den System-Effekten bildet ein Index der Stringenz als Kombination von Regeln, die die sekundären und die tertiären Effekte beeinflussen (vgl. Abschnitt 1). Die Zuordnung folgt dem Vorgehen bei den Typologien bei von Below (2011, Abschnitt 4.3) und der daran anschließenden, breit begründeten Klassifikation bei Helbig und Nikolai (2015; Abbildung 28: S. 286) in zwei grundlegende Dimensionen: Modernisierung vs. Traditionalität und De-Standardisierung vs. Standardisierung. Tabelle 3 gibt die Verteilungen für die 16 Bundesländer wieder, zunächst für den Index mit fünf Werten, dann für den unmittelbaren Vergleich in getrennten Analysen für drei Werte.

Tabelle 3: Zuordnung der Bundesländer zu den Werten des Index der Stringenz (Index aus der Aufstellung bei Helbig und Nikolai 2015, Abbildung 26; vgl. Abbildung 5.2 oben)

Index Stringenz	1		2		3		4		5	
	liberal								strikt	
	BE HB HE		MV NW RP		BB HH NI		BY		BW SN	
	SH				SL ST TH					
Welle	1	3	1	3	1	3	1	3	1	3
n	379	352	908	801	631	553	429	354	685	648
N (Welle1/Welle3)	3032/2717									
Typen Stringenz	1		2		3					
	liberal						strikt			
	BE HB HE		BB HH MV NI NW		BW BY SN					
	RP SH		SL ST TH							
Welle	1	3	1	3	1	3				
n	532	525	1386	1181	1114	1012				
N (Welle 1/Welle3)	3032/2717									
BB Brandenburg, BE Berlin, BW Baden-Württemberg, BY Bayern, HB Bremen, HE Hessen, HH Hamburg, MV Mecklenburg-Vorpommern, NI Niedersachsen, NW Nordrhein-Westfalen, RP Rheinland-Pfalz, SH Schleswig-Holstein, SL Saarland, SN Sachsen, ST Sachsen-Anhalt, TH Thüringen										

Etwas vereinfachend gesehen geben die Indexwerte Kombinationen der Einschränkung der Wahlfreiheit der Eltern, der Kontrolle von Leistungsgerechtigkeit und Leistungen und einer

frühen Sortierung wieder. Die beiden Extremtypen bilden danach Berlin, Bremen, Hessen und Schleswig-Holstein einerseits und Baden-Württemberg und Sachsen andererseits. Die drei Typen der Stringenz nehmen jeweils diese Extremwerte, aus Gründen möglichst ähnlicher Fallzahlen noch um drei andere Zuordnungen ergänzt: Der liberale System-Typ um Mecklenburg-Vorpommern und Rheinland-Pfalz, bei denen es auch weder Verbindlichkeit, noch Kontrolle gibt, und der strikte System-Typ um Bayern, das von allen anderen noch am nächsten zu Baden-Württemberg und Sachsen positioniert zu den strikten Regeln positioniert ist.

Die Verteilungen sind nicht weiter auffällig. Zwischen den Wellen gibt es kaum Verschiebungen, allenfalls steigen die eher günstigen Umstände: Der Anteil von Kindern in den Gymnasien nimmt in der 3. Welle leicht zu, aber der aus den oberen Schichten verändert sich kaum. Auch gehen bleiben kaum weniger Migranten in der erfassten Auswahl. Größere Verzerrungen scheint es also nicht zu geben, auch wenn die Fallzahl im Vergleich um etwa 10% kleiner wird. Das Analysesample umfasst dann die Fälle mit gültigen Angaben für alle Modellvariablen (listwise deletion). Zur Kontrolle wurden verschiedene andere Verfahren der Ersetzung fehlender Werte gerechnet. Die Ergebnisse unterscheiden sich im Kern nicht wesentlich voneinander, auch nicht zu vergleichbaren anderen Analysen (Esser und Relikowski 2015, Esser und Hoenig 2018).

4.2 Analysen

Für die Bestimmung der Effekte wurden bei metrischen Variablen OLS-Regressionen, bei kategorialen Variablen logistische Regressionen berechnet, für die Analysen zum Übergang auch als konditionale durchschnittliche Marginaleffekte (Average Marginal Effects, AME) und zum Vergleich der Koeffizienten für genestete Logit-Modelle das KHB-Verfahren. Bei der Berücksichtigung von Schulklasseneffekten (bei der Analyse der Leistungen in der Sekundarstufe) wurden zur Kontrolle auch Verfahren der Mehr-Ebenen-Analyse eingesetzt. Auch hier unterscheiden sich die Befunde nicht wesentlich von denen mit der Regression mit den entsprechend aggregierten Schulklassen-Merkmalen (Niveau und Streuung). Zur Analyse unterschiedlicher Effekte bei verschiedenen Leistungsstufen der Kinder wurde das Verfahren der Quantilsregression eingesetzt. In allen Modellen wird für *Geschlecht* des Kindes (FEM), den *Migrationshintergrund* (MHG, mit 1 gegen 0) und die Dauer des *Vorschulbesuchs* vor der

Einschulung (VSB) kontrolliert, die Einzelergebnisse dazu jedoch aus Platzgründen nicht eigens ausgewiesen.

5. Ergebnisse

Das Ziel der Analysen ist die Identifikation der System-Effekte der Leistungsdifferenzierung auf die Leistungen in der Sekundarstufe unter Kontrolle der dabei beteiligten allgemeinen Bedingungen und Vorgänge nach dem in Abschnitt 1 vorgestellten theoretischen Modell. Dazu werden – grob gesehen in der kausalen Reihenfolge der Kausalbeziehungen des theoretischen Modells – die Selektivitäten beim Übergang nach der Grundschule auf die verschiedenen Bildungswege (Abschnitt 5.1), die dabei eintretende kognitive Homogenisierung der Schulklassen (Abschnitt 5.2), die System-Effekte auf die Leistungen in der Sekundarstufe (Abschnitt 5.3), die Einflüsse von Homogenität und Homogenisierung der Schulklassen durch eine striktere Sortierung (Abschnitt 5.4) und schließlich die Frage behandelt, ob die Effekte der Homogenisierung die leistungsschwächeren Kinder benachteiligt oder nicht (Abschnitt 5.5).

5.1 Der Übergang

Tabelle 4 gibt die Ergebnisse zu den allgemeinen und zu den System-Effekten einer unterschiedlichen Stringenz der Regeln der Sortierung beim Übergang auf das Gymnasium nach der Grundschule wieder.

Tabelle 4: Allgemeine und System-Effekte beim Übergang auf das Gymnasium (T: strikte Differenzierung; NEPS, SC 3, Welle 1; logistische Regression; AME und KHB-Methode; Kontrolle von Geschlecht, Migrationshintergrund und Vorschulbesuch; fett: $p \geq 0.05$)

	allgemeine Effekte					System-Effekte	
soziale Herkunft	0.048	0.045	0.034	0.029	0.021	0.021	0.011
kognitive Fähigkeit		3.704	1.082	0.765	0.615	0.635	0.654
Leistung Grundschule			1.588	1.110	1.015	0.995	-0.011
Empfehlung				2.742	2.420	2.419	2.445
Aspirationen					3.333	3.373	3.420
T						0.073	-0.063
T* soziale Herkunft							0.002
T*Leistung Grundschule							0.204
C	-2.98	-4.96	-3.01	-4.08	-5.93	-5.84	-5.78
(Pseudo-)R ²	0.01	0.18	0.34	0.50	0.58	0.58	0.58
N	3032	3032	3032	3032	3032	3032	3032

Bei den *allgemeinen* Effekten zeigen sich für den Übergang auf das Gymnasium die gewohnten Befunde: Die soziale Herkunft hat ohne weitere Kontrollen einen deutlichen Einfluss. Das ändert sich kaum, wenn die kognitiven Fähigkeiten kontrolliert werden, die ihrerseits den weitaus stärkeren Einfluss haben. Die Empfehlungen und besonders die Aspirationen erklären einen weiteren Teil der Herkunftseffekte, die Aspirationen jene aus den sekundären Effekten der Bildungsentscheidungen der Eltern, die Empfehlungen die aus den tertiären Effekten der Bewertungen durch die Schule. Es bleibt aber (soziale Herkunft mit 0.021) immer noch ein eigener, so nicht weiter erklärbarer Rest an Herkunftseffekten. Einen eigenen *System*-Effekt der Stringenz der Differenzierung gibt es hingegen offenbar *nicht* (T mit 0.063, n.s.). Die Analyse mit Einschluss der Interaktions-Effekte zeigt, dass sich der Einfluss der sozialen Herkunft mit der Stringenz der Differenzierung *nicht* ändert (T*soziale Herkunft mit 0.002, n.s.), wohl aber wird die Kopplung des Übergangs an die Leistungen in der Grundschule, die Leistungsgerechtigkeit also, stärker (mit 0.204 für T*Leistung Grundschule). Es ist die auf das Konstrukt der der Stringenz bezogene Bestätigung eines ähnlichen Befundes für die Effekte speziell der Verbindlichkeit (Esser und Hoenig 2018, Tabelle 5: 436). Interessant ist, dass es nun keine direkten Effekte mehr gibt, weder solche der sozialen Herkunft, noch der kognitiven Fähigkeiten aus der Grundschule: Alles ist über die

Aspirationen bzw. Empfehlungen erklärt, und mit der Stringenz verstärkt sich der Einfluss der Leistungen in der Grundschule auf den Übergang, was ja das Ziel der Regelungen ist.

Im getrennten Vergleich von drei Gruppen von Bundesländern mit unterschiedlicher Stringenz werden einige zusätzliche und bemerkenswerte Details in den Unterschieden der Effekte erkennbar (Tabelle 5).

Tabelle 5: Der Übergang auf das Gymnasium im System- Vergleich; (T: strikte Differenzierung; NEPS, SC 3, Welle 1; logistische Regression; Kontrolle von Geschlecht, Migrationshintergrund und Vorschulbesuch; fett: $p \geq 0.05$)

	1			2		
	Differenzierung			Differenzierung		
	liberal		strikt	liberal		strikt
soziale Herkunft	0.036	0.044	0.051	0.018	0.022	0.019
kognitive Fähigkeit	3.179	3.738	4.116	-0.134	1.009	0.307
Leistungen Grundschule				0.937	0.780	1.573
Empfehlung				1.484	2.717	2.920
Aspirationen				2.464	2.888	4.546
C	-3.31	-5.57	-5.08	-3.57	-6.47	-6.43
(Pseudo-)R ²	0.14	0.20	0.20	0.41	0.59	0.69
N	533	1386	1114	533	1386	1114

Betrachtet man die Brutto-Effekte der kognitiven Fähigkeit allein (Modell 1), dann zeigt sich zunächst in der Tat ein Anstieg des Einflusses des Herkunftseffekts (von 0.036 über 0.044 auf 0.051), aber auch eine deutlich engere Kopplung an die *kognitiven* Fähigkeiten (von 3.179 über 3.738 auf 4.116). Es ist die Folge der Leistungsgerechtigkeit bei der Sortierung nach den Leistungen in der Grundschule, wie sie sich in Tabelle 4 zeigte. Dass sich damit gerade auch die Bindung an die kognitiven Fähigkeiten ändert, ist eine wichtige Voraussetzung für das eigentliche Ziel der Differenzierung: Die Sortierung nach den eher *latenten* Talenten, und eben nicht nur nach den am Ende der Grundschule *sichtbaren* Leistungen. Es wäre die zentrale kausale Grundlage für die angestrebte *strukturelle* Verstärkung der kognitiven Homogenität, über die mit der dann eher möglichen Fokussierung des Unterrichts eine Verbesserung des Leistungsniveaus eintreten sollte.

Mit der Kontrolle der primären, sekundären und tertiären Effekte über die Leistungen in der Grundschule, den Aspirationen und den Empfehlungen (Modell 2) verschieben sich die Gewichte wieder: Die Herkunftseffekte vermindern sich erheblich und es gibt keinen systematischen System-Effekt mehr und die Leistungen in der Grundschule bestimmen den Übergang mit der Stringenz deutlich, ebenso wie dann auch die Empfehlungen. Das verweist auf die Zunahme des meritokratischen Elements bei einer stringent leistungsbezogenen Differenzierung: Die Leistungen in der Grundschule und die Empfehlungen der Schule erklären jeden Effekt der kognitiven Fähigkeiten – weil sie zusammen mit den Noten offenbar bruchlos die Einflüsse der kognitiven Fähigkeiten auf den Übergang übertragen und keinen unmittelbaren Effekt darüber hinaus übriglassen. Den Schulen in den stringenten Systemen gelingt also ein hohes Ausmaß an objektiv richtiger Zuordnung, jedenfalls im Vergleich mit den offenen Systemen.

Durchbrochen wird diese stark meritokratische Vermittlung offenbar aber über die Effekte der Aspirationen der Eltern. Die machen sich gerade in den striktesten Bundesländern bemerkbar. Gerade das aber sollte eigentlich nicht sein. Womöglich liegt es daran, dass es in den strikteren Regimes eher sogar noch Möglichkeiten der Abweichungen von den Empfehlungen gibt, die von zwei Seiten mit einem ähnlichen Effekt eintreten können: Bei den oberen bzw. bildungsnäheren Schichten in der stärkeren Nutzung von Nachprüfungen o.ä., womöglich aber auch gewissen Interventionen vorher bei der Schule einerseits, bei den unteren bzw. bildungsferneren Schichten darüber, dass sie gerade dann, wenn es ernster wird mit dem Leistungsbezug, noch vorsichtiger werden als ohnehin schon und die *institutionell* immer gegebene Möglichkeit nutzen, gerade ihre talentierten Kinder *mit* einer Empfehlung *nicht* auf das Gymnasium zu lassen andererseits. Darauf gibt es empirische Hinweise (Esser und Hoenig 2018, Abschnitt 6.3).

5.2 Homogenisierung

Die engere Kopplung des Übergangs an die kognitiven Fähigkeiten wirkt sich in der Tat dann auf die strukturelle Homogenisierung der Schulklassen aus (Tabelle 6). Wir betrachten beide Aspekte der Sortierung: Die Stratifikation einer stärkeren Homogenisierung nach der sozialen Herkunft und die Differenzierung als eine stärkere Homogenisierung nach den kognitiven Fähigkeiten, jeweils auch in ihrer wechselseitigen Beeinflussung.

Tabelle 6: Soziale und kognitive Homogenisierung der Schulklassen nach der Stringenz der Leistungsdifferenzierung; (T: strikte Differenzierung; NEPS, SC 3, Welle 1; OLS; Kontrolle von Geschlecht, Migrationshintergrund, Vorschulbesuch, soziale Herkunft, kognitive Fähigkeiten, Leistungen Grundschule, Empfehlungen, Aspirationen; fett: $p \geq 0.05$)

	1		2	
	soziale Homogenisierung		kognitive Homogenisierung	
T	0.089	0.1093	0.0016	0.0013
kognitive Homogenisierung		0.4538		---
soziale Homogenisierung		---		0.0000
C	14.9	13.6	0.189	0.213
R ²	0.001	0.009	0.027	0.034
N	3005	3005	3005	3005

In Modell 1 stehen die Effekte der (strikten) Differenzierung auf die soziale, in Modell 2 auf die kognitive Homogenisierung, jeweils gemessen über die entsprechende Streuung in den Schulklassen. Danach verstärkt sich mit der Stringenz (T) die soziale Homogenität *nicht* (mit 0.0890, n.s. in Modell 1). Sie wird allerdings signifikant mit der Stringenz stärker, wenn auch der Effekt der kognitiven Homogenisierung einbezogen wird (mit 0.1093). Selbst hat die kognitive Homogenisierung keinen signifikanten Einfluss auf die Stratifikation in den Schulklassen (mit 0.4538, n.s.). Bei der kognitiven Homogenisierung (Modell 2) gibt es den strukturellen Effekt für beide Konstellationen, ohne und mit Kontrolle der sozialen Homogenisierung (mit 0.0016 bzw. 0.0013). Die soziale Homogenisierung hat dabei keinen Einfluss.

Es sieht also danach so aus, als würde eine stringenterer Regelung der Differenzierung in der Tat die angestrebten strukturellen Effekte der kognitiven Homogenisierung nach sich ziehen. Es zeigt sich aber auch, dass sich damit auch die soziale Homogenisierung verstärkt. Zunächst also scheint es so wie es die Standardposition annimmt: Die Schulklassen werden nicht nur kognitiv, sondern auch *sozial* homogener. Für die Frage hier nach den Leistungen kommt es also darauf an zu sehen, wie sich das auf die Leistungen auswirkt: Verstärkung auch des Effekts der sozialen Homogenisierung darauf? Das wird im nächsten Abschnitt behandelt.

5.3 Die Leistungen in der Sekundarstufe

Tabelle 7 beschreibt die Befunde für die Leistungen in der Sekundarstufe, links die allgemeinen Effekte, rechts die System-Effekte. Hier sind bei den Schul-Effekten nur das kognitive bzw. das soziale *Niveau* berücksichtigt, sowie der dazu gesonderte Einfluss der Schultypen bzw. der ggf. inhaltlich unterschiedlichen Curricula (über die Zugehörigkeit zum Gymnasium oder einem anderen Schultyp). Die evtl. Effekte der Homogenität und Homogenisierung der Schulklassen werden in Abschnitt 5.3 behandelt.

Tabelle 7: Leistungsdifferenzierung und Leistungen in der Sekundarstufe; (T: strikte Differenzierung; NEPS, SC 3, Welle 3; OLS; Kontrolle von Geschlecht, Migrationshintergrund und Vorschulbesuch; fett: $p \geq 0.05$)

	allgemeine Effekte					System-Effekte			
						ohne Kontrolle GYM		mit Kontrolle GYM	
soziale Herkunft Niveau soziale Herkunft	0.023	0.006 0.060	0.006 0.044	0.006 0.027	0.005 0.013	0.006 0.025	0.008 0.029	0.005 0.012	0.008 0.022
kognitive Fähigkeiten Niveau kogn. Fähigkeit			1.944	1.426 2.587	1.435 1.457	1.431 2.609	1.803 0.260	1.440 1.491	1.753 0.634
Gymnasium					0.637			0.629	0.625
T						0.065	-0.106	0.060	-0.026
T*soziale Herkunft							0.000		0.000
T*Niveau soziale Herkft.							-0.001		-0.002
T* kognitive Fähigkeit							-0.075		-0.063
T*Niveau kogn. Fähigkeit							0.465		0.420
C	-0.70	-2.87	-3.17	-3.51	-2.43	-3.75	-2.91	-2.66	-2.25
R ²	0.013	0.30	0.43	0.47	0.51	0.48	0.48	0.51	0.52
N	2717	2717	2717	2717	2717	2717	2717	2717	2717

Wieder finden sich bei den *allgemeinen Effekten* die auch vorher immer wieder gefundenen Zusammenhänge: Es gibt einen deutlichen Effekt der sozialen Herkunft, auch einen der sozialen Zusammensetzung bzw. des sozialen Niveaus der Schulklassen, aber diese Einflüsse

werden auch wieder stark vermindert, wenn die kognitiven Fähigkeiten und das kognitive Niveau der Schulklassen einbezogen werden, die selbst die stärksten eigenen Einflüsse auf die Leistungen haben. Die Zugehörigkeit zu einem Gymnasium hat einen eigenen Effekt, und die Kontrolle vermindert *beide* Schulklasseneffekte. Das bedeutet, dass das Gymnasium nicht nur curriculare Effekte transportiert, sondern auch Effekte der sozialen und kognitiven Zusammensetzung trägt, etwa der unterschiedlichen Motivation des Lehrpersonals, Referenzgruppeneffekte oder solche der besonderen Identifikation mit der jeweiligen „Anstalt“ bei Eltern und Kindern.

Für die Leistungen in der Sekundarstufe gibt es einen deutlichen *System-Effekt*: Das Leistungsniveau *erhöht* sich mit der Stringenz der Differenzierung (mit T von 0.065). Mit der Berücksichtigung der Interaktionseffekte geht dieser Effekt (als konditionaler (Haupt-)Effekt) jedoch deutlich zurück und ist nicht signifikant. Ansonsten ändert sich bis auf eine Ausnahme nichts: Weder die soziale Herkunft, noch das soziale Niveau, noch die individuellen kognitiven Fähigkeiten verändern ihren Einfluss auf die Leistungen mit der Stringenz. *Allein* der Effekt des kognitiven Niveaus allerdings wird nun *stärker* (mit 0.465). Mit der Kontrolle des Schultyps (Gymnasium) ändert sich daran nichts wesentlich. Es ist die Bestätigung des theoretischen Modells in seinem ersten Kernpunkt: Die Stringenz der Differenzierung verändert die Einflüsse der sozialen Herkunft und der sozialen Zusammensetzung der Schulklassen *nicht*, aber sie *verstärkt* den positiven Effekt des *kognitiven Niveaus* der Schulklassen – und darüber dann das Leistungsniveau insgesamt.

In Tabelle 8 ist das im System-Vergleich für die detaillierteren Unterschiede noch einmal zusammenfassend beschrieben, wieder einmal ohne und einmal mit Kontrolle des Gymnasiums als Lernumgebung.

Tabelle 8: Die Leistungen in der Sekundarstufe im System-Vergleich; (T: strikte Differenzierung; NEPS, SC 3, Welle 3; OLS; Kontrolle von Geschlecht, Migrationshintergrund, Vorschulbesuch; fett: $p \geq 0.05$)

	Differenzierung					
	ohne Kontrolle GYM			mit Kontrolle GYM		
	liberal		strikt	liberal		strikt
soziale Herkunft	0.008	0.005	0.005	0.008	0.005	0.004
Niveau soziale Herkunft	0.023	0.026	0.025	0.009	0.019	0.006
kognitive Fähigkeiten	2.031	1.219	1.411	1.924	1.253	1.439
Niveau kognitive Fähigkeiten	1.263	2.592	3.214	0.622	1.599	1.853
C	-2.84	-3.59	-3.56	-2.18	-2.65	-2.17
R ²	0.45	0.48	0.47	0.50	0.50	0.53
N	525	1181	1101	525	1181	1111

In beiden Vergleichen *sinkt* der Effekt der sozialen Herkunft mit der Stringenz der Differenzierung, und eine sonderliche Regelmäßigkeit für die Effekte des sozialen Niveaus in den Schulklassen gibt es auch nicht. Es zeigt sich jedoch, dass der Effekt des kognitiven Niveaus in den Schulklassen deutlich *zunimmt*. Bemerkenswert sind dabei die Befunde zu den liberalen Regimes: Das kognitive Niveau der *Schulklassen* hat dort offenbar *keinen* signifikanten Einfluss auf die Leistungen (mit 0.1263 bzw. 0.622, jeweils n.s.), hingegen ist der Einfluss der *individuellen* kognitiven Fähigkeiten im Vergleich dort am stärksten.

Womöglich hängt das zusammen: In den kognitiv weniger förderlichen schulischen Umgebungen der liberalen Regimes setzen sich die talentierten Kinder, wie allgemein ohnehin, immer noch durch, in den strikteren bekommen jedoch durch die dort günstigeren Lernbedingungen auch die weniger talentierten Kinder ihre Chancen auf gute Leistungen, so dass hier das individuelle Talent nicht mehr so bedeutsam sein muss, um ein gewisses Niveau zu erreichen. Anders gesagt: Bei stringenter Differenzierung gibt es *mehr* an „resilienten“ Kindern.

5.4 Homogenität und Homogenisierung

Eine striktere Differenzierung führt dem theoretischen Modell folgend zu einem höheren Leistungsniveau, weil sich die Effekte des kognitiven Niveaus in den Schulklassen – *ceteris*

paribus – damit verstärken. Das sollte Modell an der dann bei gleichem Niveau stärkeren *Homogenisierung* der Schulklassen liegen, die eine stärkere Fokussierung des jeweiligen Unterrichts ermöglicht (vgl. Abschnitt 1 oben dazu). Tabelle 9 gibt die Befunde für eine Analyse wieder, in der die vier zentralen Bedingungen zur Erklärung der Leistungen (kognitive Fähigkeiten, kognitives Niveau, Stringenz der Differenzierung und Heterogenität bzw. Homogenität in der Verteilung der kognitiven Fähigkeiten in den Schulklassen) als dummy-Variablen kombiniert sind.⁵ Referenzkategorie der dann 16 Kombinationen ist nach dem theoretischen Modell ungünstigste Fall (Kombination 1): geringe kognitiven Fähigkeiten, niedriges kognitives Niveau, offenes Regime und kognitive Heterogenität in den Schulklassen (ABL-/NAB-/T-/HET). Zu vergleichen sind die Effekte für jede der vier Kombinationen nach den kognitiven Fähigkeiten und dem kognitiven Niveau der Schulen und Schulklassen.

⁵ Analysen aller Variablen mit den entsprechenden Interaktionseffekten wurden nicht vorgenommen, weil die dazu nötige statistische Genauigkeit bei den gegebenen Fallzahlen nicht mehr gegeben ist. Die vier betrachteten Bedingungen – kognitive Fähigkeiten, kognitives Niveau und kognitive Homogenität der Schulklassen sowie die Stringenz der Differenzierung wurden jeweils am Median dichotomisiert.

Tabelle 9: Effekte der kognitiven Homogenisierung auf die Leistungen in der Sekundarstufe (NEPS, SC 3, Welle 3; dummy-Variablen für die Kombination von kognitiven Fähigkeiten (ABL), Niveau kognitive Fähigkeiten Schulklasse (NABL), Stringenz Differenzierung (T) und Heterogenität/Homogenität kognitive Fähigkeiten Schulklassen (HET/HOM); OLS; Kontrolle Geschlecht, Migrationshintergrund, Vorschulbesuch; fett: stärkster Effekt; grau hervorgehoben: siehe Text)

				Kombination		plus Kontrolle SES, SSES	plus Kontrolle GYM	n
ABL-	NABL-	HET	T-	1	---	---	---	353
		HET	T+	2	-0.067	-0.126	-0.099	234
		HOM	T-	3	-0.150	-0.137	-0.118	217
		HOM	T+	4	0.332	0.313	0.237	114
	NABL+	HET	T-	5	0.627	0.155	-0.118	97
		HET	T+	6	0.951	0.543	0.242	60
		HOM	T-	7	0.899	0.380	0.027	37
		HOM	T+	8	1.224	0.622	0.309	145
ABL+	NABL-	HET	T-	9	0.729	0.608	0.561	182
		HET	T+	10	0.933	0.622	0.675	45
		HOM	T-	11	0.530	0.395	0.419	62
		HOM	T+	12	1.176	1.030	0.695	12
	NABL+	HET	T-	13	1.201	0.711	0.416	185
		HET	T+	14	1.605	1.106	0.769	90
		HOM	T-	15	1.398	0.845	0.518	329
		HOM	T+	16	1.599	1.039	0.718	324
C					-0.11	-2.14	-3.02	
R ²					0.33	0.42	0.48	
N					2717	2717	2717	

Das Ergebnis lässt sich leicht zusammenfassen: In allen vier Konstellationen von kognitiven Fähigkeiten und kognitivem Niveau und bei allen Kontrollen erzeugt die Kombination von strikter Differenzierung und kognitiver Homogenität der Schulklassen den jeweils *stärksten* positiven Effekt (Kombinationen 4, 8, 13 und 16; mittelgrau unterlegt). Es gibt eine Ausnahme: Für die strukturell günstigste Konstellation der passenden Sortierung talentierter Kinder in Schulklassen mit dem höheren Niveau (ABL+,NAB+) ist es bei strikter Differenzierung nicht nur gleichgültig, ob die Zusammensetzung heterogen oder homogen ist, ihnen schient die Heterogenität sogar einen gewissen Vorteil zu bieten, der mit der Kontrolle von sozialer Herkunft und Schultyp sogar stärker hervortritt (Kombination 14; hellgrau unterlegt). Anders gesagt: Die *Heterogenität* ist eher etwas für die *besseren* Kinder in den

besseren Schulen, also da, wo es keine Probleme gibt, die Homogenität für alle anderen, gerade auch für die unter den ungünstigeren Bedingungen. Das aber *nur* unter der Bedingung, dass die Sortierung auch wirklich *strikt* nach den Fähigkeiten bzw. den Leistungen vorher erfolgt ist.

5.5 Trade Off?

Eine der von Beginn an besonders umstrittenen Fragen für die Effekte von kognitiver Homogenität und Homogenisierung ist die, ob sich die evtl. Gewinne unterschiedlich verteilen und dabei, wenn überhaupt, eher die leistungsstärkeren Kinder etwas davon haben. In Tabelle 10 stehen die Ergebnisse eines Vergleichs der Effekte für drei unterschiedliche Leistungsgruppen: Im unteren Quantil von 10% der Verteilung, im mittleren von 50% und im hohen von 90% der Verteilung. Der Vergleich bezieht sich auf den Effekt der jeweils für die Leistungen günstigsten Kombination von strikter Differenzierung und Homogenität bei den vier Konstellationen von kognitiven Fähigkeiten und kognitivem Niveau der Schulklassen, also die mittelgrau unterlegten Einträge 4, 8, 12 und 16 in Tabelle 8 oben). In Tabelle 9 sind die betreffenden Werte aus Tabelle 8 zum Vergleich noch einmal eingetragen und grau unterlegt.

Tabelle 10: Effekte der kognitiven Homogenisierung für unterschiedliche Leistungsstufen (NEPS, SC 3, Welle 3; Quantils-Regression; Kontrolle von Geschlecht, Migrationshintergrund, Vorschulbesuch; fett: stärkster Effekt im Vergleich; kursiv: schwächster Effekt im Vergleich)

					+Kontrolle SES, SSES				+ Kontrolle GYM				
		alle	Quantile			alle	Quantile			alle	Quantile		
		HOM T+	0.1	0.5	0.9	HOM T+	0.1	0.5	0.9	HOM T+	0.1	0.5	0.9
ABL-	NABL-	0.332	0.325	0.384	<i>0.191</i>	0.313	0.355	<i>0.268</i>	0.408	0.237	0.337	0.271	<i>0.237</i>
	NABL+	1.224	1.498	1.214	<i>1.088</i>	0.622	0.774	0.688	<i>0.564</i>	0.309	0.348	<i>0.302</i>	0.369
ABL+	NABL-	1.176	1.430	1.324	<i>0.665</i>	0.711	1.032	1.089	<i>0.719</i>	0.416	0.981	<i>0.656</i>	0.739
	NABL+	1.599	1.925	1.619	<i>1.282</i>	1.039	1.176	1.100	<i>0.832</i>	0.718	0.844	0.730	<i>0.615</i>
C		-0.11				-2.14				-3.02			
R ²		0.33				0.42				0.48			
N		2717											

Für die Analysen ohne weitere Kontrollen zeigt sich, dass in allen vier Konstellationen die leistungsstarken Kinder am wenigsten von der Kombination strikter Differenzierung und Homogenität profitieren (Werte kursiv gekennzeichnet), im Gegensatz zu den leistungsschwachen Kindern, die in drei Fällen den z.T. deutlich höheren Zuwachs haben. Die einzige Ausnahme sind die leistungsschwachen Kinder in der ungünstigen Kombination geringer kognitiver Fähigkeiten und niedrigem kognitiven Niveau der Schulklasse. Aber hier ist der Abstand zu der besten Gruppe, den Kindern mit dem mittleren Leistungsniveau, nicht groß (0.325 zu 0.384), aber der zu den leistungsstarken erheblich (0.325 zu 0.191). In keiner Konstellation gibt es bei den leistungsschwachen Kindern den geringsten Effekt.

Der Befund, dass die leistungsschwachen Kinder von Homogenität und Stringenz der Differenzierung *mehr* profitieren, wiederholt sich mit der Kontrolle der sozialen Herkunft und des sozialen Niveaus der Schulklassen bzw. dann auch mit dem Schultyp (Gymnasium) im Wesentlichen. Allerdings werden die Effekte und die Unterschiede wie üblich nun geringer, weil die mit der Herkunft und dem sozialen Niveau der Schulen und Schulklassen gegebenen Qualitätsunterschiede eine Rolle spielen, und auch etwas inkonsistenter, sowie, kontrolliert mit der Zugehörigkeit zu einem Gymnasium, die unterschiedlichen Curricula. Bemerkenswert ist, dass auch nach der sozialen Herkunft und der sozialen Zusammensetzung der Schulklassen und der Gymnasialzugehörigkeit weiter gerade die leistungsschwachen Kinder eher Vorteile von kognitiver Homogenität und Homogenisierung haben oder wenigstens nicht den geringsten Zuwachs haben. Ein Trade Off im dem Sinne, dass die Gewinne der Leistungsdifferenzierung klar zulasten der Leistungsschwachen gehen, findet sich jedenfalls nicht. Eher sogar das Gegenteil: Die strikte Leistungsdifferenzierung und die damit erfolgende kognitive Homogenisierung der Schulklassen *kompensieren* tendenziell die Ungleichheiten zwischen den leistungsschwächeren und den leistungsstärkeren Schülerinnen und Schülern sogar.

6. Zusammenfassung und Bewertung

Der Ausgangspunkt des Beitrags war die Frage, ob die nach vor wie nahezu als fraglos gültig angenommene Standardposition zu den Effekten der Leistungsdifferenzierung zutrifft, wonach das strikte Ability Tracking die Leistungen nicht verbessere, sondern eigentlich nur die Bildungsungleichheiten nach der sozialen Herkunft verstärke, also eine Art von

meritokratisch verkleidetem *Nobility Tracking* wäre. Das Ergebnis lässt sich in sechs Punkten zusammenzufassen: *Erstens*. Mit der strikten Differenzierung nach Leistung steigt, jeweils im Vergleich zu den offeneren Systemen, die Leistungsgerechtigkeit beim Übergang, ohne dass sich andere Effekte verändern, insbesondere nicht solche der sozialen Herkunft. *Zweitens*. Die kognitive Homogenität der Schulklassen nimmt mit der strikten Differenzierung in der Tat zu. *Drittens*. Mit der strikten Differenzierung steigt das Leistungsniveau in der Sekundarstufe insgesamt. *Viertens*. Das geschieht über die Verstärkung des Effekts des kognitiven Niveaus in den Schulklassen als Folge der strikten Differenzierung. *Fünftens*. Diese Verstärkung ist die Folge des leistungssteigernden Effekts der bei strikter Differenzierung stärkeren kognitiven Homogenisierung der Schulklassen. *Sechstens*. Von diesem Effekt profitieren vor allem die leistungsschwächeren Kinder.

Die Befunde entsprechen so gut wie ausnahmslos den Hypothesen des theoretischen Modells. Sie widersprechen damit in ihren Kernbestandteilen der Standardposition. Sie reißen sich in die inzwischen deutlicher werdenden Hinweise ein, dass die bisherigen Befunde, die vornehmlich die Standardposition stützen, insbesondere die aus den internationalen Vergleichsstudien wie PISA vor allem, auf Annahmen, Daten und Analysen beruhen, die zu systematischen Fehlschlüssen geradezu führen mussten (vgl. auch die Übersicht dazu allgemein bei Marks 2014 und für den speziellen Fall von Finnland: Sahlgren 2015). Es waren insbesondere drei meist übersehene Gegebenheiten und Probleme. *Erstens*. Zu den entscheidenden Mechanismen der Erklärung von schulischen Leistungen waren die erforderlichen Daten nicht gegeben, wie insbesondere die kognitiven Fähigkeiten, sowohl als individuelle Merkmale wie als Bestandteile der schulischen Lernumgebungen nach Niveau und Homogenität. *Zweitens*. Damit war es schon von der empirischen Basis her nicht möglich, den Vorgang der Sortierung nach den kognitiven Fähigkeiten am Ende der Grundschule zu kontrollieren, so dass es zu Fehlspezifikationen für die Identifikation der spezifischen System-Effekte der Differenzierung kommen musste. *Drittens*. Dazu gehörte insbesondere die Fehlzuschreibung von Effekten der kognitiven Fähigkeiten auf solche der sozialen Herkunft durch einen schon schwerwiegenden *omission variable bias*: Die zu begrüßende *Verstärkung* der *positiven* Effekte der *kognitiven Fähigkeiten* bei strikter Differenzierung wurde als eine zu kritisierende *Verschärfung* der Effekte der *sozialen Herkunft* fehlgedeutet, so dass der Eindruck in der Tat entstehen musste, dass die Differenzierung nach den Fähigkeiten und Leistungen nichts weiter wäre als eine ideologisch verdeckte Form der Stratifikation.

Ein Hintergrund dieser inzwischen schon langen Geschichte reichlich verwirrender Debatten um die Differenzierung war ohne Zweifel, dass die genaueren theoretischen Mechanismen – bis auf die frühen Arbeiten bei Sörensen und Hallinan im Anschluss an Coleman – kaum einmal explizit systematisiert wurden. Es war vielmehr, wie so oft auch anderswo, ein langes, eher induktives Stochern im Nebel unzureichender Daten und ad-hoc fabrizierter „Theorien“, und es stellte sich erst sehr langsam heraus, was man eigentlich benötigen würde. Hinzu legte – ohne Zweifel – eine kaum einmal in Frage gestellte – Voreinstellung die Akzeptanz der Standardposition nahe, wonach Meritokratie und Leistungsbezug nur Schwestern der Stratifikation seien. Und weil es die nötigen Daten lange nicht gab, konnte sich das, auch gegen eine Vielzahl von Unklarheiten und Gegenevidenzen, lange halten, eigentlich bis heute. Erst in den Daten der NEPS-Studie sind alle die erforderlichen Informationen enthalten gewesen, was den doch erstaunlich deutlichen Widerspruch der berichteten Befunde zur Standardposition erklären mag.

Weil die Befunde sich einstweilen nur auf die deutschen Bundesländer beziehen, bilden sie sicher noch keinen verallgemeinerbaren Beleg für die Effekte von kognitiver Homogenität und Homogenisierung durch eine strikte Differenzierung, gerade auch nicht im internationalen Vergleich. Aber wir haben nichts Anderes – bislang. Die bildungspolitischen Schlussfolgerungen für das deutsche Bildungssystem liegen, wie schon für Bildungsbeteiligung (vgl. Esser und Hoenig 2018), allerdings auch so schon auf der Hand: Nicht die weitere Öffnung und Abkehr von den Kriterien der Leistungsdifferenzierung hilft die Effizienz des Bildungssystems zu steigern und die sozialen (und ethnischen) Bildungsungleichheiten zu dämpfen und sogar zu senken, sondern eine (wieder) stärkere Orientierung daran. Der Grund ist einfach genug: Wenn zuerst die Fähigkeiten, Begabungen, Talente und Leistungen zählen, dann treten die Einflüsse anderer Merkmale zurück wie die der sozialen und ethnischen Herkunft, und gerade die privilegierten Familien können ihre Vorteile dann nicht weiter wie bisher offenbar so ausnutzen wie sie das mit ihren Aspirationen wollen. Ob und wie das – „sozialverträglich“ – umzusetzen wäre, ist natürlich eine ganz andere Frage.

Adoso, S.O., und J.O. Agbayewa. 2011. Effect of homogenous and heterogeneous ability grouping class teaching on student's interest, attitude and achievement in integrated science, in: *International Journal of Psychology and Counselling*, 3: 48-54.

Allmendinger, J. 2012. *Schulaufgaben. Wie wir das Bildungssystem verändern müssen, um unseren Kindern gerecht zu werden.* München: Pantheon.

Ammermüller, A. 2005. Educational opportunities and the role of institutions. ZEW Discussion Papers No. 05-44.

Baumert, J., P. Stanat, und R. Watermann. 2006. Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In: J. Baumert, P. Stanat und R. Watermann, Hrsg., *Herkunftsbedingte Disparitäten im Bildungswesen. Vertiefende Analyse im Rahmen von PISA 2000*, 95-188. Wiesbaden: VS Verlag für Sozialwissenschaften.

Becker M., O. Lüdtke, U. Trautwein, O. Köller und J. Baumert. 2012. The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? In: *Journal of Educational Psychology*, 104, 682–699.

Betts, J.R. 2011. The economics of tracking in education. In: E.A. Hanushek, S. Machin und L. Woessmann, Hrsg., *Handbook of the Economics of Education*, Band 3, 341-381. Amsterdam: North Holland.

Betts, J.R., und J.L. Shkolnik. 2000. Key difficulties in identifying the effects of ability grouping on student Achievement. In: *Economic Education Review*, 19, 21–26.

Blossfeld, H.P., S. Buchholz, J. Skopek und M. Triventi. 2016, Hrsg.. *Models of Secondary Education and Social Inequality – An International Comparison.* eduLIFE Lifelong Learning Series, Band 3. Cheltenham, UK und Northampton, MA, USA: Edward Elgar Publishing.

Blossfeld, H.P., H.G. Roßbach und J. von Maurice. 2011. *Education as a Lifelong Process – The German National Educational Panel Study (NEPS).* Wiesbaden: Springer VS.

Böhme, K., und S. Weirich. 2010. Der Ländervergleich im Fach Deutsch. In: P. Stanat, H. A. Pant, K. Böhme und D. Richter, Hrsg., *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs*, 103-116. Münster: Waxmann.

Bol, Th., J. Witschge, H.G. van de Werfhorst, und J. Dronkers. 2014. Curricular tracking and central examinations: Counterbalancing the impact of social background on student achievement in 36 countries. In: *Social Forces* 92, 1545-1572.

Diehl, C., Ch. Hunkler und C. Kristen, (Hrsg). 2016. *Ethnische Ungleichheiten im Bildungsverlauf. Mechanismen, Befunde, Debatten.* Wiesbaden: Springer VS.

Dronkers, J., und R. Korthals. 2015. Tracking, school entrance requirements and the educational performance of migrant students. In: A. Hadjar und Ch. Gross, Hrsg., *Education systems and inequalities. International comparisons*, 185-205. Bristol und Chicago: Policy Press.

Dronkers, J., R. van der Velden und A. Dunne. 2011. The effects of educational systems, school-composition, track-level, parental background and immigrants' origins on the achievement of 15-years old native and immigrant students. A reanalysis of PISA 2006. Maastricht: ROA, Research Centre for Education and the Labour Market.

Dronkers, J., R. van der Velden, und A. Dunne. 2012. Why are migrant students better off in certain types of educational systems or schools than in others? In: *European Educational Research Journal* 11, 11-44.

Duflo, E., P. Dupas, und M. Kremer. 2011. Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. In: *American Economic Review* 101, 1739-1774.

Dunne, A., 2010. Dividing lines. Examining the relative importance of between- and within-school differentiation during lower secondary education. PhD-Thesis, Department of Political and Social Sciences. Florence: European University Institute.

Esser, H. 2013. Ethnische Bildungsungleichheit und Bildungssysteme: Der Blick in die Bundesländer. In: N. McElvany, M.M. Gebauer, W. Bos und H.G. Holtappels, Hrsg., *Jahrbuch der Schulentwicklung*, Band 17, 80-101. Weinheim und Basel: Beltz Juventa.

Esser, H. 2016a. Bildungssysteme und ethnische Bildungsungleichheit. In: Diehl, C., Ch. Hunkler & C. Kristen, Hrsg., *Ethnische Ungleichheiten im Bildungsverlauf. Mechanismen, Befunde, Debatten*, 331-396. Wiesbaden: Springer VS.

Esser, H. 2016b. Educational Systems and Educational Inequality. The Model Ability Tracking and Empirical Findings. In: H.P. Blossfeld, S. Buchholz, J. Skopek, und M. Triventi, Hrsg., *Models of Secondary Education and Social Inequality – An International Comparison*. eduLIFE Lifelong Learning Series, Band 3, 25-44. Cheltenham, UK und Northampton, MA, USA: Edward Elgar Publishing.

Esser, H., und K. Hoenig. 2018. Leistungsgerechtigkeit und Bildungsungleichheit. Effekte der Verbindlichkeit der Grundschulempfehlungen beim Übergang auf das Gymnasium. Ein Vergleich der deutschen Bundesländer mit den Daten der „National Educational Panel Study“ (NEPS). 2018. In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 70, 419-447.

Esser, H., und I. Relikowski. 2015. Is Ability Tracking (Really) Responsible for Educational Inequalities in Achievement? A Comparison between the Country States Bavaria and Hesse in Germany. IZA-Discussion Paper Nr. 9082. Bonn: Forschungsinstitut Zukunft der Arbeit. Institute for the Study of Labor.

Felouzis, F., und S. Charmillon. 2013. School tracking and educational inequality: A comparison oft 12 educational systems in Switzerland. In: *Comparative Education* 49, 181-205.

Felouzis, F., und S. Charmillon. 2017. Schulische Ungleichheit in der Schweiz. Genfer Forschungsgruppe zur Bildungspolitik (Groupe Genevois d'Analyse des Politiques Éducatives – GGAPE). Psychologische und Erziehungswissenschaftliche Fakultät Universität Genf. *Social Change in Switzerland*. Nr . 8. Lausanne.

Galindo-Rueda, F., und A. Vignoles. 2007. The heterogeneous effect of selection in UK secondary schools. In: L. Wößmann und P.E. Peterson, Hrsg., *Schools and the equal opportunity problem*, 103-128. Massachusetts: MIT Press.

Gamoran, A. 2009. *Tracking and Inequality: New Directions for Research and Practice*. WCER Working Paper Nr. 2009-6. Madison: University of Wisconsin-Madison: Wisconsin Centre for Education Research

Garmoran, A. und R.D. Mare. 1989. Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality? In: *American Journal of Sociology*, 94, 1146-1183.

Ganzeboom, H. B., P. M. de Graaf, & D. J. Treiman. 1992. A standard international socio-economic index of occupational status. In: *Social Science Research* 21: 1-56.

Hadjar, A., und R. Becker. 2016. Education systems and meritocracy: social origin, educational and status attainment. In: A. Hadjar und Ch. Gross, Hrsg., *Education systems and inequalities. International comparisons*, 231-258. Bristol und Chicago: Policy Press.

Hallinan, M.T. 1994. Tracking: From Theory to Practice. In: *Sociology of Education*, 67, 79-84.

Hanushek, E.A., und L. Wößmann. 2006. Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. In: *The Economic Journal*, 116, C63-C76.

Hanushek, E.A., und L. Wößmann. 2011. The economics of international differences in educational achievement. In: E.A. Hanushek, S. Machin und L. Wößmann, Hrsg., *Handbook of the Economics of Education*, Band. 3, 89-200. Amsterdam: North-Holland.

Helbig, M., und R. Nikolai. 2015. *Die Unvergleichbaren. Der Wandel der Schulsysteme in den deutschen Bundesländern seit 1949*. Bad Heilbrunn: Klinkhardt.

Holtappels, H.G. 2017. *Entwicklung und Qualität des Schulsystems. Neue empirische Befunde und Entwicklungstendenzen*. Münster und New York: Waxmann.

Horn, D. 2008. Age of selection counts: a cross-country comparison of educational institutions. Mannheim: Arbeitspapiere - Mannheimer Zentrum für Europäische Sozialforschung; 107.

Horn, D. 2013. Diverging performances: the detrimental effects of early educational selection on equality of opportunity in Hungary, *Research in Social Stratification and Mobility*, 32, 25-43.

Huang, M.H. 2009. Classroom homogeneity and the distribution of student math performance: A country-level fixed-effects analysis. In: *Social Science Research*, 38, 781-791.

Jackson, M., und J.O. Jonsson. 2013. Why does inequality of educational opportunity vary across countries? Primary and secondary effects in comparative context. In: M. Jackson, Hrsg., *Determined to succeed? Performance versus choice in educational attainment*, Stanford, CA: Stanford University Press, 306-337.

Kerr, S.K., T. Pekkarinen und R. Uusitalo. 2013. School tracking and development of cognitive skills. In: *Journal of Labor Economics*, 31, 577 – 602.

Lang, F. R., S. Kamin, M. Rohr, C. Stünkel und B. Williger. 2014. Erfassung der fluiden kognitiven Leistungsfähigkeit über die Lebensspanne im Rahmen des Nationalen Bildungspanels: Abschlussbericht zu einer NEPS-Ergänzungsstudie (NEPS Working Paper No. 43). Bamberg: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.

Linchewski, L., und B. Kutscher. 1996. Tell Me With Ehom You´re Learning. And I´ll Tell You How Much You´ve Learned: Mixed-Ability versus Same-Ability Grouping in Mathematics. In: *Journal for Research in Mathematics Education* 29, 533-554.

Manning, A., und J.-St. Pischke. 2006. Comprehensive Versus Selective Schooling in England & Wales: What Do We Know? Centre for the Economics of Education London School of Economics.

Marks, G.N. 2014. Education, social background and cognitive ability. The decline of the social. London: Routledge.

Oakes, J. 2005, *How Schools Structure Inequality*, 2. Aufl., New Haven und London: Yale University Press (zuerst 1985).

Pfeffer, F, 2008. Persistent Inequality in Educational Attainment and its Institutional Context. In: *European Sociological Review*, 24, 543-565.

Piopiunik, M. 2014. The effects of early tracking on student performance: Evidence from a school reform in Bavaria. In: *Economics of Education Review* 42, 12-33.

Sacerdote, B. 2011. Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In: E.A. Hanushek, S. Machin und L. Woessmann, Hrsg., *Handbook of the Economics of Education*, Band 3, 250-277. Amsterdam: North Holland.

Sahlgren, G.H. 2015. *Real Finnish Lessons. The true story of an educational superpower.* Centre for Policy Studies: London.

Schallock, T. 2016. Tracking. Konsequenzen der Leistungsgruppierung von Schülerinnen und Schülern für die Leistungsentwicklung. In: C. Diehl, C. Hunkler und C. Kristen, Hrsg., *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten*, 287-330. Wiesbaden: Springer VS.

Schipolowski, S. und K. Böhme. 2010. Die sprachlichen Kompetenzen in den Ländern. In: O. Köller, M. Knigge und B. Tesch, Hrsg., *Sprachliche Kompetenzen im Ländervergleich*, 87-96. Münster: Waxmann.

Sörensen, A.B. 1970. Organizational Differentiation of Students and Educational Opportunity? *Sociology of Education*, 43, 355-376.

Sörensen, A.B., und M. T. Hallinan. 1977. A Reconceptualization of School Effects. In: *Sociology of Education*, 50, 273-289.

Solga, H., und R. Becker. 2012. Einleitung: Soziologische Bildungsforschung – Eine kritische Bestandsaufnahme. In: R. Becker und H. Solga, Hrsg., *Soziologische Bildungsforschung*. Sonderheft 52 der Kölner Zeitschrift für Soziologie und Sozialpsychologie, 7-43. Wiesbaden: Springer.

Stadelmann-Steffen, I. 2012. Education Policy and Educational Inequality – Evidence from the Swiss Laboratory. In: *European Sociological Review* 28: 379-393.

van de Werfhorst, H.G., und J.J.B. Mijs. 2010. Achievement inequality and the institutional structure of educational systems: A comparative perspective. In: *Annual Review of Sociology*, 36: 407-428.

von Below, S. 2011. Bildungssysteme im historischen und internationalen Vergleich. In: R. Becker, Hrsg., *Lehrbuch der Bildungssoziologie*, 2. Aufl., 139-162. Wiesbaden: Springer VS.

Waldinger, F. 2007. Does tracking affect the importance of family background on students' test scores? Unpublished manuscript. London: LSE.

Wößmann, L. 2010. Institutional Determinants of School Efficiency and Equity: German States as Microcosm for OECD Countries. In: *Jahrbücher für Nationalökonomie und Statistik*, 230: 234-270.

Wößmann, L., E. Lüdemann, G. Schütz, und M.R. West. 2009. *School accountability, autonomy and choice around the world*. Celtenham: Edward Elgar Publishing.

Zimmer, R. 2003. A new twist in Theo educational tracking debate. In: *Economics of Education Review* 22, 307-315.