

Frank Lipowsky, Tobias C. Stubbe, Gabriele Faust[†],
Josef Künsting, Swantje Hadelers & Wilfried Bos

Was leisten Schülerinnen und Schüler der privaten BIP-Kreativitätsgrundschulen im nationalen Vergleich?

Zusammenfassung

Der vorliegende Beitrag untersucht die Leistungen von Schülerinnen und Schülern der privaten BIP-Kreativitätsgrundschulen am Ende des 4. Schuljahres in den Domänen Mathematik, Naturwissenschaften und im Lesen. Bei den privaten BIP-Kreativitätsgrundschulen handelt es sich um Grundschulen vorwiegend im Bundesland Sachsen und in anderen ostdeutschen Bundesländern, welche ihrem Konzept nach in besonderer Weise zur Förderung der Begabung, der Intelligenz und der Kreativität beitragen wollen.

Um die Leistungen der BIP-Schülerinnen und Schüler national einordnen zu können, wurden diese mit IGLU-Instrumenten erfasst und untersucht. Ein Vergleich mit der IGLU-Stichprobe 2001 deckt auf, dass die BIP-Schülerinnen und Schüler erhebliche Leistungsvorsprünge im Lesen, in Mathematik und in den Naturwissenschaften aufweisen. Da es sich bei den BIP-Schülerinnen und Schülern um eine stark sozial selektierte Stichprobe handelt, wird ein Propensity-Score-Matching Verfahren durchgeführt, um jedem Schüler bzw. jeder Schülerin aus den BIP-Schulen einen Schüler bzw. eine Schülerin mit vergleichbarem so-

Prof. Dr. Frank Lipowsky (corresponding author), Fachgebiet Empirische Schul- und Unterrichtsforschung, Universität Kassel, Nora-Platiel-Str. 1, 34109 Kassel, Deutschland
E-Mail: lipowsky@uni-kassel.de

Prof. Dr. Tobias C. Stubbe, Lehrstuhl Schulpädagogik/Empirische Schulforschung, Georg-August-Universität Göttingen, Waldweg 26, 37073 Göttingen, Deutschland
E-Mail: tstubbe@uni-goettingen.de

Prof. Dr. Josef Künsting, Fakultät für Psychologie, Pädagogik und Sportwissenschaft, Universität Regensburg, Universitätsstraße 31, 93053 Regensburg, Deutschland
E-Mail: josef.kuensting@ur.de

Dr. des. Swantje Hadelers, Sprach- und Literaturdidaktik im Elementar- und Primarbereich, Universität Bremen, Bibliotheksstraße 1–3, 28359 Bremen, Deutschland
E-Mail: s.hadelers@uni-bremen.de

Prof. Dr. Wilfried Bos, Institut für Schulentwicklungsforschung, Technische Universität Dortmund, Vogelpothsweg 78, 44227 Dortmund, Deutschland
E-Mail: officebos@ifs.tu-dortmund.de

zialen Hintergrund aus der IGLU-Stichprobe zuzuordnen. Ein Vergleich der gematchten Stichproben ergibt deutlich geringere Leistungsvorteile der BIP-Schülerinnen und Schüler als ein Vergleich der ungematchten Stichproben. Sofern man auch die Klassenkomposition kontrolliert, ergibt sich nur noch für das Fach Mathematik ein Leistungsvorsprung der BIP-Schülerinnen und Schüler.

Schlagworte

BIP-Kreativitätsgrundschulen; Privatschule; Grundschule; Leistungsvergleich; Propensity Score Matching

What academic performance do students of the private BIP creativity schools achieve compared to a nationwide sample of students in public schools?

Abstract

This paper examines the performance of students attending the private BIP creativity schools at the end of Grade 4 in the domains mathematics, science, and reading. The private BIP creativity schools are elementary schools with a special concept that seeks to foster the giftedness, the intelligence, and the personality of students. BIP schools are predominantly located in the German federal state of Saxonia and in other eastern states of Germany.

In order to make a comparison of our findings on a national level possible, the performance of the BIP students was assessed and examined using instruments from the IGLU study. A comparison with the IGLU sample of 2001 reveals that BIP students show considerably higher performance levels in reading, mathematics, and science. Since the sample of BIP students is highly selective, a Propensity Score Matching was applied to match each BIP student with a student from the IGLU sample who has a similar social background. A comparison of the matched samples shows that the advantages of the BIP students decrease clearly in comparison to the unmatched samples. When the composition of the classes is additionally controlled, the advantage of the BIP students is limited only to mathematics.

Keywords

BIP creativity schools; Private school; Elementary school; Comparison of performance; Propensity Score Matching

1. Einleitung

Das Interesse von Eltern, ihre Kinder auf Privatschulen zu schicken, ist ungebrochen und hat in den letzten Jahren sogar noch weiter zugenommen. So ist die Anzahl allgemeinbildender Privatschulen im Zeitraum von 1992/1993 bis

2011/2012 von 1991 auf 3 396 um 70.6 % kontinuierlich gestiegen. Ihr Anteil an allen allgemeinbildenden Schulen betrug im Schuljahr 2011/2012 9.8 %, während er im Schuljahr 1992/1993 noch bei 4.5 % lag. Die Anzahl der Schüler¹ an allgemeinbildenden privaten Schulen nahm im gleichen Zeitraum von 445 609 auf 725 898 zu, was einem Zuwachs von 62.9 % entspricht. Jeder 12. Schüler an allgemeinbildenden Schulen besuchte 2011/2012 in Deutschland damit eine Privatschule (8.4 %), während es im Jahr 1992/1993 nur jeder 21. Schüler war (4.8 %) (vgl. Statistisches Bundesamt, 2012). Zu einem besonders starken Zuwachs kam es im Primarbereich. Zwischen 1995/1996 und 2010/2011 stieg die Zahl der Privatschüler an allgemeinbildenden Grundschulen um 83.2 %, während die Anstiege im Sekundarbereich I und im Sekundarbereich II mit 37 % und 59.8 % deutlich geringer ausfielen (vgl. Weishaupt et al., 2012, S. 235).

Innerhalb des Bildungswesens erfüllen private Schulen wichtige Funktionen: Sie bieten alternative pädagogische Angebote und/oder versprechen eine relativ homogene Schulkultur, die einer besonderen Idee oder Weltanschauung verpflichtet ist (Leschinsky, 2003). Allerdings fehlt es trotz der wachsenden Resonanz, die Privatschulen im Primarbereich erfahren, an Studien, die die Leistungen von Grundschulern an Privatschulen systematisch untersuchen.² Hier setzt der vorliegende Beitrag an.

Im Unterschied zu vielen Studien, die die Leistungen von Schülern an öffentlichen Schulen mit denen von Schülern an privaten Schulen unterschiedlichster Konzeption vergleichen (vgl. z. B. Dronkers & Avram, 2009; Lubienski, Lubienski & Crane, 2008), liegt den in diesem Beitrag fokussierten privaten BIP-Kreativitätsgrundschulen ein relativ homogenes Schulkonzept zu Grunde. So wird eine Konfundierung unterschiedlicher Schulkonzepte innerhalb der Vergleichsstichprobe privater BIP-Kreativitätsgrundschulen weitgehend vermieden. Damit bestehen günstigere Voraussetzungen für eine Evaluation der Schulkonzeption eines privaten Schulträgers als in bisherigen Studien.

2. Die BIP-Kreativitätsgrundschulen

Die privaten *BIP*-Kreativitätsschulen wurden von Hans-Georg und Gerlinde Mehlhorn eingerichtet und aufgebaut (Mehlhorn & Mehlhorn, 2003). Sie sind bis auf wenige Tage im Jahr ganzjährig und ganztägig geöffnete Einrichtungen mit ei-

1 Wenn in diesem Beitrag von Schülern oder Lehrern die Rede ist, sind jeweils immer Vertreter beiderlei Geschlechts gemeint.

2 In einer der wenigen Studien, die Privatschuleffekte im Grundschulalter untersuchte, verglichen Heise und Kollegen (2010) die Entwicklung von Grundschulkindern in einer Montessorigrundschule mit der Entwicklung von Kindern in einer traditionellen Grundschule von der ersten bis zur vierten Klassenstufe. Während sich die Montessorischüler in der Kreativität und in der Geometrie etwas günstiger entwickelten als die Schüler in der öffentlichen Schule, waren in der Intelligenz, der Mathematikleistung insgesamt und in der Rechtschreibleistung keine Unterschiede zwischen beiden Gruppen festzustellen, im Sachrechnen waren die Schüler der öffentlichen Grundschule überlegen.

nem erweiterten Bildungsangebot, das die umfassende Förderung der Begabung, der Intelligenz und der Persönlichkeit der Schüler beansprucht.

Die schulische Umsetzung der Konzeption begann in den 90er Jahren des letzten Jahrhunderts. Seit 1997 entstanden in mehreren neuen Bundesländern nach diesem Konzept freie Grundschulen, die inzwischen größtenteils als Ersatzschulen anerkannt sind. Bis heute wurden zehn Grundschulen und zwei Gymnasien eingerichtet. Der Besuch der BIP-Schulen kostet Schulgeld, das gegenwärtig bei maximal 370 Euro pro Monat liegt. Dies spricht dafür, dass die BIP-Kreativitätsgrundschulen insbesondere von Schülern aus sozial privilegierteren Elternhäusern besucht werden.

Die BIP-Kreativitätsgrundschulen unterscheiden sich in vielfacher Hinsicht von öffentlichen Grundschulen. Die besonderen Merkmale der BIP-Schulen werden im Folgenden näher beschrieben.³

2.1 Merkmale der Schule und des Curriculums

Zu den Besonderheiten der BIP-Schulen zählen kleine Klassen (max. 22 Kinder pro Klasse), die Öffnung der Schulen in den Ferien, der Ganztagesbetrieb mit Hortbetreuung sowie die über das Curriculum staatlicher Schulen hinausreichenden Angebote und Fördermaßnahmen. Diese beinhalten u. a. ein speziell gestaltetes kreativitätspädagogisches Förderprogramm. Die Kinder erhalten verpflichtend zusätzlichen Unterricht mit jeweils einer Stunde pro Woche in den sogenannten Krea-Fächern Strategisches Spiel/Schach, Darstellendes Spiel, Digitale Medien/Informatik, Tanz und Bewegung, Entdecken-Erforschen-Erfinden sowie sprachliches, bildkünstlerisches und musikalisches Gestalten. Dazu kommt, dass die Schüler an den BIP-Schulen ab dem ersten Schuljahr verpflichtend am umfangreichen Fremdsprachenangebot der Schulen teilnehmen, wobei sie im ersten Schuljahr Angebote in mindestens zwei Fremdsprachen (in der Regel Englisch und Arabisch), ab dem dritten Schuljahr teilweise auch zusätzliche Angebote in Französisch oder Spanisch erhalten (Schöppe, 2010). Außerdem haben die Schüler neben dem regulären Unterricht die Möglichkeit, Instrumentalunterricht zu besuchen. Insgesamt wird das wöchentliche Stundenpensum der Schüler durch den Unterricht in den Krea-Fächern und in den Fremdsprachen um bis zu zwölf Stunden in elf Fächern (je eine Stunde in den acht Krea-Fächern und zwei Fremdsprachen, zwei Stunden Englisch) erweitert.

3 Die hier genannten Merkmale der BIP-Konzeption beziehen sich auf Informationen aus folgenden Quellen: Mehlhorn (2001); Mehlhorn & Mehlhorn (2003); Schöppe (2010); <http://www.mehlhornschoolen.de/grundkonzept.pdf> [letzter Zugriff am 07.04.2010]; <http://www.bip-muenchen.de/bildungskonzept/grundschulkonzept.html> [letzter Zugriff am 07.04.2010]; <http://www.creativitieschools.com/> [letzter Zugriff am 07.04.2010]; http://www.probip-dresden.de/clients/probip/probip.nsf/content/BIPKreativitaet_Definition.htm [letzter Zugriff am 07.04.2010]; <http://www.krea-schulzentrum.de/konzept.html> [letzter Zugriff am 07.04.2010]. Einige Merkmale der BIP-Konzeption haben sich zwischenzeitlich leicht verändert.

Die BIP-Kreativitätsgrundschulen bezeichnen sich selbst als „Leistungsschulen“. Ziel ist, dass mindestens vier von fünf Kindern am Ende der Grundschulzeit Übergangsempfehlungen zum Gymnasium erhalten. Die Schulen wollen anstrengende, lustbetonte und konzentrierte Tätigkeiten anregen und fördern. Damit einhergehend werden hohe Leistungsanforderungen an die Schülerschaft gestellt. Beispielsweise werden Hauptfächer schon ab dem ersten Schuljahr benotet und individuelle Förderpläne für diejenigen Schüler erstellt, die in den Fächern Deutsch, Mathematik und Sachunterricht eine Note schlechter als 2 haben.

Neben anderen Differenzierungsmaßnahmen wie klassen- und jahrgangsübergreifendem Unterricht werden die Schüler in den Fächern Mathematik und Deutsch von zwei Lehrpersonen gemeinsam (Teamteaching) oder/und in zwei Gruppen von jeweils einer Lehrperson getrennt unterrichtet (Teilungsunterricht). Beim Teamteaching fungiert eine Lehrperson als verantwortliche Hauptlehrperson und die andere als Nebenlehrperson. Als Teilungskriterien für den getrennten Unterricht werden häufig die Leistungsfähigkeit und/oder die Arbeitsgeschwindigkeit eines Schülers herangezogen. Diese Teilung der Klassen ist allerdings nicht starr, sondern wird im Verlauf des Schuljahres der Lern- und Leistungsentwicklung der Schüler entsprechend angepasst. Der Teilungsunterricht soll der individuellen Förderung dienen, wobei ein besonderer Fokus auf die intensive Einzelförderung eher leistungsschwächerer Schüler gelegt werden soll.

Die Kinder bekommen in den BIP-Schulen keine täglichen Hausaufgaben, da alle Schularbeiten in der Schule unter Anleitung der Pädagogen geleistet werden sollen. Unterrichtsausfall gibt es an den BIP-Schulen laut eigenen Aussagen nicht, da bei fehlenden Lehrern immer eine qualifizierte, in der Regel fachgerechte Vertretung eingesetzt werde und die Kinder somit nicht außerplanmäßig nach Hause geschickt werden.

2.2 Merkmale der Lehrpersonen

Mit Blick auf die *Lehrpersonen* ist die BIP-spezifische Aus- und Fortbildung der Lehrpersonen zu Kreativitätspädagogen als verpflichtende Zusatzqualifizierung zu nennen. Diese müssen die Lehrpersonen bei Neuanstellung berufsbegleitend absolvieren. Die Kosten werden von den Lehrpersonen selbst getragen. Die Zusatzqualifizierung umfasst mindestens 1860 Stunden und ist laut Aussage der Ausbildungsinstitution vergleichbar mit einem fünf- bis sechssemestrigen universitären Studium. Zudem verpflichten sich die Lehrpersonen während ihrer Tätigkeit als Lehrer an einer BIP-Schule zur stetigen Weiterbildung.

BIP-Kreativitätsschullehrer sind aufgrund ihrer Anstellung an freien Schulen keine Beamten, sondern Angestellte, die bei Neuanstellungen zunächst einen befristeten Vertrag mit einjähriger Laufzeit bekommen.

Von den Lehrern der BIP-Schulen werden eine Identifikation mit den Werten und Zielen des Schulkonzepts und die Gestaltung besonderer Lehr- und Lernbedingungen erwartet (s. o.).

Bisherige Erfahrungen mit der Schulkonzeption deuten nach Ansicht der Schulgründer darauf hin, dass BIP-Schüler bei den Bildungsempfehlungen für die weiterführenden Schulen gegenüber Kindern aus öffentlichen Grundschulen besser abschneiden. Darüber hinaus wird von herausragenden Leistungen einzelner Kinder u. a. in Schach, Musik, Bildender Kunst, Mathematik und IT berichtet (vgl. Mehlhorn & Mehlhorn 2003, S. 35f.).

Neben der Frage nach belastbaren Evidenzen für Leistungsvorsprünge von BIP-Schülern stellt sich mit Blick auf die besonderen sozioökonomischen Lernvoraussetzungen der BIP-Schüler auch die Frage, inwieweit mögliche Leistungsvorsprünge der BIP-Schüler tatsächlich auf eine bessere schulische Förderung zurückzuführen oder lediglich das Ergebnis günstigerer Kontextbedingungen sind. In diesem Beitrag werden diese Fragen auf der Basis querschnittlicher Daten zu beantworten versucht, indem BIP-Schüler am Ende des 4. Schuljahres in ihren Leistungen mit Schülern öffentlicher Grundschulen verglichen werden.

3. Privatschulen als differentielle Lern- und Entwicklungsmilieus?

Stellt man sich die Frage, ob BIP-Kreativitätsgrundschulen besonders günstige Voraussetzungen für die Entwicklung von Schülern bieten, muss zunächst die Referenzgruppe definiert werden. Ohne statistische Kontrolle sozialer Hintergrundmerkmale auf Schüler- und Klassenebene sind Leistungsvorsprünge von BIP-Schülern gegenüber Schülern öffentlicher Schulen aufgrund der günstigeren individuellen Startvoraussetzungen plausibel und sehr wahrscheinlich. Relevanter ist die Frage, ob diese Leistungsvorsprünge auch dann noch Bestand haben, wenn man die individuellen sozialen Hintergrundbedingungen der Schüler und im zweiten Schritt auch die Kompositionsmerkmale der Klassen kontrollierend berücksichtigt.

Theoretisch werden Leistungsvorsprünge von Privatschülern mit einer höheren Schulautonomie, einem stärkeren Wettbewerb, einem kognitiv herausfordernden Unterricht, einem günstiger ausgeprägten Schulklima, mit höheren Leistungserwartungen von Lehrpersonen und Eltern, einem höheren Commitment der Lehrpersonen an die Schule und mit Wechselwirkungen zwischen diesen Variablen zu erklären versucht (Chubb & Moe, 1990; Coleman, Hoffer & Kilgore, 1982; McEwan, 2000; Opdenakker & Van Damme, 2006; Lubienski, Lubienski & Crane, 2008; zsf. Weiß, 2011).

Die Privatschulforschung stellt jedoch vielfach fest, dass bei Kontrolle der unterschiedlichen individuellen Voraussetzungen und der Schülerzusammensetzung keine Leistungsunterschiede zwischen Schülern an privaten und öffentlichen Schulen⁴ mehr nachzuweisen sind. Einige Studien berichten sogar Leistungsvorteile

4 Häufig werden in der Privatschulforschung drei Typen von Schulen unterschieden: (1) Privat finanzierte und geleitete Schulen, die unabhängig sind von staatlicher Einflussnahme und Anerkennung (private-independent) und die weniger als 50 % ihrer erforderlichen finanziellen Mittel vom Staat erhalten (2) Private Schulen, die zu einem erheb-

öffentlicher Schulen gegenüber Privatschulen (OECD, 2004, 2007; Opdenakker & Van Damme, 2006; Lubienski & Lubienski, 2006; Lubienski, Weitzel & Lubienski, 2009; zsf. Weiß, 2011; Wenglinsky, 2007).

Dronkers und Robert (2003, 2008a, 2008b), welche die PISA 2000- und die PISA 2006-Daten analysierten, ermittelten, dass der größte Teil an Leistungsunterschieden zwischen öffentlichen und privaten Schulen auf Unterschiede in den Lernvoraussetzungen und den Kompositionsbedingungen zurückgeführt werden kann. Dementsprechend kann auch für die hier untersuchten BIP-Schüler gefragt werden, wie sich der Einfluss der Schulzugehörigkeit verändert, wenn man zum einen die individuellen Lernvoraussetzungen und zum anderen die klassenspezifischen Kontextbedingungen kontrollierend berücksichtigt.

4. Fragestellungen

Unter Berücksichtigung der angeführten Forschungsarbeiten gehen wir folgenden Fragestellungen nach:

Fragestellung 1: Wie schneiden die BIP-Schüler beim Leistungsvergleich mit einer repräsentativen IGLU-Stichprobe am Ende des 4. Schuljahres im Lesen, in Mathematik und in den Naturwissenschaften ab, wenn die unterschiedlichen individuellen und kontextbezogenen Voraussetzungen der einzelnen Schüler und der Klassen nicht kontrolliert werden?

Fragestellung 2: Wie fallen die Leistungsunterschiede aus, wenn die BIP-Schüler solchen Schülern der IGLU-Stichprobe gegenüber gestellt werden, die über vergleichbare soziale Lernvoraussetzungen verfügen?

Fragestellung 3: Zeigen sich Leistungsvorsprünge der BIP-Schüler, wenn man neben individuellen Voraussetzungen auch die soziale Zusammensetzung der Klassen kontrollierend berücksichtigt?

5. Methode

5.1 Stichprobe

Die BIP-Stichprobe besteht insgesamt aus 279 Schülern der vierten Jahrgangsstufe aus 17 Klassen der BIP-Grundschulen in Leipzig, Chemnitz, Gera und Dresden. Die Daten wurden im Frühjahr 2005 und im Frühjahr 2006 erhoben.

Die für den Vergleich herangezogene nationale IGLU 2001-Stichprobe unterscheidet sich je nach untersuchter Forschungsfrage: In einem *ersten Analyseschritt* werden die Leistungen der BIP-Schüler mit der gesamten nationalen IGLU-

lichen Teil (mehr als 50 %) vom Staat finanziell alimentiert und unterstützt werden, sich aber in privater Trägerschaft befinden (private government-dependent) und (3) öffentlich finanzierte und geleitete Schulen (public). Bei den Privatschulen in Deutschland handelt es sich meist um private government-dependent Schulen. Dies gilt auch für die BIP-Kreativitätsgrundschulen.

Stichprobe aus dem Jahr 2001 verglichen. Diese Stichprobe setzt sich insgesamt aus 7633 Schülern zusammen, die alle am Lesetest teilgenommen haben. Die Fallzahlen für die Domänen Naturwissenschaften und Mathematik fallen deutlich geringer aus, weil diese Tests an einem zweiten Testtag durchgeführt wurden, an dem sich nicht alle Länder der Bundesrepublik Deutschland beteiligt haben ($N_{\text{Lesen}} = 7633$; $N_{\text{Naturwissenschaften}} = 5943$; $N_{\text{Mathematik}} = 5943$).

Die folgende Tabelle 1 verdeutlicht anhand von einigen Parametern, wie sich die BIP- und IGLU-Stichproben in sozialen und kognitiven Hintergrundvariablen voneinander unterscheiden. Der Vergleich offenbart, dass die BIP-Schüler aus sozial privilegierten und materiell besser gestellten Familien stammen als die IGLU-Schüler. So ist der höchste soziökonomische Index der Familien (HISEI) der BIP-Schüler im Durchschnitt bedeutend höher ausgeprägt als der der IGLU-Schüler. Auch das mittlere Haushaltseinkommen der BIP-Familien weist einen deutlich höheren Wert auf als das der IGLU-Familien. Von den BIP-Schülern sind nur etwas mehr als 3 % im Ausland geboren, während es von den IGLU-Schülern jeder fünfte ist. Eine entsprechende Verteilung zeigt sich auch beim Geburtsland der Eltern: Nur in ca. 5 % aller Fälle sind beide Elternteile der BIP-Schüler im Ausland geboren, während es in der IGLU-Gesamtstichprobe ca. dreimal soviel sind.

Tabelle 1: Vergleich der gesamten BIP-Stichprobe mit der IGLU-Gesamtstichprobe in sozialen Hintergrundmerkmalen und kognitiven Lernvoraussetzungen (nicht imputierte Daten)

| | BIP-Stichprobe ($N = 279$) | IGLU 2001-Gesamtstichprobe ($N = 7633$) |
|----------------------------------------------------------------------------------------------------------|---------------------------------|----------------------------------------------|
| | M (SD) oder Anteil | M (SD) oder Anteil |
| HISEI | 64.5 (15.1) | 48.7 (16.2) |
| Einkommen (sechsstufig) ^a | 4.3 (1.8) | 3.2 (1.6) |
| Anteil der Haushalte, in denen ein zweites Auto vorhanden ist | 70.4 % | 48.5 % |
| Anteil der Haushalte, in denen eine Tageszeitung vorhanden ist | 76.5 % | 62.9 % |
| Anteil der Haushalte, in denen eine Geschirrspülmaschine vorhanden ist | 87.8 % | 68.5 % |
| Anteil der Haushalte, in denen Kunstwerke vorhanden sind | 85.1 % | 78.1 % |
| Anzahl der Bücher (fünfstufig) ^b | 4.0 (1.0) | 3.1 (1.2) |
| Geburtsland des Kindes: Anteil der Kinder, die im Ausland geboren sind | 3.5 % | 20.8 % |
| Migrationshintergrund: Anteil der Familien, in denen beide Elternteile nicht in Deutschland geboren sind | 5.0 % | 14.3 % |
| Geschlecht: Anteil der Jungen | 53.0 % | 50.2 % |

^aEinkommen Antwortformat: (1) unter 20000 Euro, (2) 20 000 bis 29 999 Euro, (3) 30 000 bis 39 999 Euro, (4) 40 000 bis 49 999 Euro, (5) 50 000 bis 59 999 Euro, (6) 60 000 Euro oder mehr. ^bAnzahl der Bücher Antwortformat: 1 = „keine oder sehr wenige Bücher“ 0–10 Bücher; 2 = „genug um ein Regalbrett zu füllen“ 11–25 Bücher; 3 = „genug um ein Regal zu füllen“ 26–100 Bücher; 4 = „genug um zwei Regale zu füllen“ 101–200 Bücher; 5 = „genug um drei oder mehr Regale zu füllen“ über 200 Bücher.

Die Analyseschritte 2 und 3 sehen einen Vergleich von BIP-Schülern mit IGLU-Schülern vor, die über ähnlich günstige Voraussetzungen verfügen wie die BIP-Schüler. Hierzu wird ein Propensity Score Matching (s. u.) durchgeführt. Mit Hilfe dieses Verfahrens wird jedem der 279 BIP-Schüler ein Schüler mit vergleichbaren sozialen Hintergrundvoraussetzungen aus der IGLU-Stichprobe zugeordnet. Im Folgenden werden das angewandte Propensity-Score-Matching Verfahren und deskriptive Parameter der beiden Substichproben genauer dargestellt.

5.2 Propensity Score Matching

Um Einflüsse der Schulzugehörigkeit (Privatschule vs. öffentliche Schule) auf die Entwicklung von Schülern zu untersuchen, kommen grundsätzlich (Mehrebenen-) Regressionsanalysen oder Matching-Verfahren in Frage. Bei ersteren wird der Effekt der Schulzugehörigkeit für die gesamte Population geschätzt (*average treatment effect* – ATE). Dieses Vorgehen gilt allerdings als nicht unproblematisch, wenn die Vergleichsgruppe (im vorliegenden Fall die IGLU-Stichprobe) sehr groß und überdies heterogen zusammengesetzt ist, während die Untersuchungsgruppe (hier die BIP-Stichprobe) vergleichsweise klein und relativ homogen und zudem durch besondere Merkmale charakterisiert ist (s. o.). Daher bietet sich für die Bearbeitung der zweiten und dritten Forschungsfrage die zweite Option an. Bei Matching-Verfahren wird der Effekt der Schulzugehörigkeit nur für diejenigen Schüler bestimmt, die aufgrund ihrer Hintergrundmerkmale typisch für die BIP-Schulen sind (*average treatment effect for the treated* – ATT). So ist es möglich, sich durch die Wahl des Analyseverfahrens einem experimentellen Versuchsdesign anzunähern.

In diesem Beitrag wird mit dem von Rosenbaum und Rubin (1983, 1985) entwickelten *Propensity Score Matching* (PSM) das zurzeit gängigste Matching-Verfahren genutzt. Dabei wird versucht, eine Vergleichsgruppe zu bilden, deren Mitglieder aufgrund solcher Merkmale, welche nicht schon durch das Treatment selbst beeinflusst wurden (Caliendo & Kopeinig, 2008), eine ähnliche Wahrscheinlichkeit hätten, zur Untersuchungsgruppe zu gehören. Für die vorliegende Untersuchung geht es also darum, in der Gruppe der IGLU-Schüler solche Schüler zu identifizieren, die aufgrund ihrer sozialen Hintergrundbedingungen eine vergleichbare Wahrscheinlichkeit hätten, eine BIP-Schule zu besuchen. Für das Matching sind solche Hintergrundvariablen zu bestimmen und heranzuziehen, die die Schulwahl der Eltern beeinflusst haben bzw. mit der Schulwahl der Eltern in Verbindung stehen könnten. Annehmen lässt sich, dass die Wahl einer Privatschule in erheblichem Maße von Merkmalen der sozialen Herkunft der Eltern, dem Einkommen sowie dem ökonomischen und kulturellen Kapital der Familien abhängig sein dürfte (vgl. Bosetti, 2004). Daher wurden für das Matching in dieser Studie Hintergrundvariablen zum sozialen Status und zum Migrationshintergrund der Schüler herangezogen. Als Indikatoren für das ökonomische Kapital wurden das Einkommen (sechsstufig), der Besitz eines zweiten Autos bzw. einer

Geschirrspülmaschine (jeweils dichotom: 0 = nein, 1 = ja) sowie der höchste sozio-ökonomische Index der Familie (HISEI) einbezogen. Als Indikatoren für das kulturelle Kapital einer Familie wurden die Anzahl der im Haushalt vorhandenen Bücher (ohne Zeitschriften, Zeitungen und Schulbücher, fünfstufig), das Abonnement einer Tageszeitung sowie der Besitz von Kunstwerken (jeweils dichotom: 0 = nein, 1 = ja) genutzt. Der Migrationshintergrund des Kindes wurde über das Geburtsland des Kindes (0 = Deutschland, 1 = Ausland) und über das Geburtsland der Eltern (dreistufig; 0 = kein, 1 = ein, 2 = beide Elternteile im Ausland geboren) erfasst.

Im Vorfeld des Propensity Score Matching musste wie bei allen multivariaten Verfahren zunächst der Umgang mit fehlenden Werten in den Datensätzen festgelegt werden. In den letzten Jahren hat sich auch in der deutschen Erziehungswissenschaft das modellbasierte Schätzen von fehlenden Werten mit Hilfe der Imputation als zu präferierendes Vorgehen etabliert (Lüdtke, Robitzsch, Trautwein & Köller, 2007).

Für diesen Beitrag wurden in R 2.12.0 (R Development Core Team, 2011) mit dem Package mice (van Buuren & Groothuis-Oudshoorn, 2009) alle fehlenden Werte in einem kombinierten Datensatz mit allen IGLU- und allen BIP-Schülern einfach ersetzt. Dabei wurden neben den Variablen, die für das Propensity Score Matching benötigt werden (s. o.), zusätzlich die Leistungsscores in den drei getesteten Domänen und die Schultypzugehörigkeit (0 = IGLU, 1 = BIP) genutzt.

Das Matching wurde mit dem R-Package MatchIt (Ho, Imai, King & Stuart, 2011) durchgeführt. Zunächst wird dabei eine logistische Regression mit der Schultypzugehörigkeit als abhängige und den für das Matching ausgewählten Indikatoren als unabhängigen Variablen berechnet. Die erhaltenen bedingten Wahrscheinlichkeiten, eine BIP-Schule zu besuchen, werden dann genutzt, um IGLU-Schüler zu identifizieren, die auf Grund ihrer Hintergrundmerkmale eine entsprechend große Wahrscheinlichkeit für den Besuch einer BIP-Schule aufweisen.

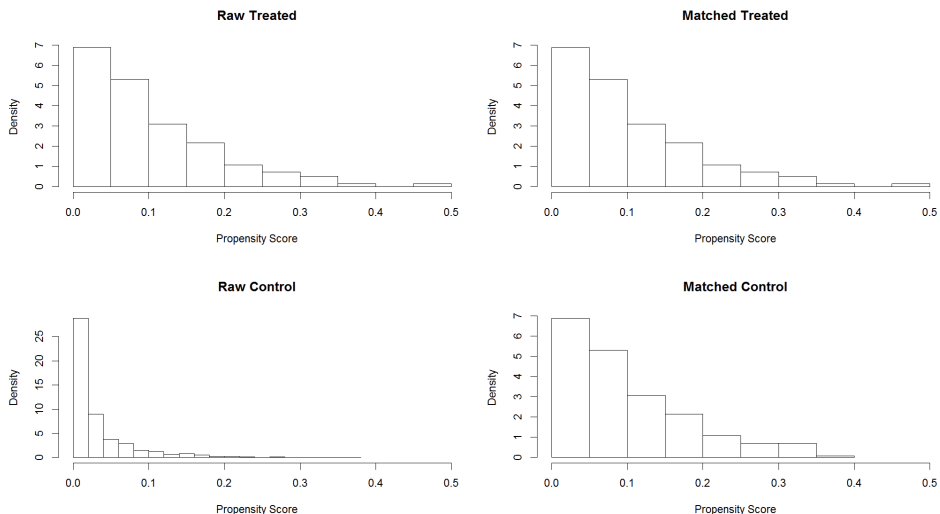
Für diesen Beitrag wurde die Methode Nearest Neighbor Matching ohne Zurücklegen genutzt. Dabei wird nacheinander für jeden Fall der Treatmentgruppe der Fall der Kontrollgruppe ausgewählt, der diesem am ähnlichsten ist. „Ohne Zurücklegen“ bedeutet, dass jeder Fall der Kontrollgruppe nur einem Fall der Treatmentgruppe zugeordnet werden kann.

Ein wichtiger abschließender Schritt ist die Kontrolle der gematchten Datensätze dahingehend, für wie viele Fälle der Treatmentgruppe kein „Zwilling“ in der Kontrollgruppe gefunden werden konnte und ob sich die beiden Gruppen tatsächlich nicht mehr hinsichtlich der für das Matching genutzten Variablen unterscheiden (vgl. vertiefend Baumert, Becker, Neumann & Nikolova, 2009). Die entsprechenden Ergebnisse werden im folgenden Abschnitt dargestellt.

5.3 Ergebnisse des Matching-Verfahrens

Für jeden der 279 BIP-Schüler konnte ein IGLU-Zwilling gefunden werden. Die Abbildung 1 verdeutlicht die Überlappungsbreite der beiden Stichproben vor und nach dem Matching. Die beiden linken Histogramme verdeutlichen, wie sich die Propensity Scores vor dem Matching verteilen, die beiden rechten Histogramme verdeutlichen die Verteilung nach dem Matching. Erwartungskonform zeigt sich, dass in der IGLU-Gesamtstichprobe (Raw Control) Schüler mit geringen Propensity Scores deutlich in der Mehrheit sind, während sich nach dem Matching die Propensity Scores der beiden Stichproben kaum noch unterscheiden (Matched Treated vs. Matched Control).

Abbildung 1: Verteilungen der *Propensity Scores* in den beiden Stichproben (BIP = Treated, IGLU = Control) vor (Raw) und nach (Matched) dem Matching



Auch die folgende Abbildung 2 dokumentiert die Güte des Matchings: Nach dem Matching (Matched Data) liegen die Mittelwertsunterschiede nahe bei Null, d. h. die Unterschiede zwischen beiden Stichproben sind gering, während sie vor dem Matching (All Data) teilweise beträchtlich waren.

Abbildung 2: Mittelwertsunterschiede der Hintergrundvariablen vor und nach dem Matching



Die folgende Tabelle 2 gibt einen Überblick über die Mittelwerte und Standardabweichungen bzw. die Häufigkeiten der Modellvariablen nach dem durchgeführten Propensity Score Matching. Wie der Tabelle zu entnehmen ist, finden sich zwischen den beiden Stichproben nur noch geringe Unterschiede, was für die Qualität des durchgeführten Matchings spricht.

Tabelle 2: Vergleich der BIP-Stichprobe mit den „IGLU-Zwillingen“ in denjenigen Variablen, die für das Matching herangezogen wurden (fehlende Daten wurden vorab imputiert)

| | BIP-Stichprobe (<i>N</i> = 279) | IGLU-„Zwillinge“ (<i>N</i> = 279) |
|----------------------------------------------------------------------------------------------------------|---------------------------------------|---------------------------------------|
| | <i>M</i> (<i>SD</i>) oder Anteil | <i>M</i> (<i>SD</i>) oder Anteil |
| HISEI | 63.8 (15.5) | 62.9 (15.4) |
| Einkommen (sechsstufig) ^a | 4.3 (1.8) | 4.4 (1.6) |
| Anteil der Haushalte, in denen ein zweites Auto vorhanden ist | 70.6 % | 74.6 % |
| Anteil der Haushalte, in denen eine Tageszeitung vorhanden ist | 76.3 % | 77.4 % |
| Anteil der Haushalte, in denen eine Geschirrspülmaschine vorhanden ist | 87.5 % | 91.4 % |
| Anteil der Haushalte, in denen Kunstwerke vorhanden sind | 86.7 % | 87.5 % |
| Anzahl der Bücher (fünfstufig) ^b | 4.0 (1.0) | 4.0 (1.0) |
| Geburtsland des Kindes: Anteil der Kinder, die im Ausland geboren sind | 3.2 % | 2.5 % |
| Migrationshintergrund: Anteil der Familien, in denen beide Elternteile nicht in Deutschland geboren sind | 6.8 % | 4.7 % |
| Geschlecht: Anteil der Jungen | 53.0 % | 54.1 % |

Anmerkung. Abweichungen zu den in Tabelle 1 berichteten Werten für die 279 BIP-Schüler ergeben sich durch die hier vorab durchgeführte Imputation fehlender Werte.

^aEinkommen Antwortformat: (1) unter 20 000 Euro, (2) 20 000 bis 29 999 Euro, (3) 30 000 bis 39 999 Euro, (4) 40 000 bis 49 999 Euro, (5) 50 000 bis 59 999 Euro, (6) 60 000 Euro oder mehr. ^bAnzahl der Bücher Antwortformat: 1 = „keine oder sehr wenige Bücher“ 0–10 Bücher; 2 = „genug um ein Regalbrett zu füllen“ 11–25 Bücher; 3 = „genug um ein Regal zu füllen“ 26–100 Bücher; 4 = „genug um zwei Regale zu füllen“ 101–200 Bücher; 5 = „genug um drei oder mehr Regale zu füllen“ über 200 Bücher.

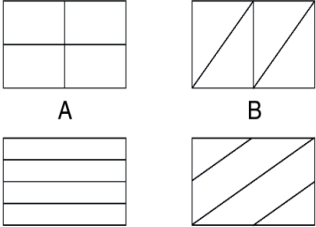
5.4 Erfassung der Schülerleistungen

Die Leistungen der Schüler wurden mit den IGLU-Testinstrumenten erhoben. Im Folgenden sind einige Beispielitems der 2001er Studie dargestellt. Die Mathematikaufgabe in Abbildung 3 und die Naturwissenschaftsaufgabe in Abbildung 4 stammen ursprünglich aus der *Trends in International Mathematics and Science Study* (TIMSS) 1995 (Martin et al., 1997; Mullis et al., 1997) und wurden 2001 im Rahmen der nationalen Erweiterung von IGLU eingesetzt (Prenzel, Geiser, Langeheine & Lobemeier, 2003; Walther, Geiser, Langeheine & Lobemeier, 2003).

Das Beispielitem aus dem Bereich Mathematik (vgl. Abbildung 3) gehört zum Inhaltsbereich *Geometry* und zum kognitiven Anforderungsbereich *Knowing*. Es handelt sich dabei um eine Aufgabe, die im Rahmen von IGLU 2001 mittelschwer war und von 62 % der Viertklässler korrekt gelöst wurde (in der Stichprobe der BIP-Schüler beträgt die Lösungswahrscheinlichkeit 94 %).

Abbildung 3: Beispielitem Mathematik

Hier siehst du vier große Rechtecke.
Drei Rechtecke sind in vier gleiche Teile zerschnitten, eines nicht.
Welches Rechteck ist nicht in vier gleiche Teile zerschnitten?



A B
C D

Bei dem abgebildeten naturwissenschaftlichen Beispielitem (vgl. Abbildung 4) handelt es sich um eine der schwierigsten Aufgaben im Test, die bundesweit nur 31 % der Kinder korrekt gelöst haben (BIP: 50 %). Das Item wird dem Inhaltsbereich *Environmental Issues and the Nature of Science* und dem Anforderungsbereich *Investigating the Natural World* zugeordnet.

Abbildung 4: Beispielitem Naturwissenschaften

Um herauszufinden, ob Samen im Licht oder im Dunkeln besser wachsen, kannst du einige Samen auf ein feuchtes Papier legen und

- sie an einen warmen, dunklen Ort stellen.
- einen Teil an einen hellen Ort und einen anderen Teil an einen dunklen Ort stellen.
- sie an einen warmen, hellen Ort stellen.
- sie an einen hellen oder dunklen, aber kühlen Ort stellen.

Zur Erfassung der Lesekompetenz wurden der literarische Text „Mäuse auf dem Kopf“ und der Sachtext „Flusswanderung“ aus IGLU 2001 genutzt. Die vollständigen Texte und die zugehörigen Aufgaben wurden in Voss (2006) veröffentlicht.

Abbildung 5: Beispielitem Lesen: „Mäuse im Kopf“

Wie wird in der Geschichte gezeigt, was die Mäuse über das Geschehene dachten?

- Es wird erzählt, was Labon über die Mäuse dachte.
- Es wird beschrieben, wo die Mäuse lebten.
- Es wird erzählt, was die Mäuse zueinander sagten.
- Es wird beschrieben, wie die Mäuse aussahen.

Abbildung 6: Beispielitem Lesen: „Flusswanderung“

Wo beginnt der Flusswanderweg?

- Banheim
- Gründorf
- Altenburg
- Parklandschaft am Fluss

In den Abbildungen 5 und 6 ist exemplarisch jeweils eine Aufgabe zu den beiden Texten dargestellt. Das Beispielitem aus dem Mäuse-Text gehört zur Kompetenzstufe III (‘Implizit im Text enthaltene Sachverhalte aufgrund des Kontextes erschließen’) und wurde im Rahmen von IGLU 2001 von 58 % der Schüler korrekt beantwortet (BIP: 66 %). Deutlich einfacher ist die Aufgabe „Flusswanderung“, die bundesweit von 79 % der Viertklässler gelöst wurde (BIP: 95 %) und der Kompetenzstufe I (‘Gesuchte Wörter in einem Text erkennen’) zuzuordnen ist.

Der Skalierung der Leistungswerte lag folgendes Verfahren zugrunde: Zunächst wurden die Itemschwierigkeiten und die Personenfähigkeiten auf Basis eines kombinierten Datensatzes mit allen IGLU- und allen BIP-Schülern unter Berücksichtigung aller im Rahmen von IGLU 2001 eingesetzten Lesetestaufgaben ermittelt (*concurrent calibration*). Danach wurden die Personenscores so transformiert, dass über alle IGLU-Schüler hinweg die Mittelwerte 500 und die Standardabweichungen 100 betragen. So lassen sich die drei Gruppen (BIP-Schüler, IGLU-Gesamtstichprobe, IGLU-„Zwillinge“) direkt miteinander vergleichen.

5.5 Einbezug von Klassenkompositionsmerkmalen

Leistungsvorsprünge von Privatschülern lassen sich auch auf die günstigeren Kompositionsmerkmale von Klassen an Privatschulen zurückführen (s. o., Dronkers & Robert, 2003, 2008a, 2008b). Dabei sind mehrere Erklärungsansätze voneinander zu unterscheiden (vgl. Bellin, 2009). *Didaktische und curriculare Ansätze* gehen davon aus, dass Schüler in leistungsstärkeren Klassen mehr und besseren Unterricht erhalten, von Lehrern und Mitschülern stärker kognitiv gefordert und angeregt werden und in besonderer Weise von dem leistungsförderlichen Lernklima in leistungsstärkeren Klassen profitieren. *Soziologische Ansätze* betonen die höheren Bildungsaspirationen von Mitschülern und Eltern in leistungsstärkeren Klassen und den dadurch veränderten Referenzrahmen für soziale Vergleiche, während *psychologische Ansätze* unter anderem die Erwartungshaltung der Lehrperson in Abhängigkeit von der Leistungsstärke der Klasse fokussieren: Lehrpersonen entwickeln in leistungsschwächeren Klassen eher „eine unangemessene niedrige Erwartungshaltung“ (Bellin, 2009, S. 90) gegenüber ihren Schülern und beeinflussen über Stigmatisierungsprozesse die Lernhaltung der Schüler. In leistungsstarken Klassen wäre demnach ein gegenteiliger Effekt zu erwarten (Bellin, 2009).

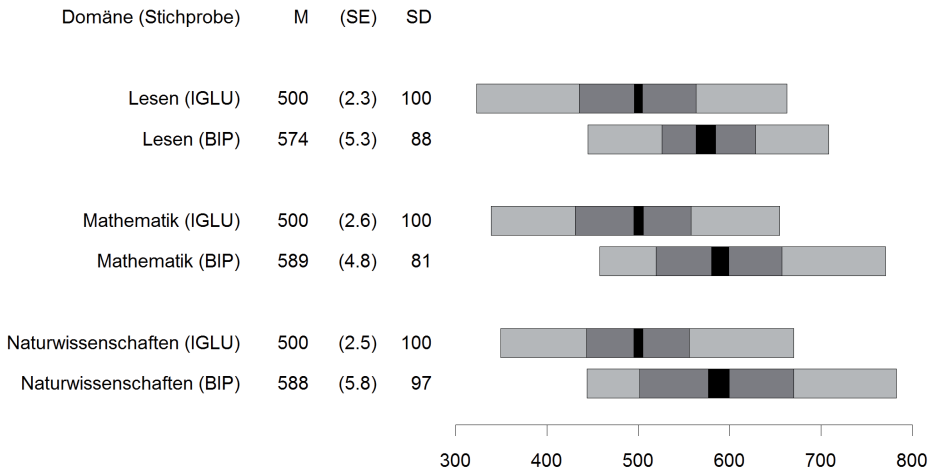
Tatsächlich liegen zahlreiche empirische Hinweise dafür vor, dass die mittlere Leistungsfähigkeit einer Klasse als ein relevantes Kompositionsmerkmal die individuelle Leistungsentwicklung der einzelnen Schüler in der Klasse positiv beeinflusst. Das bedeutet: Mit einem Anstieg der Leistungsfähigkeit einer Klasse sind auch bessere individuelle Leistungen der Lernenden verbunden, auch wenn die Lernenden ursprünglich über gleiche oder ähnliche individuelle Startvoraussetzungen verfügten (vgl. z. B. Beckerman & Good, 1981; Dar & Resh, 1986; Van Damme, De Fraine, Van Landeghem, Opdenakker & Onghena, 2002; Rindermann, 2007; Tiedemann & Billmann-Mahecha, 2004). Daher wird in den nachfolgenden Analysen mit dem mittleren HISEI-Wert der jeweiligen Klassen auch ein Indikator für die soziale Zusammensetzung der Klassen kontrolliert. Dass sich hierin die beiden gematchten Stichproben deutlich unterscheiden, verdeutlicht der Blick auf die beiden Mittelwerte und die Streuungen. Während der Mittelwert der HISEI-Mittelwerte der 17 Klassen, aus denen die 279 BIP-Schüler stammen, bei $M = 64.0$ liegt ($SD = 3.2$), beträgt der Mittelwert der HISEI-Klassenmittelwerte derjenigen 189 Klassen, aus denen die 279 IGLU-Schüler kommen, $M = 47.9$ ($SD = 5.7$).

6. Ergebnisse

In einem ersten Schritt wird überprüft, wie die Schüler der BIP-Schulen im Vergleich zur gesamten IGLU-Stichprobe abschneiden (Fragestellung 1). Wie beschrieben, wurden die Leistungen der BIP-Schüler hierzu auf der gleichen Skala verankert wie die Leistungen der IGLU-Gesamtstichprobe. Abbildung 7 zeigt er-

wartungskonform erhebliche Leistungsvorsprünge der BIP-Schüler im Lesen, in den Naturwissenschaften und in Mathematik.

Abbildung 7: Leistungsergebnisse der BIP-Schüler im Vergleich zur gesamten IGLU-Stichprobe



Anmerkung. Der schwarze Bereich der Balken kennzeichnet das Konfidenzintervall ($M \pm 2 SE$). Dunkelgrau ist der Bereich vom 25. bis zum 75. Perzentil eingefärbt und hellgrau der Bereich vom 5. bis zum 95. Perzentil.

Berücksichtigt man, dass eine halbe Standardabweichung (50 Punkte) etwa dem Lernzuwachs eines Schuljahres entspricht (Bos et al., 2007), so lässt sich gegenüber der IGLU-Gesamtstichprobe von einem Vorsprung der BIP-Schüler von ca. 1½ bis 1¾ Jahren ausgehen. Die Unterschiede sind signifikant ($p < .01$) und praktisch bedeutsam. Die ermittelten Cohens d -Werte betragen für das Lesen $d = 0.79$, für Mathematik $d = 0.98$ und für die Naturwissenschaften $d = 0.88$.

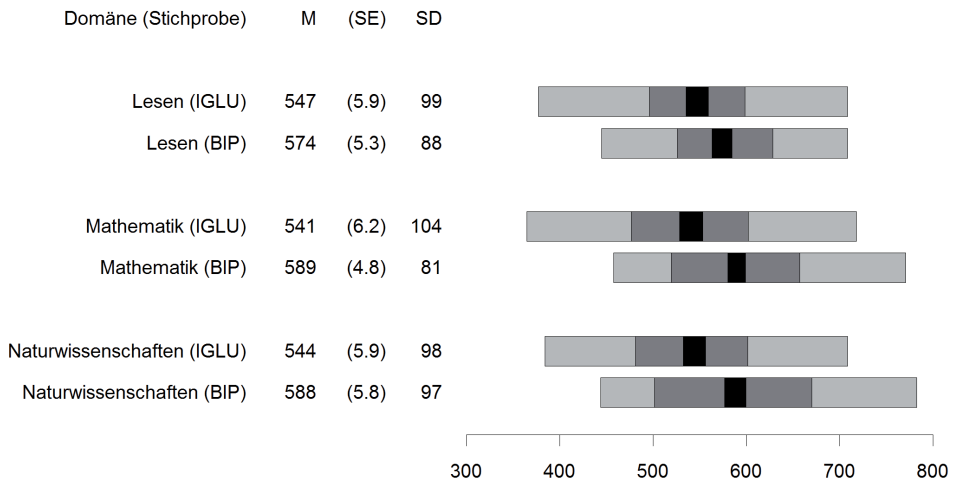
Der Vergleich der Streuungen (Abbildung 7) zeigt, dass die Leistungen der BIP-Schüler insgesamt homogener ausfallen als die der IGLU-Schüler. Im Lesen beträgt die Standardabweichung 88 Punkte, in Mathematik 82 Punkte und in den Naturwissenschaften 96 Punkte. Die Streuung der IGLU-Gesamtstichprobe beträgt dagegen jeweils 100 Punkte.

Insgesamt erlauben diese Differenzen zwischen den beiden Stichproben noch keine Rückschlüsse auf den Effekt der Konzeption der BIP-Schulen auf die Leistungsfähigkeit, da ja – wie bereits oben dargestellt – erhebliche Unterschiede in der sozialen Herkunft der beiden Schülerpopulationen bestehen. Entsprechend der Fragestellung 2 ist daher zunächst zu klären, ob die BIP-Schüler auch dann noch bessere Leistungen zeigen, wenn man sie mit ausgewählten „IGLU-Zwillingen“ vergleicht, die über ähnliche soziale Lernvoraussetzungen verfügen wie die BIP-Schüler.

Um diese Frage zu überprüfen, wurden die Leistungen der BIP-Schüler mit denen der gematchten IGLU-Zwillinge verglichen. Die Abbildung 8 stellt die entspre-

chenden Ergebnisse für die gesamte Matching-Stichprobe ($N = 279$ BIP-Schüler und $N = 279$ IGLU-Schüler) dar. Regressionsanalysen (ohne Tabelle) zeigen, dass die BIP-Schüler auch im Vergleich zu den IGLU-Zwillingen in allen drei Domänen nach wie vor signifikante ($p < .01$), wenngleich weniger große Leistungsvorsprünge aufweisen. Im Lesen beträgt der Unterschied 27 Punkte, was einer Effektstärke von $d = 0.28$ entspricht. Der Unterschied in Mathematik liegt bei 49 Punkten und entspricht einer Effektstärke von $d = 0.52$. In Naturwissenschaften beträgt der Unterschied 44 Punkte bei einer Effektstärke von $d = 0.45$.

Abbildung 8: Leistungsvergleich der BIP-Schüler mit den gematchten „IGLU-Zwillingen“



Anmerkung. Der schwarze Bereich der Balken kennzeichnet das Konfidenzintervall ($M \pm 2 SE$). Dunkelgrau ist der Bereich vom 25. bis zum 75. Perzentil eingefärbt und hellgrau der Bereich vom 5. bis zum 95. Perzentil.

Keine Berücksichtigung fanden in den bisherigen Analysen mögliche Effekte von Unterschieden in der Klassenkomposition. Um entsprechende Unterschiede zu kontrollieren, wäre analog zum obigen Vorgehen ein Matching auf Klassenebene das anzustrebende Verfahren. Da jedoch keine der IGLU-Klassen über eine so günstige Komposition verfügt wie die BIP-Klassen und ein Matching demzufolge nicht möglich ist, nutzen wir zur Beantwortung der dritten Forschungsfrage den auf Schülerebene gematchten Datensatz und ziehen als Indikator für die Klassenzusammensetzung den mittleren HISEI-Wert einer Klasse heran. Jedem der 558 Schüler wird also der über alle Schüler der Klasse aggregierte HISEI-Wert der eigenen Klasse zugeordnet.

Tabelle 3: Einflüsse der Klassenkomposition und der Schultypzugehörigkeit auf die Leistungen im Lesen, in Mathematik und in Naturwissenschaften Ende des 4. Schuljahres (Ergebnisse von linearen Regressionsanalysen, Einschlussverfahren)

| Prädiktoren | AV: Leistungen im/in ... | | | | | |
|--------------------|--------------------------|----------|------------|----------|---------------------|----------|
| | Lesen | | Mathematik | | Naturwissenschaften | |
| | Modell 1 | | Modell 2 | | Modell 3 | |
| | <i>B</i> | <i>T</i> | <i>B</i> | <i>T</i> | <i>B</i> | <i>T</i> |
| BIP | <i>ns</i> | -0.89 | 35.85** | 2.66 | <i>ns</i> | 0.55 |
| mittlerer HISEI | 3.00** | 3.55 | <i>ns</i> | 1.19 | 2.80** | 3.18 |
| Adjustiertes R^2 | .04 | | .06 | | .06 | |

Anmerkungen. *B* = unstandardisierter Regressionskoeffizient, *T* = *T*-Wert, *ns* = nicht signifikant; BIP: 1 = BIP-Schüler; 0 = „IGLU-Zwilling“, *N* = 558 Schüler.

** $p < .01$.

Die Ergebnisse zeigen, dass der Effekt des Schultyps (private BIP-Schule vs. öffentliche Schule) bis auf das Fach Mathematik unter die Signifikanzgrenze sinkt, sofern man als Indikator für die Klassenzusammensetzung jedem Schüler den mittleren HISEI-Wert seiner Klasse zuordnet. Im Lesen und in den Naturwissenschaften fallen die Unterschiede zwischen BIP-Schülern und Schülern der IGLU-Stichprobe somit nicht mehr signifikant aus.

7. Diskussion

Im Rahmen dieses Beitrags wurde die Konzeption der sogenannten BIP-Kreativitätsgrundschulen, eines privaten Schulkonzepts mit Schulstandorten vorwiegend in Ostdeutschland, einer Evaluation unterzogen, indem die Leistungen der Viertklässler an den BIP-Schulen mit den Leistungen einer repräsentativen IGLU-Stichprobe verglichen wurden. Wie erwartet zeigen die durchgeführten Analysen zunächst deutliche Leistungsvorsprünge der BIP-Schüler gegenüber der repräsentativen IGLU-Stichprobe. Die Leistungsvorsprünge machen dabei in den Domänen Mathematik, Naturwissenschaften und Lesen am Ende des vierten Schuljahres ca. 1½ Schuljahre aus.

Um die unterschiedlichen Startvoraussetzungen der beiden Teilstichproben zu berücksichtigen, wurde ein Propensity Score Matching unter Einbezug von Variablen des sozialen Hintergrunds der Schüler durchgeführt. Hierdurch ließen sich in der repräsentativen IGLU-Stichprobe 279 Schüler identifizieren, die aus vergleichbar privilegierten Elternhäusern stammen wie die BIP-Schüler. Vergleicht man diese beiden gematchten Stichproben, so ergeben sich weiterhin signifikante Leistungsunterschiede zugunsten der BIP-Schüler in allen drei Domänen. In Mathematik und in den Naturwissenschaften entsprechen die Leistungsvorsprünge knapp einem Schuljahr und im Lesen etwa einem halben Schuljahr.

Kontrolliert man mit dem mittleren HISEI-Wert zusätzlich ein Klassenkompositionsmerkmal, so ist nur in Mathematik noch ein Leistungsunterschied zugunsten der BIP-Schüler festzustellen. Der Unterschied beträgt 36 Punkte, was in etwa einem Vorsprung von 8 Monaten entspricht.

Die hier ermittelten Befunde lassen sich in Einklang bringen mit Befunden aus der internationalen Privatschulforschung (s. o.), wonach unter Kontrolle der besonders günstigen sozialen Hintergrundmerkmale von Privatschülern häufig keine oder nur geringe Leistungsvorsprünge gegenüber Schülern in öffentlichen Schulen nachzuweisen sind. McEwan (2000) fasst eine Vielzahl von Studien zusammen, die u. a. mit experimentellen oder quasi-experimentellen Studiendesigns Bildungsprogramme untersuchten, die einkommensschwachen Familien über die Verlosung von Bildungsgutscheinen den Besuch von Privatschulen ermöglichten. In diesen Studien wurde die Leistungsentwicklung von Schülern, die über ein Losverfahren einen Platz an einer Privatschule erlangten, mit der Leistungsentwicklung von Schülern verglichen, die keinen Platz an einer Privatschule erhielten und weiterhin an einer öffentlichen Schule verblieben. Insgesamt sprechen die Befunde dieser Studien keine einheitliche Sprache, was die Vorteile eines Wechsels auf die Privatschule anbelangt. McEwan (2000) kann aber in seinem Überblick etwas größere Effekte für Mathematik ausmachen, während sich im Lesen in den meisten der zusammengefassten Studien keine signifikanten Effekte zugunsten derjenigen Schüler ergaben, die auf eine Privatschule wechselten.

Auch in dieser Studie traten Unterschiede zugunsten der Schüler an den privaten BIP-Schulen nur in Mathematik und nicht im Lesen und in den Naturwissenschaften auf. Die damit vergleichsweise geringen Leistungsunterschiede zwischen den BIP-Schülern und den „Zwillingen“ an den staatlichen Schulen decken sich auch mit ersten Ergebnissen der dieser Querschnittstudie folgenden Längsschnittstudie PERLE, welche die Persönlichkeits- und Lernentwicklung von Kindern an BIP-Grundschulen und staatlichen Grundschulen über die gesamte Grundschulzeit untersuchte. Bislang liegen aus der PERLE-Längsschnittstudie Ergebnisse zur Leistungsentwicklung bis zum Ende des 2. Schuljahres vor (vgl. Lipowsky, Faust & Kastens, 2013).⁵

Demnach zeigte sich lediglich für das Rechtschreiben, welches in der vorliegenden Querschnittstudie nicht getestet wurde, eine günstigere Entwicklung zugunsten der BIP-Schüler bis zum Ende des 2. Schuljahres. Für das Lesen konnten – analog zu den hier berichteten Ergebnissen der Querschnittstudie – auch für die erste Phase des Längsschnitts keine Vorteile abgesichert werden (Lotz, Schoreit &

5 Zu beachten ist auch, dass sich die Stichprobe der Längsschnittstudie in anderer Weise zusammensetzte als die Stichprobe der vorliegenden Untersuchung. Die Kontrollgruppe in der Längsschnittstudie stammte aus Schulen in Sachsen, welche von der Schulverwaltung zur Teilnahme an der Studie vorgeschlagen worden waren. Die bislang vorliegenden Ergebnisse der Längsschnittstudie zum Vergleich der BIP-Schüler mit den Schülern der staatlichen Schulen beruhen zudem nicht auf einem Propensity Score Matching, sondern auf Mehrebenenanalysen, in die die Schülerleistungen unmittelbar nach der Einschulung als Eingangsvoraussetzungen kontrollierend einbezogen wurden.

Kempter, 2013). In Mathematik ergab sich in der Längsschnittstudie kein Effekt des Schultyps (BIP vs. staatlich) auf die Leistungsentwicklung bis zum Ende des 2. Schuljahres (Karst & Lipowsky, 2013), die Leistungen in den Naturwissenschaften wurden in der Längsschnittstudie nicht erfasst.

Berücksichtigt man die erweiterten Lerngelegenheiten und Bildungsangebote an den BIP-Schulen, kann demnach in den untersuchten Domänen nicht von einer Überlegenheit der BIP-Schulen gesprochen werden. Einmal mehr bestätigen sich damit Befunde der Schul- und Unterrichtsforschung, wonach Merkmale der Schule (hier: die Schulträgerschaft) insgesamt weniger Erklärungskraft für den Bildungserfolg von Schülern haben als Merkmale des Unterrichts und der Lehrperson (Hattie, 2009). Demzufolge ist es vorstellbar, dass die Unterrichtsqualität innerhalb der beiden Systeme breiter streut als zwischen den Systemen. Weitere Analysen aus dem Längsschnittprojekt, die auch Videodaten aus dem Unterricht integrieren können, werden Aufschluss darüber geben, ob sich die Unterrichtsqualität innerhalb der Systeme deutlicher unterscheidet als zwischen den beiden Systemen.

Eltern dürften ihre Entscheidung, auf welche Schule sie ihr Kind schicken, eher am sogenannten Typ A-Effekt und nicht am Typ B-Effekt ausrichten. Der Typ A-Effekt gibt an, welche Leistungsunterschiede zwischen Schülern bestehen, die die gleichen individuellen Lernvoraussetzungen haben, aber unterschiedliche Schulen besuchen. Der Typ B-Effekt bezeichnet dagegen die Differenz zwischen den Leistungen von Schülern, die vergleichbare individuelle Voraussetzungen haben und auch Schulen mit vergleichbaren Kontextbedingungen besuchen (Ditton, 2009). In dieser Studie konnte ein Typ A-Effekt für Lesen, Mathematik und Naturwissenschaften ermittelt werden, während ein Typ B-Effekt nur für Mathematik nachzuweisen war. Eltern dürften gleichwohl ein vergleichsweise geringes Interesse an der Gültigkeit des Typ B-Effekts haben, also daran, ob die besseren Leistungen ihrer Kinder durch Kontextmerkmale wie die Klassenzusammensetzung oder durch besonders gut ausgeprägte Schul- und Unterrichtsmerkmale bestimmt sind. Relevanter dürfte für sie der Typ A-Effekt sein, also die Frage, auf welcher Schule ihr Kind die besseren Leistungen erbringt, unabhängig davon, welche Ursache hierfür verantwortlich ist. Aus bildungspolitischer Sicht sind solche elterlichen Präferenzen und Entscheidungskriterien nicht unproblematisch, denn in der Konsequenz leisten sie einer Homogenisierung und Segregation der Schülerschaft bereits im Grundschulalter Vorschub. Einen fairen Vergleich bietet in jedem Fall die Ermittlung des Typ B-Effekts, der deshalb für die Bildungsforschung und die Bildungspolitik die zentrale Referenzgröße darstellt.

In diesem Zusammenhang gilt es noch einmal auf die Grenzen dieser Untersuchung hinzuweisen: Es handelt sich um eine Querschnittstudie, die die *Entwicklungen* der beiden Schülergruppen nicht nachzeichnen kann. Hinzu kommt, dass die BIP-Schüler in dieser Studie 2005 und 2006 getestet und mit IGLU-Schülern verglichen wurden, deren Leistungen bereits 2001 erfasst wurden. Kohorteneffekte und damit einhergehende Verzerrungen lassen sich daher nicht ganz ausschließen.

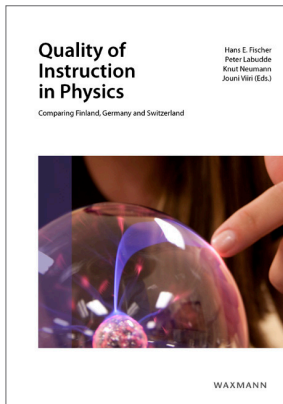
Aus der Sicht der künftigen Forschung ist interessant, inwieweit sich die Ergebnisse dieser Querschnittstudie durch die Befunde der Längsschnittstudie über das Ende des 2. Schuljahres hinaus bestätigen lassen.

Literatur

- Baumert, J., Becker, M., Neumann, M. & Nikolova, R. (2009). Frühübergang in ein grundständiges Gymnasium – Übergang in ein privilegiertes Entwicklungsmilieu? *Zeitschrift für Erziehungswissenschaft*, 12(2), 189–215.
- Beckerman, T. & Good, T. (1981). The classroom ratio of high- and low-aptitude students and its effect on achievement. *American Educational Research Journal*, 18(3), 317–327.
- Bellin, N. (2009). *Klassenkomposition, Migrationshintergrund und Leistung. Mehrebenenanalysen zum Sprach- und Leseverständnis von Grundschulern*. Wiesbaden: VS.
- Bos, W., Valtin, R., Hornberg, S., Buddeberg, I., Goy, M. & Voss, A. (2007). Internationaler Vergleich 2006: Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 109–160). Münster: Waxmann.
- Bosetti, L. (2004). Determinants of school choice. Understandig how parents choose elementary schools in Alberta. *Journal of Educational Policy*, 19(4), 387–405.
- Caliendo, M. & Kopeinig, S. (2008). Some practical guidance for the implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22(1), 31–72.
- Chubb, J. E. & Moe, T. M. (1990). *Politics, markets & America's schools*. Washington DC: The Brookings Institution.
- Coleman, J. S., Hoffer, T. & Kilgore, S. (1982). *High school achievement: Public, catholic, and private schools compared*. New York: Basic Books.
- Dar, Y. & Resh, N. (1986). Classroom intellectual composition and academic achievement. *American Educational Research Journal*, 23, 357–374.
- Ditton, H. (2009). Familie und Schule – eine Bestandsaufnahme der bildungssoziologischen Schuleffektforschung von James S. Coleman bis heute. In R. Becker (Hrsg.), *Lehrbuch der Bildungssoziologie* (S. 237–256). Wiesbaden: VS.
- Dronkers, J. & Avram, S. (2009). Choice and effectiveness of private and public schools. A new approach. *Zeitschrift für Pädagogik*, 55(6), 895–909.
- Dronkers, J. & Robert, P. (2003). *Effectiveness of public and private schools in a comparative perspective*. San Domenico: EUI Working Paper SPS No. 2003/13. Zugriff am 14.02.2014 unter <http://cadmus.eui.eu/dspace/bitstream/1814/1360/1/sps2003-13.pdf>
- Dronkers, J. & Robert, P. (2008a). School choice in the light of the effectiveness differences of various types of public and private school in 19 OECD countries. *Journal of School Choice*, 2(3), 260–301.
- Dronkers, J. & Robert, P. (2008b). Differences in scholastic achievement of public, private government-dependent and private independent schools. A cross-national analysis. *Educational Policy*, 22(4), 541–577.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Heise, E., Böhme, E. & Körner, S. (2010). Montessori-orientierter und traditioneller Grundschulunterricht: Ein Vergleich der Entwicklung von Rechtschreibung,

- Rechnen, Intelligenz und Kreativität. *Psychologie in Erziehung und Unterricht*, 57, 273–289.
- Ho, D. E., Imai, K., King, G. & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1–28.
- Karst, K. & Lipowsky, F. (2013). Leistungsentwicklung in Mathematik und deren Determinanten. In F. Lipowsky, G. Faust & C. Kastens (Hrsg.), *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren* (S. 51–75). Münster: Waxmann.
- Leschinsky, A. (2003). Der institutionelle Rahmen des Bildungswesens. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick* (S. 148–213). Reinbek b. Hamburg: Rowohlt.
- Lipowsky, F., Faust, G. & Kastens, C. (Hrsg.). (2013). *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren*. Münster: Waxmann.
- Lotz, M., Schoreit, E. & Kempfer, I. (2013). Die Entwicklung des Leseverständnisses und der Rechtschreibfähigkeit. In F. Lipowsky, G. Faust & C. Kastens (Hrsg.), *Persönlichkeits- und Lernentwicklung an staatlichen und privaten Grundschulen. Ergebnisse der PERLE-Studie zu den ersten beiden Schuljahren* (S. 77–97). Münster: Waxmann.
- Lubienski, C., Weitzel, P., Lubienski, S. T. (2009). Is there a “consensus” on school choice and achievement? *Educational Policy*, 23(1), 161–193.
- Lubienski, S. T. & Lubienski, C. (2006). School sector and academic achievement: A multilevel analysis of NAEP mathematics data. *American Educational Research Journal*, 43(4), 651–698.
- Lubienski, S. T., Lubienski, C. & Crane, C. C. (2008). Achievement differences and school type: The role of school climate, teacher certification, and instruction. *American Journal of Education*, 115(1), 97–138.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung: Probleme und Lösungen. *Psychologische Rundschau*, 58, 103–117.
- Martin, M. O., Mullis, I. V. S., Beaton, A. E., Gonzalez, E. J., Smith, T. A. & Kelly, D. L. (1997). *Science achievement in the primary school years: IEA’s Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill: Boston College.
- McEwan, P. (2000). The potential impact of large-scale voucher programs. *Review of Educational Research*, 70(2), 103–149.
- Mehlhorn, H.-G. (2001). Begabungsentwicklung durch Kreativitätsförderung in Kindergarten und Schule. *Materialien Forum Bildung*, 7, 22–44. Zugriff am 14.02.2014 unter <http://www.blk-bonn.de/papers/forum-bildung/band07.pdf>
- Mehlhorn, G. & Mehlhorn, H.-G. (2003). Kreativitätspädagogik – Entwicklung eines Konzepts in Theorie und Praxis. *Bildung und Erziehung*, 56(1), 23–45.
- Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L. & Smith, T. A. (1997). *Mathematics achievement in the primary school years: IEA’s Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill: Boston College.
- OECD – Organisation for economic Co-operation and Development. (2004). *Lernen für die Welt von Morgen: Erste Ergebnisse von PISA 2003*. Paris: OECD.
- OECD – Organisation for economic Co-operation and Development. (2007). *PISA 2006 – Schulleistungen im internationalen Vergleich. Naturwissenschaftliche Kompetenzen für die Welt von morgen*. Bielefeld: Bertelsmann.
- Opdenakker, M. C. & Van Damme, J. (2006). Differences between secondary schools: A study about school context, group composition, school practice, and school effects

- with special attention to public and catholic schools and types of schools. *School Effectiveness and School Improvement*, 17(1), 87–117.
- Prenzel, M., Geiser, H., Langeheine, R. & Lobemeier, K. (2003). Das naturwissenschaftliche Verständnis am Ende der Grundschule. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 143–187). Münster: Waxmann.
- R Development Core Team. (2011). *R: A language and environment for statistical computing*. Wien: R Foundation for Statistical Computing. Verfügbar unter <http://www.R-project.org>
- Rindermann, H. (2007). Die Bedeutung der mittleren Klassenfähigkeit für das Unterrichtsgeschehen und die Entwicklung individueller Fähigkeiten. *Unterrichtswissenschaft*, 35(1), 68–89.
- Rosenbaum, P. & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–50.
- Rosenbaum, P. & Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39, 33–38.
- Schöppe, K. (2010). Begabungen der Kinder frühzeitig und langfristig entwickeln. *Neue Praxis der Schulleitung*, (100), F 5.5, 1–20. Zugriff am 14.02.2014 unter <http://www.bip-mehlhornschoolen.de/fileadmin/uploads/pdf/Artikel-Raabe-Verlag.pdf>
- Statistisches Bundesamt. (2012). *Bildung und Kultur. Private Schulen. Schuljahr 2011/2012*. Fachserie 11, Reihe 1.1. Wiesbaden.
- Tiedemann, J. & Billmann-Mahecha, E. (2004). Kontextfaktoren der Schulleistung im Grundschulalter. Ergebnisse aus der Hannoverschen Grundschulstudie. *Zeitschrift für Pädagogische Psychologie*, 18(2), 113–124.
- Van Buuren, S. & Groothuis-Oudshoorn, K. (2009). *MICE: Multivariate imputations by chained equations in R*. Zugriff am 14.02.2014 unter <http://www.stefvanbuuren.nl/publications/MICE%20in%20R%20-%20Draft.pdf>
- Van Damme, J., De Fraine, B., Van Landeghem, G., Opendakker, M.-C. & Onghena, P. (2002). A new study on educational effectiveness in secondary schools in Flanders. An introduction. *School Effectiveness and School Improvement*, 13(4), 383–397.
- Voss, A. (2006). *Print- und Hypertextlesekompetenz im Vergleich. Eine Untersuchung von Leistungsdaten aus der Internationalen Grundschul-Lese-Untersuchung (IGLU) und der Ergänzungsstudie Lesen am Computer (LaC)*. Münster: Waxmann.
- Walther, G., Geiser, H., Langeheine, R. & Lobemeier, K. (2003). Mathematische Kompetenzen am Ende der vierten Jahrgangsstufe. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 189–226). Münster: Waxmann
- Weishaupt, H., Baethge, M., Füssel, H.-P., Hetmeier, H.-W., Rauschenbach, T., Rockmann, U., Seeber, S. & Wolter, A. (Hrsg.). (2012). *Bildung in Deutschland 2012. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: Bertelsmann.
- Weiß, M. (2011). *Allgemeinbildende Privatschulen in Deutschland. Bereicherung oder Gefährdung des öffentlichen Schulwesens?* Berlin: Friedrich-Ebert-Stiftung. Zugriff am 20.07.2012 unter <http://library.fes.de/pdf-files/studienfoerderung/07833.pdf>
- Wenglinsky, H. (2007). Are private high schools better academically than public high schools? Center on Education Policy (CEP). Zugriff am 14.02.2014 unter <http://www.edline.com/uploads/pdf/PrivateSchoolsReport.pdf>



Hans E. Fischer, Peter Labudde,
Knut Neumann, Jouni Viiri (Eds.)

Quality of Instruction in Physics Comparing Finland, Germany and Switzerland

2014, 216 pages, pb., € 29,90

ISBN 978-3-8309-3055-6

E-Book: € 26,99

This book reports the findings from the tri-national video study Quality of Instruction in Physics (QuIP). Within the scope of the QuIP study, physics instruction was investigated in a total of 103 classes from Finland, North Rhine-Westphalia (Germany) and German-speaking Switzerland. The main aim was to identify typical patterns of physics instruction of the three samples and to investigate conditions under which these patterns are successful with respect to students' learning, interest and motivation.



WAXMANN