

*Anna Makles/Horst Weishaupt*

## **Sozialindex für Schulen – Möglichkeiten und Probleme der Konstruktion am Beispiel einer Untersuchung in Nordrhein-Westfalen**

### **1** Einleitung

In den letzten Jahren haben mehrere Entwicklungen im Schulwesen dazu geführt, der sozialen und migrationsspezifischen Zusammensetzung der Schüler einer Schule besondere Aufmerksamkeit zu schenken. Sicher waren die Ergebnisse der internationalen Leistungsstudien zu der vergleichsweise sehr hohen sozialen Selektivität des deutschen Schulsystems ein Auslöser für dieses Interesse. Um die soziale Situation der Schulen zu erfassen, mussten neue Überlegungen angestellt werden, denn die Schulstatistik sieht die Erhebung des Bildungsniveaus und des beruflichen Status der Eltern der Schüler nicht vor.

Es entstanden deshalb in den letzten Jahren mehrere Ansätze, einen „Sozialindex“ für Schulen zu konstruieren. Damit sind drei Zielsetzungen verbunden:

- (1) Es sollen damit standortspezifisch unterschiedliche Ressourcenzuweisungen (insbesondere für Grund- und Hauptschulen) begründet werden können, damit durch die soziale und ethnische Zusammensetzung der Schüler bedingte Standortnachteile von Schulen ausgeglichen werden können.
- (2) Im Zusammenhang der landesweiten Lernstandserhebungen entstand ein Interesse daran, über einen Sozialindex für einzelne Schulen Standortunterschiede zu erfassen und dadurch faire Vergleiche zwischen den Schulen zu ermöglichen.
- (3) Im Zusammenhang der Einführung von Schulinspektionen wünschen die Inspektoren Informationen, um die Rahmenbedingungen der Arbeit an den Schulen besser einschätzen zu können.

Die Intentionen zur Entwicklung eines „Sozialindex“ stehen folglich in engem Zusammenhang mit Ansätzen neuer Steuerung der Schulen über autonome Mittelverwaltung, zentrale Leistungsüberprüfungen, Bildungsstandards und Maßnahmen der Qualitätssicherung.

In diesem Beitrag sollen zunächst Einwände gegen die bisherigen Ansätze zur Konstruktion von Sozialindizes dargestellt werden, dann der Ansatz und Teile von Ergebnissen einer eigenen Untersuchung<sup>1</sup> für zwei Regionen in Nordrhein-Westfalen vorgestellt und Überlegungen zur Weiterentwicklung dieser Forschungsbemühungen angestellt werden.

1 Dieser Beitrag beinhaltet Teilergebnisse der Projektgruppe ‚Schulindex‘ von *A. Makles* und *H. Weishaupt* des Verbund-Projektes „Prüfung der Leistungsfähigkeit und Verbesserung der Struktur amtlicher Daten im Rahmen der Konstruktion von schulbezogenen Sozialindizes“. Dieses Projekt wurde vom Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (LDS NRW) – Herr PD Dr. Jörg-Peter Schräpler/Herr Kurt Imhäuser koordiniert. Universitäre Kooperationspartner waren die Technische Universität Dortmund – Institut für Schulentwicklungsforschung (IFS) –

## 2 Ansätze für einen Sozialindex für Schulen

Ein von der Stadt Hamburg gewähltes aufwändiges Verfahren zur Ermittlung eines „Sozialindex“ von Schulen ist die Befragung von Eltern und Schülern.<sup>2</sup> Diese Erhebung wurde inzwischen erneut durchgeführt. Auch in den von uns untersuchten Regionen in Nordrhein-Westfalen, Stadt Dortmund und Kreis Coesfeld, bilden entsprechende Untersuchungen den Bezugspunkt der eigenen Analysen. Die Elternbefragung erscheint zunächst als „Königsweg“ zur Ermittlung eines Sozialindex von Schulen. Zu berücksichtigen ist jedoch, dass durch Antwortausfälle die Ergebnisse verfälscht werden können. Das Land Hessen ist nach anfänglichen Bestrebungen, dieses Verfahren auch an hessischen Schulen einzusetzen, davon abgekommen, weil befürchtet wird, dass Eltern strategisch antworten, um dadurch Vorteile für die Schule zu erreichen. Deshalb bemüht sich Hessen inzwischen um ein nichtreaktives Verfahren zur Erfassung der sozialen Situation der Schulen.

Im Rahmen der Lernstandserhebungen in Nordrhein-Westfalen wurden die Schulleiter gebeten, selbst die soziale Lage der Schulen anhand einiger Vorgaben zu klassifizieren. Hier zeigte sich bereits in der Tendenz strategisches Verhalten, denn die Angaben der Schulleiter wichen nicht nur häufig von den Ergebnissen der Elternbefragungen ab: sie bewerteten die soziale Situation an den Schulen häufig schlechter, als sie sich als Ergebnis der Elternbefragung darstellte.

Von besonderem Interesse für unser eigenes Vorhaben waren Ansätze zur Entwicklung eines Sozialindex anhand von Daten der amtlichen Statistik. Der Kanton Zürich berechnete einen Index auf der Grundlage der Daten für Gemeinden und Gemeindeteile.<sup>3</sup> Für Nordrhein-Westfalen wurde für eine regional differenzierte Personalzuweisung ein Index auf Ebene der Schulaufsichtsbezirke (Kreise und kreisfreie Städte) entwickelt: „Mit dem Sozialindex für Nordrhein-Westfalen können die Schülerzahlen gewichtet werden, um Ressourcen der sozialen Belastung entsprechend disproportional zu den Schülerzahlen zu verteilen. [...] Der Konstruktion des Sozialindex liegt die Annahme zu Grunde, dass Kreise mit vielen Arbeitslosen, Sozialhilfeempfängern und Migranten sowie einem geringen Anteil von Wohnungen in Einfamilienhäusern sozial stärker belastet sind, als Kreise mit einer einheimischen, erwerbstätigen, den Lebensunterhalt selbständig bestreitenden und Einfamilienhäuser bewohnenden Bevölkerung.“<sup>4</sup> Wie die Zuweisungen nach dem Sozialindex an die Kreise und Städte auf die einzelnen Schulen verteilt werden, ist in Nordrhein-Westfalen der Schulaufsicht überlassen. Für ein einheitliches Vorgehen fehlt bisher ein geeigneter standortbezogener Index.

Die für den Sozialindex in Nordrhein-Westfalen herangezogene Kombination von einem hohen Anteil von Arbeitslosen, Sozialhilfeempfängern, Ausländern und einem niedrigen Anteil von

---

Prof. Dr. Wilfried Bos und die Ruhr-Universität Bochum – ZEFIR – Zentrum für interdisziplinäre Regionalforschung – Prof. Dr. Klaus Peter Strohmeier. Gefördert wurde das Vorhaben durch das Bundesministerium für Bildung und Forschung (BMBF). Der Bericht wurde im Februar 2010 in der Reihe Bildungsforschung, Band 31, unter dem Titel „Zur Konstruktion von Sozialindizes – Ein Beitrag zur Analyse sozialräumlicher Benachteiligungen von Schulen als Voraussetzung für qualitative Schulentwicklung“ vom BMBF veröffentlicht.

- 2 Vgl. Bos, W./Pietsch, M./Gröhlich, C./Janke, N., Ein Belastungsindex für Schulen als Grundlage der Ressourcenzuweisung am Beispiel von KESS 4, in: Bos, W./Holtappels, H.G./Pfeiffer, H./Rolf, H.-G., Schulz-Zander, R. (Hrsg.), Jahrbuch der Schulentwicklung, Bd. 14. Juventa, Weinheim 2006.
- 3 Vgl. Bildungsdirektion des Kantons Zürich, Sozialindex: Theoretischer Hintergrund und Beschreibung der Berechnung 2004.
- 4 Frein, T./Möller, G./Petermann, A./Wilpricht, M., Bedarfsgerechte Stellenzuweisung: Das neue Instrument Sozialindex, in: Schulverwaltung, Ausgabe Nordrhein-Westfalen, Nr. 6/2006, S. 188.

Personen in Einfamilienhäusern<sup>5</sup> erscheint vor dem Hintergrund der stadtsoziologischen Forschung nach Shevky & Bell<sup>6</sup> plausibel. Relativ aktuelle Daten für die Hauptschulbezirke der Stadt Hannover zeigen auch Korrelationen mit der Gymnasialbesuchsquote<sup>7</sup> (als einem guten Sozialstrukturindikator, der in Hannover durch eine Sondererhebung wohnortbezogen vorlag), die den erwarteten Zusammenhängen entsprechen (an Stelle des Anteils der Personen in Einfamilienhäusern wurde als äquivalenter Indikator die Wohnfläche je Person verwendet) (vgl. Tabelle 1).

Tabelle 1: Korrelationen sozialstatistischer Merkmale mit der Gymnasialbesuchsquote in den Hauptschuleinzugsgebieten von Hannover

Variable	Korrelation
Arbeitslosenquote	-0,90
Anteil Sozialhilfeempfänger	-0,89
Anteil Ausländer an 6- bis 14-Jährigen	-0,81
Anteil Türken an Ausländern	-0,73
Erwerbstätigkeit Frauen	0,41
Wohnfläche je Einwohner	0,44
Wanderungssaldo	0,55
Erwerbstätigkeit	0,59
Wahlbeteiligung	0,81

*Quelle:* Weishaupt, H., Schulzeck, U. unter Mitarbeit von Borde, M. (2005): *Regionalisierte Stichprobenziehungen und Analysen bei large scale assessments – Möglichkeiten zur Typisierung regionaler Lebenslagen mit Daten der amtlichen Statistik*. Ergebnisbericht an das Bundesministerium für Bildung und Forschung, Wuppertal (unveröffentlichtes Manuskript), S. 31.

Dennoch ist die Wahl der Indikatoren für den Sozialindex in Nordrhein-Westfalen nicht unproblematisch. Denn die Betrachtung der Korrelationen von sozialstatistischen Merkmalen der Kreise und kreisfreien Städte in Deutschland mit der gymnasialen Bildungsbeteiligung führt zu abweichenden Befunden (vgl. Tabelle 2).

Besonders auffällig ist, dass im Kreisvergleich der Anteil der Ausländer positiv mit dem Gymnasialbesuch korreliert.<sup>8</sup> Offensichtlich können statistische Merkmale im innerstädtischen Vergleich etwas anderes indizieren als im interregionalen.

- 
- 5 Der Kanton Zürich verwendet für den Sozialindex den Anteil der Arbeitslosen, der Ausländer, der Wohnungen in Einfamilienhäusern und der Personen, die in den letzten fünf Jahren nicht umgezogen sind (Sesshaftigkeitsquote), vgl. Bildungsdirektion des Kantons Zürich (2004).
  - 6 Shevky, E./Bell, W., Sozialraumanalyse, in: Atteslander, P./Hamm, B. (Hrsg.), Materialien zur Siedlungssoziologie. Köln 1974.
  - 7 Weishaupt, H./Schulzeck, U./Borde, M., Regionalisierte Stichprobenziehungen und Analysen bei large scale assessments – Möglichkeiten zur Typisierung regionaler Lebenslagen mit Daten der amtlichen Statistik. Ergebnisbericht an das Bundesministerium für Bildung und Forschung, Wuppertal (unveröffentlichtes Manuskript), 2005.
  - 8 Bezogen auf die 54 Kreise in Nordrhein-Westfalen korreliert nur der Anteil der Arbeitslosen schwach (-.30) mit der Gymnasialbesuchsquote. Der Ausländeranteil (.08) und der Anteil der Ein- und Zweifamilienhäuser (-.01) sind unkorreliert. Auch wenn die Einkünfte der Steuerpflichtigen als Sozialstrukturindikator gewählt werden, ändert sich nichts an diesem Ergebnis.

Tabelle 2: Korrelationen sozialstatistischer Merkmale mit dem Gymnasialbesuch für Kreise und kreisfreie Städte in Deutschland

Variablen	alte Länder	neue Länder
Anteil 6- bis 18-Jähr.	-0,67	-0,51
Personen je Haushalt	-0,62	-0,53
Zusammengefasste Geburtenziffer	-0,52	-0,26
Anteil Wohngebäude mit 1 und 2 Wohnungen	-0,51	-0,54
Anteil Arbeiter an SV-Beschäftigten	-0,51	-0,43
Familienwanderungssaldo (unter 18, 30- bis 50-Jähr.)	-0,15	-0,52
PKW je 1000 Einwohner	-0,08	-0,54
Arbeitslosenquote	0,15	-0,05
Ausländeranteil	0,46	0,40
Umsatzsteuer je Einwohner	0,52	0,47

Quelle: INKAR 2003, eigene Berechnung

Im Vergleich zwischen Regionen überlagern Stadt-Land-Unterschiede die sozialen Disparitäten. Auch zeigen sich gelegentlich noch Unterschiede zwischen Ost- und Westdeutschland (z. B. bei der Frauenerwerbsquote), die auf regionale Bedeutungsunterschiede hinweisen (eine hohe Frauenerwerbsquote ist in Westdeutschland vermutlich ein Hinweis auf eine gut qualifizierte Bevölkerung in einer Dienstleistungsregion, in Ostdeutschland zeigt sie eine gute Beschäftigungssituation an). Die für den Sozialindex in Nordrhein-Westfalen ausgewählte Datenbasis ist folglich theoretisch kaum begründbar, denn es ist fraglich, ob es sich um sinnvolle Sozialstrukturindikatoren für einen regionalen Vergleich handelt. Die Datenbasis bezieht sich im Übrigen nicht auf Schüler, sondern die Wohnbevölkerung. Insofern stellt sich die Frage, weshalb nicht die Schulstatistik herangezogen wurde, um – z. B. anhand des Ausländeranteils, der Übergangsquote zum Gymnasium (zur Hauptschule), der Überweisungsquote auf die Förderschule mit dem Förderschwerpunkt Lernen – die soziale Belastung von Schulen zu erfassen. Daher sind die Ansätze zu einem Sozialraumindex für Schulen, die sich auf Daten der Bevölkerungs- und Arbeitsmarktstatistik stützen, kritisch einzuschätzen.

Dennoch sind sozialraumanalytische Ansätze, die sich auf das sozioökonomische Umfeld von Bildungseinrichtungen beziehen, für die Interpretation von Bildungsdisparitäten unverzichtbar. Dies wird von der Studie von Baumert, Carstensen und Siegle (2005) unterstrichen, die erstmals belegen konnte, dass etwa 3 % der Leistungsvariation in Mathematik zwischen Schulen auf wirtschaftliche und soziale Kontextbedingungen zurückgeführt werden können, die deshalb praktisch und politisch nicht unbedeutend sind.<sup>9</sup> In spezifischer Weise als leistungsvariant erwiesen sich in ihrem Untersuchungsansatz: die regionale Arbeitslosenquote, die Quote der Sozialhilfeempfänger und der Anteil an Schulabgängern mit Hochschulreife. Die Autoren müssen aber einräumen: „Über welche Vermittlungsprozesse sich diese Regionalstrukturmerkmale auf den individuellen Kompetenzerwerb in Schulen auswirken, ist offen und ungeklärt.“<sup>10</sup>

9 Vgl. Baumert, J./Carstensen, C.H./Siegle, T., Wirtschaftliche, soziale und kulturelle Lebensverhältnisse und regionale Disparitäten des Kompetenzerwerbs, in: Prenzel, M./Baumert, J./Blum, W./Lehmann, R./Leutner, D./Neubrand, M. et al. (Hrsg.), PISA 2003: Der zweite Vergleich der Länder in Deutschland – Was wissen und können Jugendliche?, S. 323–365, Münster 2005.

10 Ebd., S. 360.

Auch in dieser Studie ist allerdings die verwendete Datenbasis nicht unproblematisch, nicht wegen der gewählten Indikatoren, sondern weil keine Informationen über die Einzugsbereiche der einzelnen Schulen verwendet wurden, und sich die sozioökonomischen Daten auf die Kreise und kreisfreien Städte beziehen, in denen die Schulen gelegen sind.

Diese Beispiele sollen verdeutlichen, dass die Entwicklung eines Sozialindex für Schulen einen erheblichen Entwicklungsaufwand erfordert und manche Fallstricke dieser Aufgabe zunächst übersehen wurden. Unter Berücksichtigung dieser Erfahrungen versuchen wir mit unserem Vorhaben einen Schritt bei der Konstruktion eines schulbezogenen Sozialindex voranzukommen.

### 3 Fragestellung und Konzept der Untersuchung

Gegenstand des Verbund-Projektes war die Untersuchung der Möglichkeit, regelmäßig anfallende amtliche Daten mit dem Ziel aufzubereiten, forschungsökonomisch praktikable und valide schulspezifische Sozialindizes zu entwickeln. Dazu wurden amtliche kleinräumige Daten mit Daten der amtlichen Schulstatistiken und Sekundärdaten verknüpft und in Form eines konsistenten Datensatzes aufbereitet. Auf Basis verschiedener Datenquellen und Verfahren wurden dann beispielhaft Sozialindizes für Schulen in Dortmund und dem Kreis Coesfeld berechnet und mit entsprechenden Ergebnissen der durch das IFS Dortmund durchgeführten Befragung der Eltern an den Schulen in den beiden Regionen evaluiert. In Dortmund und Coesfeld wurden hierzu alle Grundschulen und weiterführenden Schulen (ohne Förderschulen) in die Untersuchung des IFS einbezogen. Insgesamt nahmen in Dortmund 92 Grundschulen und 54 weiterführende Schulen und im Kreis Coesfeld 43 Grundschulen und 29 Schulen der Sekundarstufe I teil.

Der Schülerhintergrund wurde in Anknüpfung an aktuelle Modelle sozialer Ungleichheit über einen ressourcen-orientierten Kapitalansatz der Reproduktion sozialer Ungleichheit erfasst.<sup>11</sup> Dieser Ansatz wurde durch Daten zum Migrationshintergrund der Schülerinnen und Schüler ergänzt. Zur Berechnung des Index zur sozialen Belastung von Schulen wurden die Einzeldaten über die soziale Lage von Schülerinnen und Schülern auf Schulebene aggregiert. Der so entstandene Index wurde zur Evaluation unserer Ergebnisse herangezogen. Daneben konnten durch ein Entgegenkommen des nordrhein-westfälischen Kultusministeriums zusätzlich auch Ergebnisse der Lernstandserhebungen der Schulen herangezogen werden.

Derzeit sind die Möglichkeiten, mit Hilfe der Daten der amtlichen Statistik Indizes auf kleinräumiger Basis zu konstruieren, sehr beschränkt. Die zu diesem Zweck nutzbaren, vorhandenen amtlichen Daten beziehen sich in der Regel auf kleinräumige Gebiete, die den Einzugsbezirken bzw. den Sozialräumen der Schulen nicht entsprechen. Zudem sind viele schulrelevante Informationen in unterschiedlichen Statistiken verteilt oder nicht frei verfügbar.

Ziel des Projektes war es daher, die Leistungsfähigkeit der amtlichen Statistik für die Konstruktion eines Sozialindex von Schulen zu prüfen. Insofern handelt es sich bei den hier präsentierten Ergebnissen um Teilergebnisse einer Machbarkeitsstudie, in der vor allem verschiedene methodische Verfahren mit der Perspektive gegeneinander getestet wurden, die entwickelte Datenstruktur später auch für einen bundesweiten Schulindex nutzen zu können.

11 Vgl. Bourdieu, P., Ökonomisches Kapital, kulturelles Kapital, soziales Kapital, in: Kreckel, R. (Hrsg.), Soziale Ungleichheiten, Sonderband 2 der Zeitschrift 'Soziale Welt', S. 183–198, Göttingen 1983.

#### 4 Anforderungen an einen Sozialindex für Schulen auf Basis von amtlichen Daten sowie kommerziellen Sekundärdaten

Der Sozialindex für Schulen soll analog zum Kreis-Sozialindex<sup>12</sup> die „soziale Belastung“ von Schulen messen und die soziale Situation vor Ort widerspiegeln. Eine wesentliche Voraussetzung für die Konstruktion eines derartigen Index ist die möglichst genaue Kenntnis der Zusammensetzung der Schülerschaft der Schulen, des Schulumfelds sowie die Berücksichtigung der individuellen Schulausstattung. Insofern bietet sich hier zunächst eine konzeptionelle Trennung in Schülerinformation, Sozialrauminformation und Schulinformation an.

- *Schülerinformation:* Von Interesse ist unter dieser Perspektive vor allem die spezifische Zusammensetzung der Schülerschaft innerhalb einer Schule. Die amtliche Schulstatistik erfasst aber nicht das Bildungsniveau und den Beruf der Eltern der Schüler. Allerdings liegen auf der Schulebene aggregierte Schülerangaben z. B. zum Aussiedler- und Ausländeranteil an der Schule vor (siehe Schulinformation), die zur Konstruktion des Index verwendet werden können.
- *Schulinformation:* Die Schulinformation charakterisiert die Schule und berücksichtigt die heterogene Verteilung der Ausstattung mit Personal, Technik und Räumen sowie schülerspezifische Indikatoren. Hierzu zählen z. B. der Ausländeranteil an der Schule, Nationalitäten, die Schüler/Lehrer-Relation, Schüler/Klasse-Relation, Schüler- und Klassenfrequenzen, tatsächlich erteilte Unterrichtsstunden, Altersstruktur der Lehrer, Abgängerzahlen, Übergängerzahlen, Schultypen (z. B. Ganztagschule), Förderangebote der Schulen für bestimmte Zielgruppen und Schulträger. Diese Informationen liegen in der amtlichen Schulstatistik vor und können zur Erstellung von Schulprofilen genutzt werden.
- *Sozialrauminformation:* Eine erweiterte PISA-Auswertung bestätigte, dass neben den familiären Merkmalen auch die wirtschaftlichen, sozialen und kulturellen Rahmenbedingungen der jeweiligen Schulstandorte einen Einfluss auf die Kompetenzen der getesteten Schüler haben. Die Ergebnisse wurden in dieser Studie auf Ebene der Kreise und kreisfreien Städte berechnet,<sup>13</sup> jedoch sind Städte und Kreise in sich oftmals hochgradig differenziert. Die gesellschaftliche Ungleichheit findet sich in einer kleinräumigen sozialen Fragmentierung der Wohnbevölkerung wieder – häufig liegt eine starke räumliche Trennung zwischen reichen („bildungsnahe“) und armen („bildungsferne“) Wohngebieten innerhalb einer Stadt vor. Zudem kann gezeigt werden, dass Bildungsentscheidungen der Eltern nicht nur von sozio-ökonomischen Faktoren wie Status und Beruf, sondern auch von der sozialen und wirtschaftlichen Situation in dem jeweiligen kleinräumigen Umfeld beeinflusst werden.<sup>14</sup> Für eine möglichst konkrete Bestimmung der „sozialen Belastung“ durch Merkmale des Sozialraums muss daher der Fokus schärfer auf das direkte Einzugsgebiet der Schule gerichtet werden.

Obwohl alle angesprochenen Dimensionen ausführlich im Pilotprojekt berücksichtigt und bearbeitet wurden, konzentriert sich dieser Beitrag auf die Ergebnisse, die bei der Bildung eines Sozialindex auf Basis der Schüler- und Schulinformation erzielt wurden.

<sup>12</sup> Vgl. Frein et al. (Anm. 4) und die dortige Verwendung der Begriffe *Sozialindex* und *soziale Belastung*.

<sup>13</sup> Vgl. Baumert et al. (Anm. 9).

<sup>14</sup> Vgl. z. B. Ditton, H., Schulwahlentscheidungen unter sozial-regionalen Bedingungen, in: Böhm-Kasper, O./Schuchart, C./Schulzeck, U. (Hrsg.), Kontexte von Bildung, Waxmann, S. 21–38, Münster 2007.

## 5 Bildung eines Schul-Sozialindex

Um die Leistungsfähigkeit der Daten der amtlichen Schulstatistik zur Konstruktion von schulbezogenen Sozialindizes zu beurteilen, stellte das LDS NRW (heute: IT.NRW) Datensätze für die Schulen der Regionen Dortmund und Coesfeld zur Verfügung. Die Datensätze betreffen das Schuljahr 2006/2007 und enthalten Informationen zu den Schülern, Klassen, Lehrern und Schulen. Der aus diesen Daten konstruierte ‚Schulindex‘ wurde dem Sozialindex des IFS (nachfolgend IFS-Index genannt) und Ergebnissen von Lernstandserhebungen gegenüber gestellt.

Ziel eines schulbezogenen Sozialindex ist unter anderem, Schulen vergleichbar zu machen und schulspezifische Charakteristika hervorzuheben bzw. zu berücksichtigen. Aus inhaltlichen und methodischen Überlegungen werden aus dem verfügbaren Datenbestand der amtlichen Schulstatistik für das Schuljahr 2006/2007 daher Indikatoren gebildet, die folgende Voraussetzungen erfüllen sollten:

- (1) Die Information, die zur Konstruktion eines Indikators dient, lässt sich sinnvoll auf die Ebene der Schule aggregieren. So lässt sich beispielsweise der Anteil der Schüler mit Migrationshintergrund für eine Schule darstellen, während das durchschnittliche Alter der Schüler keine aussagekräftige Information liefert.
- (2) Die Information ist in der Lage, Unterschiede und Gemeinsamkeiten zwischen den Schulen zu beschreiben bzw. aufzudecken. Dies beinhaltet auch die Forderung, dass die Informationen für alle Schulen gleichermaßen vorliegen müssen und vergleichbar sind. Nicht zuletzt führt dies auch dazu, dass gesonderte Analysen für Grund- und weiterführende Schulen durchgeführt werden müssen.
- (3) Die Information liefert möglicherweise einen Hinweis auf die hohe bzw. niedrige soziale Belastung der Schule. So kann z. B. festgehalten werden, dass der Anteil der weiblichen Lehrer keinen Hinweis auf die soziale Belastung einer Schule liefert, während der Anteil nicht versetzter Schüler ein Indikator für die soziale Belastung einer Schule sein könnte.

Zur Berechnung eines schulbezogenen Sozialindex für die Grundschulen und weiterführende Schulen wurden nach den oben dargestellten Voraussetzungen unter den zahlreich vorhandenen Informationen folgende Variablen hypothesengeleitet und explorativ ausgewählt:

- *Migrantenanteil*: Aus der Anzahl der ausländischen Schüler und der Anzahl der Aussiedler unter den Schülern wird die Summe der Schüler gebildet, die einen Migrationshintergrund<sup>15</sup> aufweisen. Durch Division durch die gesamte Anzahl der Schüler an der Schule wird die Variable Migrantenanteil (*MIGR*) gebildet.
- *Nicht versetzte Schüler*: Die Anzahl der Schüler, die nicht versetzt wurden,<sup>16</sup> wird hier verwendet, um den Anteil bzw. die Quote der nicht versetzten Schüler zu bestimmen (*NVER*). Dieser

<sup>15</sup> Diese Definition nach der nordrhein-westfälischen Schulstatistik, die durch die Erfassung der Aussiedler unter den Schülern ermöglicht wird, weicht von den gängigen Definitionen und der geplanten Erhebung des Migrationshintergrunds im Rahmen des schulstatistischen Kerndatensatzes der KMK ab. Danach soll der Geburtsort und die zu Hause gesprochene Sprache berücksichtigt werden. Diese Informationen lagen für NRW zum Zeitpunkt der Untersuchung noch nicht vor.

<sup>16</sup> Nichtversetzungen ohne freiwillige Zurückstufung.

ergibt sich aus der Summe aller Schüler, die nicht versetzt wurden, dividiert durch alle Schüler dieser Schule.

- *Übergängeranteil auf Gymnasien (nur Grundschulen)*: Entsprechend der am häufigsten vertretenen Schulformen der weiterführenden Schulen (Hauptschule, Realschule, Gesamtschule und Gymnasium) wird eine Übergangsquote der Grundschüler für die entsprechende Schulform berechnet. Die Übergängerquote auf das Gymnasium (*U20R*) wurde für die weitere Analyse als Repräsentant ausgewählt.
- *Abgängeranteil mit bestem Abschluss (nur weiterführende Schulen)*: In Abhängigkeit von den betrachteten Schulformen (Hauptschule, Realschule, Gesamtschule und Gymnasium) wird ein Abgängeranteil konstruiert. Er kennzeichnet den Anteil der Schüler, die den für diese Schulform maximal erreichbaren Abschluss erreicht haben (*BEST*). Dies vereinfacht die Auswahl eines Repräsentanten der 7 bzw. 8 möglichen Abschlüsse. Für Haupt- und Realschulen bedeutet dies, dass der Anteil der Schüler betrachtet wird, die die Fachoberschulreife inkl. Berechtigung zum Besuch der Jahrgangsstufe 11 der gymnasialen Oberstufe erreicht haben. Für Gesamtschulen und Gymnasien ist dies die Hochschulreife.
- *Förderschwerpunkte*: Auf Basis der vorliegenden Daten wird der Anteil der Schüler bestimmt, die in speziellen Schwerpunkten gefördert werden (*FOER*).<sup>17</sup>
- *Klassengrößen*: Um die Information über die Klassengrößen interschulisch vergleichbar zu machen, werden für die Grundschulen auf Basis der absoluten Schülerzahl und der absoluten Klassenzahl über alle Stufen hinweg die durchschnittlichen Klassengrößen (*DKLS*) der Regelklassen bestimmt. Die Beschränkung auf die Regelklassen erfolgt aus Gründen der Vergleichbarkeit. Für die weiterführenden Schulen werden auf Basis der absoluten Schülerzahl der Klassen 05 bis 10 und der absoluten Klassenzahl von Jahrgang 05 bis einschließlich 10 die durchschnittlichen Klassengrößen (*DKLS*) der Regelklassen bestimmt.
- *Schüler-Lehrer-Relation*: Das Schüler/Lehrer-Verhältnis wird über die Schüler pro Vollzeitlehrer-Quote (*SPVL*) der Schule bestimmt. Diese Kennzahl wird direkt durch das LDS NRW zur Verfügung gestellt.
- *Lehrbelastung*: Die zusätzliche ‚Belastung‘ der Lehrer wird durch den relativen Anteil der Unterrichtsstunden, die ohne Lehrbefähigung erteilt wurden, bestimmt (*LOBR*). Diese berechnet sich aus der Summe der ohne Lehrbefähigung erteilten Stunden dividiert durch die Summe aller erteilten Unterrichtsstunden.
- *Altersstruktur der Lehrer*: Die Altersverteilung der Lehrer wird anhand des durchschnittlichen Alters der Lehrer (*DALT*) ermittelt.

Die hier genannten Variablen wurden dazu verwendet, mit Hilfe einer Faktorenanalyse einen Schulindex zu konstruieren, der die soziale Belastung von Schulen messen bzw. abbilden soll. Die Ergebnisse dieser Schritte sowie der Faktorenanalyse werden nachfolgend für Grund- und weiterführende Schulen getrennt dargestellt und diskutiert.

17 Da nur sechs Schulen auch schwerstbehinderte Schüler fördern, werden Förderschwerpunkte für Behinderte bzw. geförderte behinderte Schüler aus Gründen der Vergleichbarkeit nicht betrachtet.

### 5.1 Schul-Sozialindex für Grundschulen

Die Eingangsgröße der Faktorenanalyse stellen die oben beschriebenen standardisierten Variablen dar. Die Kommunalitäten aller acht für die Grundschulen vorhandenen Variablen waren zu Beginn sehr gering, die größte Kommunalität zeigte sich bei der Variablen *MIGR* (Migrantenanteil) mit ( $h^2_{MIGR} = 0,333$ ), ein Ergebnis, welches insgesamt in Bezug auf das Ziel der Dimensionsreduktion und eine hohe Varianzklärung nicht zufriedenstellend ist. Das Maß der Stichprobeneignung, das Kaiser-Meyer-Olkin-Kriterium, war mit  $KMO = 0,626$  ebenfalls nur ‚mittelmäßig‘. Daneben trat das Problem auf, dass die Faktorextraktion nicht konvergierte,<sup>18</sup> was dazu führte, dass die Wahl der Variablen überdacht werden musste, um zu einer eindeutigen und interpretierbaren Lösung zu gelangen. Nach mehreren Versuchen, dieses Problem zu eliminieren, wurde daher eine neue Faktorenlösung mit nur vier Variablen, die die größte Kommunalität aufwiesen, bestimmt. Das Ergebnis dieser Faktorenlösung ist nachfolgend dargestellt, welches jedoch auf Grund des relativ schlechten Kaiser-Meyer-Olkin-Kriteriums ( $KMO = 0,644$ ) ebenfalls nicht zufriedenstellend ist.

Die Kommunalitäten sowie die Ladungen der Faktoren, die in Tabelle 3 dargestellt sind, verdeutlichen die eher wenig zufriedenstellende Faktorenlösung. Zwar wird ein latenter Faktor für diese vier Variablen extrahiert, jedoch ist der kumulierte Varianzklärungsanteil mit rund 31,7 % zu gering (vgl. letzte Spalte), als dass man davon sprechen kann, dass der Faktor die Variablen hinreichend abbildet. Die Kommunalitäten nach Extraktion (dritte Spalte) zeigen zudem, dass die Variablen *MIGR* und *U20R* den größten Beitrag zur Extraktion dieses Faktors liefern. Deutlich wird dies auch, wenn die Korrelationen zwischen den Variablen und dem Faktor betrachtet werden. Dort erreicht die Variable Migrantenanteil eine Korrelation von 0,85, während die Korrelationen der anderen Variablen geringer sind (Übergängerquoten auf Gymnasien: 0,49, Förderungsschwerpunkte: 0,34 und Schüler pro Vollzeitlehrer-Quote: 0,18). Letzteres wirft die Frage auf, ob der Erklärungsgehalt der beiden Variablen *MIGR* und *U20R* nicht für sich völlig ausreichend ist, so dass die Variablen *FOER* und *SPVL* entfernt werden können.

Tabelle 3: Faktorenlösung für Grundschulen

<i>Faktorenlösung</i>							
<i>Variable</i>	<i>Kommunalitäten</i>		<i>Faktorladungen</i>	<i>Eigenwerte (kum. %)</i>			
	<i>Anfang</i>	<i>Extraktion</i>		<i>Faktor</i>	<i>Anfang</i>	<i>Extraktion</i>	
<i>MIGR</i>	0,291	0,490	0,700	1	47,292	31,748	
<i>U20R</i>	0,277	0,453	-0,673	2	68,656		
<i>FOER</i>	0,143	0,203	0,451	3	87,890		
<i>SPVL</i>	0,094	0,124	-0,352	4	100,000		

Mit Hilfe einer Regression wurden in einem nächsten Schritt die Faktorwerte aus einer Linearkombination der Faktorladungen, die in der vierten Spalte von Tabelle 3 angegeben sind, und den Ausprägungen der Variablen berechnet. Diese Werte wurden, um einen direkten Vergleich mit dem IFS-Index zu erreichen, mit -1 vormultipliziert, so dass eine dem IFS-Index entsprechende Kodierung und Interpretation der Werte möglich ist. Der so konstruierte Index kann nun dazu verwendet werden, zunächst zu überprüfen, ob alle vier Variablen tatsächlich zwischen sozial belasteten oder unbelasteten Schulen diskriminieren können, insbesondere da die Kommunalitäten der

18 So genannter Heywood-Fall: Ein Schätzproblem, bei dem Kommunalitäten größer als 1 geschätzt werden. Einige Softwarepakete brechen die Faktorenlösung beim Auftreten dieses Problems ab (daher keine Konvergenz), andere setzen die Kommunalitäten automatisch kleiner 1. Gerade Letzteres verzerrt jedoch das endgültige Ergebnis und die Interpretation aller Kommunalitäten.

Variablen Förderanteil und Schüler pro Vollzeitlehrer-Quote sehr gering sind. Um dieser Frage nachzugehen, wurde der Index in Quintile unterteilt, so dass sich in jedem Quintil ca. 20 % der Indexwerte befinden. Weisen die Quintile deutliche Mittelwertunterschiede in den betrachteten Variablen auf und weisen diese zudem eine zunehmende oder abnehmende Tendenz auf, so kann davon ausgegangen werden, dass diese Variable zwischen einzelnen Bereichen der Indexwerte diskriminieren kann. Tabelle 4 zeigt diese klassifizierten Indexwerte und die Klassenmittelwerte der jeweiligen Variablen.

Tabelle 4: Schulindex-Quintile für Grundschulen

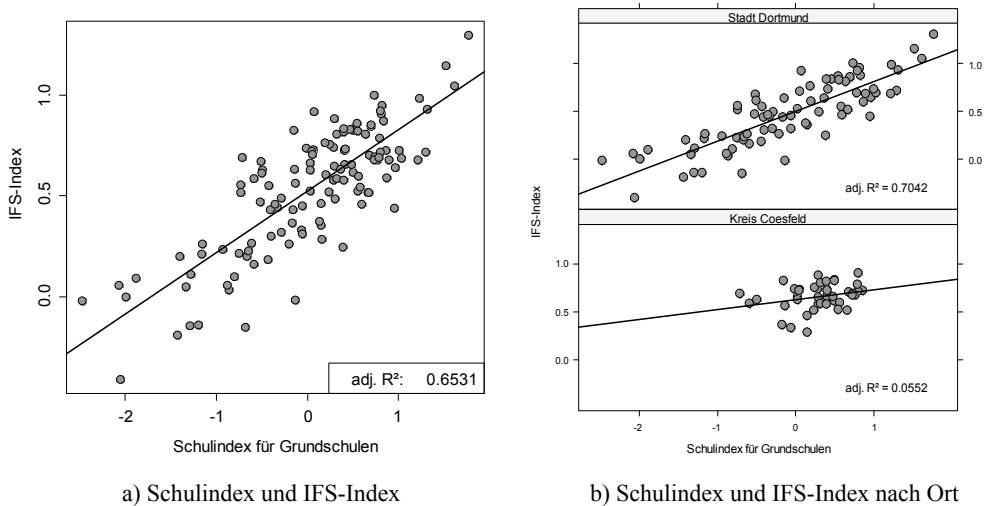
Schulindex (Quintile)	<i>MIGR</i> Mittelwert	<i>U20R</i> Mittelwert	<i>FOER</i> Mittelwert	<i>SPVL</i> Mittelwert
1 (belastet)	0,4437	0,2253	0,0330	16,4000
2	0,1879	0,3132	0,0241	16,4714
3	0,0808	0,3711	0,0139	16,7037
4	0,0685	0,4338	0,0089	19,8185
5 (unbelastet)	0,0405	0,5805	0,0062	19,9963

Bei dieser Analyse zeigt sich, dass der mittlere Migrantenanteil (*MIGR*) und der mittlere Förderanteil (*FOER*) mit steigendem Quintil geringer werden, während der mittlere Übergängeranteil auf Gymnasien (*U20R*) und die mittlere Schüler-Lehrer-Relation (*SPVL*) mit steigendem Quintil ansteigen. Insbesondere die starken Mittelwertunterschiede in den ersten drei Variablen der Tabelle zeigen auf, dass alle diese Variablen im Hinblick auf das Zielkonstrukt plausibel diskriminieren. Lediglich die steigenden Mittelwerte der Schüler-Lehrer-Relation scheinen unplausibel, da eine hohe Schüler-Lehrer-Quote auf eine unbelastete Schule hindeutet. Nähere Untersuchungen der Daten haben jedoch gezeigt, dass die Schüler pro Vollzeitlehrer-Quote und die Klassengrößen nicht wie erwartet negativ, sondern positiv oder gar nicht mit dem IFS-Index korrelieren. Das deutet darauf hin, dass die klassische Annahme, dass große Klassen auf eine starke Belastung der Schulen hindeuten, nicht bestätigt werden kann, sondern eher der Hinweis darauf gegeben wird, dass die Klassengrößen und somit das Schüler-Lehrer-Verhältnis im Zusammenhang mit dem IFS-Index etwas anderes abbildet. So kann z. B. bei Klassen mit vielen Migrantenkindern eine geringere Schüleranzahl beobachtet werden als bei Klassen mit vergleichsweise weniger Migrantenkindern.

Vor dem Hintergrund, dass alle vier Variablen bezüglich des Konstrukts soziale Belastung diskriminieren und auch fraglich ist, dass nur zwei Variablen, wie der Migrantenanteil (*MIGR*) und die Übergängerquote auf Gymnasien (*U20R*), die theoretisch vielschichtige soziale Belastung einer Schule erklären, wird die Faktorenlösung mit vier Variablen beibehalten.

Um die interne Validität des konstruierten Index zu überprüfen, wurde dieser in einem nächsten Schritt mit dem IFS-Index verglichen. Dabei richtete sich die Evaluation insbesondere auf die Frage- bzw. Problemstellung des Projektes, dessen Ziel es war, zu überprüfen, wie gut sich der IFS-Index durch die amtliche Schulstatistik bzw. einen mit Hilfe der amtlichen Schulstatistik konstruierten Index reproduzieren lässt.

Abbildung 1: Evaluation des Schulindex für Grundschulen mit dem IFS-Index, Streudiagramm inkl. Regressionsgerade und Bestimmtheitsmaß



a) Schulindex und IFS-Index

b) Schulindex und IFS-Index nach Ort

Quellen: IFS und eigene Berechnungen

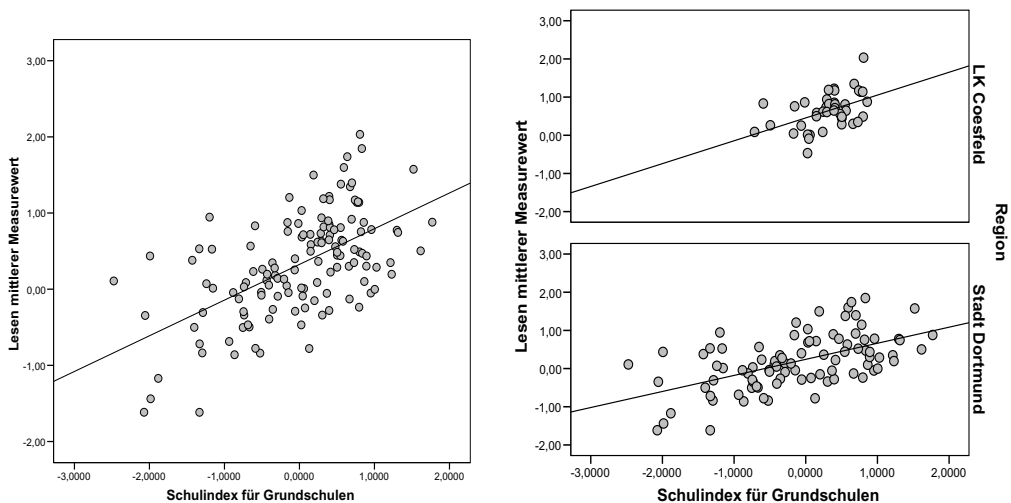
Abbildung 1 a) zeigt zunächst ein Streudiagramm der beiden Indizes sowie das dazugehörige korrigierte Gütemaß der Einfachregression. Bei einem adjustierten Gütemaß in Höhe von  $R^2_{adj} = 0,653$  und einer Korrelation in Höhe von 0,81 lässt sich das Ergebnis durchaus als zufriedenstellend bewerten. Trotz einiger kritischer Punkte aus methodischer Sicht erreicht der extrahierte Faktor eine hohe Korrelation mit dem IFS-Index. Ortsspezifisch variiert die Anpassung des Schulindex jedoch sehr stark (vgl. Abbildung 1 b). So lässt sich in Dortmund mit einem Gütemaß von  $R^2_{adj}(DO) = 0,704$  eine noch höhere Anpassung erreichen als bei der globalen Betrachtung, während die Anpassung für Coesfeld mit einem Wert von  $R^2_{adj}(COE) = 0,055$  verhältnismäßig schlecht ausfällt. Der Grund für diese geringe Anpassung im Kreis Coesfeld ist in einer mangelnden Variation der Variablen insgesamt zu sehen und bestätigt zudem die anfangs erwähnten (vermuteten) Stadt-Land-Unterschiede.

Neben der Frage, ob die amtliche Schulstatistik als Datenbasis zur Konstruktion eines Sozialindex geeignet ist, wird nachfolgend auch dargestellt, ob der konstruierte Schulindex mit Ergebnissen von Lernstandserhebungen korreliert. Zur Beantwortung dieser Frage wurde eine Evaluation des Schulindex mit den VERA-Daten des Jahres 2007 durchgeführt.<sup>19</sup> Es zeigt sich (vgl. Abbildung 2), dass der gebildete Schulindex eine Korrelation in Höhe von  $r = 0,580$  mit den Vergleichsarbeiten erreicht, welche zwar mittelmäßig, jedoch auf dem Niveau von  $\alpha \approx 0,000$  hoch signifikant ist. Auch vor dem Hintergrund, dass bei den Lernstandserhebungen das Leseverständnis, also eine andere Merkmalsdimension abgebildet wird, ist diese Korrelation als zufriedenstellend zu betrachten. Ortsspezifisch, d. h., bezogen auf die Pilotregionen, sinkt die Korrelation für die Dortmunder Grundschulen auf  $r(DO) = 0,561$  und für die Coesfelder Schulen auf  $r(COE) = 0,483$ , wobei das hohe Signifikanzniveau in beiden Fällen erhalten bleibt. Die positive Korrelation zeigt auf, dass der Schulindex für Grundschulen, welcher maßgeblich durch die

<sup>19</sup> Die Evaluation wurde aus Datenschutzgründen durch Dritte, das IFS in Dortmund, vorgenommen. Außer den Korrelationen bzw. Bestimmtheitsmaßen können keine weiteren Kennzahlen berichtet werden.

Variable Migrantenanteil bestimmt wird, durchaus mit den Ergebnissen der Lesekompetenzen in Beziehung gebracht werden kann.

Abbildung 2: Evaluation des Schulindex für Grundschulen mit VERA-Ergebnissen, Streudiagramm inkl. Regressionsgerade



a) Schulindex und Lernstandserhebung

b) Schulindex und Lernstandserhebung nach Ort

Quelle: IFS

Betrachtet man dieses Ergebnis vor dem Hintergrund der Analyse des Ministeriums für Schule und Weiterbildung des Landes Nordrhein-Westfalen,<sup>20</sup> so fällt auf, dass die Lesekompetenz in Schulen mit vergleichsweise ungünstigen Rahmenbedingungen – als Beispiele hierfür werden der Sprachhintergrund der Kinder, das Bildungsniveau der Eltern und die sozialen Verhältnisse in den Familien angeführt – geringer ist, als in Schulen mit vergleichsweise günstigeren Rahmenbedingungen. Noch deutlicher zeigen sich Unterschiede bei der Untersuchung der Lesekompetenz in Abhängigkeit vom Sprachhintergrund der Kinder. So lösen beispielsweise nur 5,5 % der getesteten Grundschüler der dritten Klasse, deren dominante Sprache nicht Deutsch ist, Aufgaben mit dem höchsten Schwierigkeitsgrad, gegenüber 23,9 % der Schüler, bei denen die dominante Sprache Deutsch ist.

Allerdings ist immer anzunehmen, dass die schulischen Leistungen nicht nur von den sozialen Bedingungen der Schule abhängig sind, sondern auch von der Gestaltung des Unterrichts und von Förderangeboten und weiteren Maßnahmen an den Schulen beeinflusst werden. Dadurch entstehen zusätzliche Einflüsse auf die Ergebnisse der Lernstandserhebung, die zu einer niedrigen Korrelation des Schulindex mit dem Ergebnis der Lernstandserhebung führen sollten. Deshalb kann die Korrelation von  $r = 0,580$  als ein zufriedenstellendes und plausibles Ergebnis angesehen werden.

20 Vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Hrsg.), Ergebnisse der Vergleichsarbeiten (VERA), Klasse 3, für das Land Nordrhein-Westfalen im Jahr 2007.

## 5.2 Schul-Sozialindex für weiterführende Schulen

Analog zu der Vorgehensweise bei der Konstruktion eines Schuldaten-basierten Index für Grundschulen wurde eine Faktorenanalyse für die weiterführenden Schulen durchgeführt. Bei der ersten Lösung mit allen acht Variablen zeigte sich, dass die Variablen *NVER* (Anteil nicht versetzter Schüler), *FOER* (Förderschwerpunkt) und *DALT* (durchschnittliches Alter der Lehrer) mit  $h_j^2 < 0,3$  eine vergleichsweise zu geringe Kommunalität aufwiesen, so dass sie aus der weiteren Analyse ausgeschlossen wurden. Das Ergebnis der Lösung mit fünf Variablen ist nachfolgend dargestellt (vgl. Tabelle 5). Das KMO-Kriterium zeigt mit einem Wert von  $KMO = 0,820$  im Vergleich zu der Lösung für die Grundschulen, dass die Faktorenanalyse mit den Variablen für die weiterführenden Schulen insgesamt ein zufriedenstellendes Ergebnis liefert. Die in Tabelle 5 dargestellten Kommunalitäten sind vor und nach der Extraktion insgesamt höher als die Kommunalitäten der Grundschulvariablen. Einzig die Kommunalität der Variablen *MIGR* fällt vor und nach der Extraktion vergleichsweise gering aus, was bei der ersten Lösung mit allen Variablen noch nicht der Fall war. Die Varianzklärung des extrahierten Faktors ist sowohl anfänglich als auch nach der Extraktion mit über 66 % bzw. über 58 % deutlich höher ausgefallen als für die Grundschulvariablen.

Entsprechend der hohen Kommunalitäten korrelieren die einzelnen Variablen stark mit dem extrahierten Faktor. Die höchste Korrelation erreicht die Variable *LOBR*, die die zusätzliche Belastung der Lehrer abbilden soll, mit 0,90. Die weiteren Korrelationen sind nachfolgend dargestellt: Durchschnittliche Klassengröße: 0,89, Schüler pro Vollzeitlehrer-Quote: 0,78, bestmöglicher Abschluss: 0,82 und Migrantenanteil: 0,65.

Tabelle 5: Faktorlösung für weiterführende Schulen

Variable	Faktorlösung					
	Kommunalitäten		Faktorladungen	Eigenwerte (kum. %)		
	Anfang	Extraktion		Faktor	Anfang	Extraktion
<i>DKLS</i>	0,653	0,699	0,836	1	66,524	58,687
<i>MIGR</i>	0,353	0,374	-0,612	2	79,047	
<i>BEST</i>	0,540	0,594	0,771	3	88,017	
<i>SPVL</i>	0,483	0,547	0,739	4	96,063	
<i>LOBR</i>	0,683	0,720	-0,848	5	100,000	

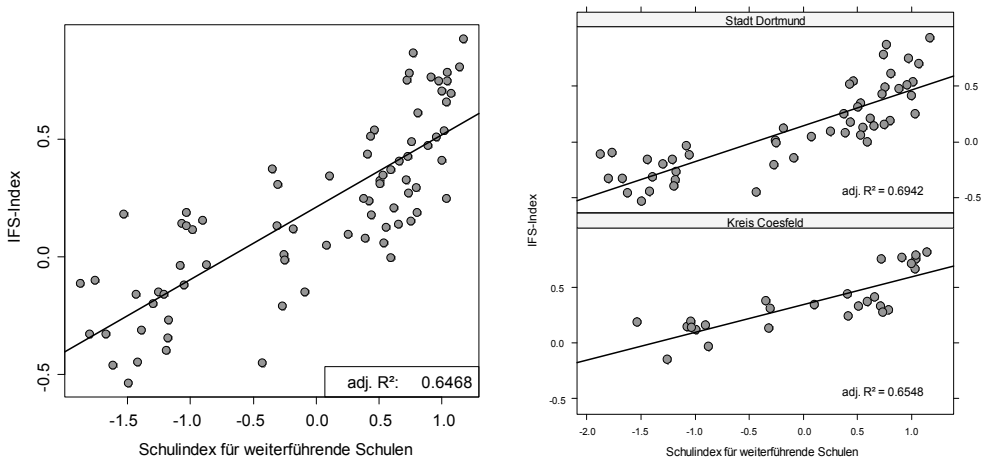
Wie auch in Abschnitt 5.1 für die Grundschulen dargestellt, wurde aus den Faktorladungen und den standardisierten Ausprägungen der Variablen ein Index konstruiert, der ebenfalls zunächst in fünf gleich große Klassen unterteilt wurde (vgl. Tabelle 6), um den Beitrag der einzelnen Variablen zu überprüfen. Bezogen auf die Interpretation der einzelnen Quintile lässt sich festhalten, dass insbesondere die Variablen *MIGR*, *BEST* und *LOBR* stark zwischen hohen und niedrigen Indexwerten diskriminieren können. So haben Schulen mit niedrigen Indexwerten (Quintil 1, belastete Schulen) einen höheren Migrantenanteil, niedrigere Quoten bei den besten Abschlüssen und einen höheren Anteil an Unterrichtsstunden, die ohne Lehrbefähigung erteilt wurden. Wie bei den Grundschulen zeigt sich auch bei den weiterführenden Schulen, dass die Schüler pro Vollzeitlehrer-Quote und die durchschnittliche Klassengröße einen der Vermutung entgegengesetzten Zusammenhang zum IFS-Index aufweisen. Dies lässt sich dadurch erklären, dass die Klassengrößen je nach Schulform stark variieren. So lassen sich z. B. bei den Hauptschulen im Durchschnitt die kleinsten Klassen ausmachen, während Gymnasien im Mittel eine größere Schüleranzahl je Klasse haben. Die Folge dieses Zusammenhangs ist, dass die Klassengröße und das Schüler/Lehrer-Verhältnis als Approximation für die nicht metrisch erfassbare Schulform verwendet werden können.

Tabelle 6: Quintile des Schulindex für weiterführende Schulen

Schulindex (Quintile)	<i>DKLS</i> Mittelwert	<i>MIGR</i> Mittelwert	<i>BEST</i> Mittelwert	<i>SPVL</i> Mittelwert	<i>LOBR</i> Mittelwert
1 (belastet)	21,2404	0,3529	0,1716	12,8294	0,3938
2	23,9790	0,2236	0,1997	15,1556	0,2303
3	27,5927	0,1354	0,3978	17,8059	0,1025
4	27,7935	0,0851	0,5945	19,2889	0,0810
5 (unbelastet)	28,7538	0,0508	0,8191	18,6722	0,0510

Der aus diesen fünf dargestellten Variablen konstruierte metrische Schulindex für weiterführende Schulen ist ebenfalls mit dem IFS-Index evaluiert worden (vgl. Abbildung 3). Mit einem adjustierten  $R^2$  von fast 0,65 und einer Korrelation von 0,81 ist die Anpassung des Schulindex an den IFS-Index nur marginal schlechter als für die Grundschulen. Zu beachten ist jedoch, dass bei den weiterführenden Schulen auch eine hohe ortsspezifische Anpassung erreicht wird.

Abbildung 3: Evaluation des Schulindex für weiterführende Schulen mit dem IFS-Index, Streudiagramm inkl. Regressionsgerade und Bestimmtheitsmaß



a) Schulindex und IFS-Index

b) Schulindex und IFS-Index nach Ort

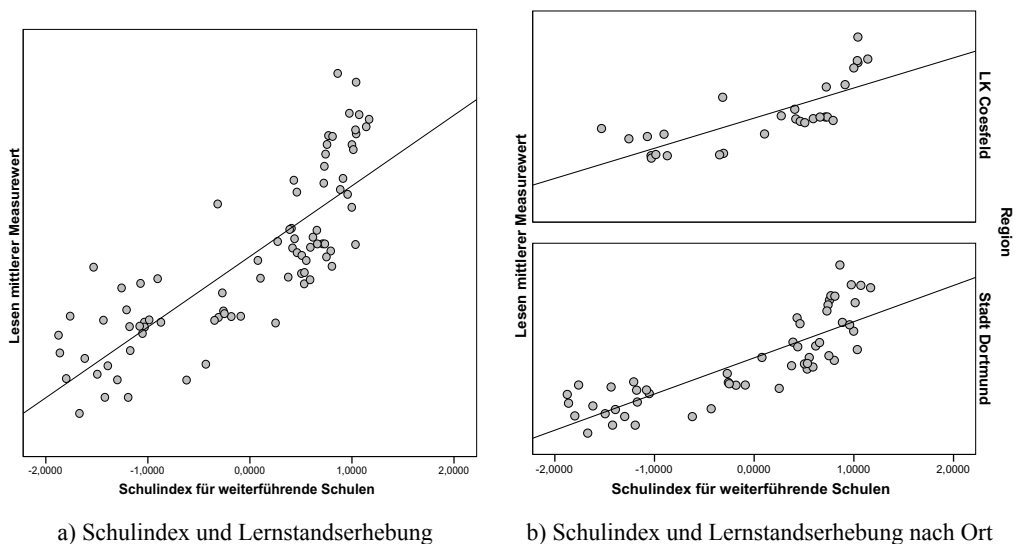
Quelle: IFS und eigene Berechnungen

In Coesfeld liegt der Erklärungsgehalt bei  $R^2_{adj}(COE) = 0,655$ , in Dortmund knapp darüber bei  $R^2_{adj}(DO) = 0,694$ . Die vergleichsweise hohe Anpassung für Coesfelder Schulen unterstützt die Vermutung, dass die Merkmale bei den weiterführenden Schulen eine größere Variation aufweisen als für die Grundschulen. Auffällig ist, sowohl insgesamt als auch ortsspezifisch, dass es Wertebereiche des Schulindex gibt, die nicht angenommen werden. Dieser Bereich liegt zwischen ca. -1,0 und -0,5 und ist deutlich als Lücke in den dargestellten Grafiken (vgl. Abbildung 3) zu erkennen. Bezogen auf einen Mittelwert von ca. 0 lässt sich daher festhalten, dass viele Schulen stark unterdurchschnittliche Indexwerte aufweisen. Der konstruierte Schulindex trennt die Grundgesamtheit also in zwei Gruppen. Weitere Analysen haben gezeigt, dass es sich bei dieser Gruppe von Beobachtungen um die Indexwerte der Hauptschulen handelt, da diese überwiegend in dem negativen Wertebereich des Schulindex liegen.

Für die weiterführenden Schulen bestand nach Abschluss der Indexkonstruktion ebenfalls die Möglichkeit, den konstruierten Index den Lernstandsdaten gegenüberzustellen. Die Korrelation zwischen dem Schulindex und den Ergebnissen der Lernstandsdaten war im Vergleich zu den Ergebnissen für die Grundschule mit  $r = 0,818$  ( $\alpha \approx 0,000$ ) sehr hoch. Dabei erreicht die Stadt Dortmund eine noch höhere Korrelation mit  $r(DO) = 0,835$ , die Korrelation für den Kreis Coesfeld war etwas geringer mit  $r(COE) = 0,766$ .

Auf Basis der Auswertungen, welche durch das MSW NRW durchgeführt wurden, lässt sich für die weiterführenden Schulen festhalten, dass eine diesem Index ähnliche Rangfolge der Lesekompetenz nach Schulform bestätigt werden kann.<sup>21</sup> So werden die Aufgaben mit dem höchsten Schwierigkeitsgrad auch eher von Gymnasiasten gelöst, während nur 1 % der Hauptschüler Aufgaben mit dem höchsten Niveau lösen. Daneben bestehen auch bei den weiterführenden Schulen erhebliche Unterschiede in der Lesekompetenz, je nach Standorttyp, was dem Ergebnis der Grundschulen sehr ähnelt und ebenfalls einen Hinweis auf die Entstehung dieser hohen Korrelation liefert.

Abbildung 4: Evaluation des Schulindex für weiterführende Schulen mit Ergebnissen der Lernstandserhebung, Streudiagramm inkl. Regressionsgerade



Quelle: IFS

Zusammenfassend lässt sich festhalten, dass die Konstruktion eines Indizes unter Verwendung der Faktorenanalyse für weiterführende Schulen bessere Ergebnisse liefert als für die Grundschulen. Die starke Streuung und die Bildung von zwei Gruppen innerhalb der Indexwerte lassen jedoch vermuten, dass eventuell noch andere Einflüsse eine Rolle spielen, die nicht berücksichtigt wurden. Grundsätzlich wird jedoch für die weiterführenden Schulen, bezogen auf die Korrelationen und Gütemaße für die Grundgesamtheit und die ortsspezifischen Stichproben, ein ähnliches, teilweise sogar ein besseres Ergebnis als für die Grundschulen erzielt.

21 Vgl. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Hrsg.), Ergebnisse der Lernstandserhebungen, Klasse 8, für das Land Nordrhein-Westfalen im Jahr 2007.

## 6 Kritische Anmerkungen und Fazit

Das Ziel dieser Untersuchung war, zu überprüfen, ob die amtliche Schulstatistik Informationen beinhaltet, die auf der Ebene der Einzelschule zur Konstruktion von validen Sozialindizes verwendet werden können. Wie in Abschnitt 5 erwähnt wurde, ist die Menge der Variablen, die sinnvoll auf die Schulebene aggregiert werden können natürlich beschränkt. Jedoch konnte – trotz vereinzelt vorhandener großer Unterschiede im Stadt-Kreis-Vergleich – auch anhand weniger Variablen demonstriert werden, dass sich ein durch eine aufwändige Befragung erhobener Index gut reproduzieren lässt.

Bezüglich der Methodenwahl Faktorenanalyse sollte festgehalten werden, dass die Reduktion auf eine Dimension der sozialen schulischen Belastung kritisch beurteilt werden sollte, nicht nur vor dem Hintergrund der nicht zufriedenstellenden Ergebnisse für die Grundschulen, sondern auch unter Beachtung der unterschiedlichen Indikatoren (z. B. Schüler- und Schulvariablen), die in die Untersuchung einfließen. Zu überdenken ist in diesem Zusammenhang, ob nicht mehrere Dimensionen, die die entsprechenden Gruppen von Variablen abbilden, der sozialen Belastung von Schulen angenommen werden sollten, so dass eine Mehr-Faktorlösung angestrebt werden sollte. Aufgrund der homogenen Ausprägungen im Kreis Coesfeld erscheint es zudem sinnvoll, eventuell andere Variablen oder Verfahren zur Indexkonstruktion heranzuziehen. Mit den auf die Schuleinzugsbereiche approximierten Daten der Bevölkerungs- und Arbeitsmarktstatistik, die parallel zu dem hier vorgestellten Vorgehen ebenfalls als alternativer Zugang zu einem Schul-Sozialindex analysiert wurden, konnte aber kein qualitativ besseres Ergebnis erreicht werden. Nicht genutzt wurden die Daten der sozialversicherungspflichtig Beschäftigten, die noch erweiterte Analysemöglichkeiten eröffnen könnten.

Unabhängig von einigen methodischen Kritikpunkten lässt sich zusammenfassend feststellen, dass eine Rekonstruktion des IFS-Index auf Basis der amtlichen Schulstatistik sowohl für Grundschulen als auch für weiterführende Schulen bereits mit wenigen Informationen möglich ist und zu zufriedenstellenden Ergebnissen führt. Insbesondere vor dem Hintergrund der Entstehung des IFS-Index, der auf einer kosten- und zeitintensiven Befragung an den Schulen beruht, sollten im Rahmen der Umstellung auf eine Individualstatistik weitere Ansatzpunkte der amtlichen Statistik genutzt werden, die Situation an den Schulen differenzierter zu untersuchen und genauer zu vergleichen. Daneben muss jedoch auch beachtet werden, dass zwei Regionen betrachtet wurden, die nicht nur im Bezug auf das Land NRW sehr unterschiedlich sind, sondern auch für sich betrachtet starke Unterschiede aufweisen. So weist der Kreis Coesfeld in vielen Merkmalen eine große Homogenität auf, während für die Stadt Dortmund eine teilweise extreme Variation in einigen Variablen festzustellen ist. Um beurteilen zu können, ob die Informationen, die die amtliche Schulstatistik liefert, in Bezug auf Handlungsalternativen genügen, sind weiterreichende Untersuchungen mit größeren Stichproben und zusätzliche Referenzgrößen notwendig, was in letzter Konsequenz zu der Anforderung einer Datenbasis auf Landesebene führt. Aus unserer Sicht würde sich aber die weitere Entwicklungsarbeit in die begonnene Richtung lohnen, weil so ein ohne zusätzliche Erhebungen jährlich aktualisierbares Verfahren zur Ermittlung der sozialen Belastung von Schulen erreicht werden kann.

*Verf.: Prof. Dr. Horst Weishaupt, DIPF – Deutsches Institut für Internationale Pädagogische Forschung, Schloßstr. 29, 60486 Frankfurt am Main, E-Mail: weishaupt@dipf.de*

*Dipl.-Ök. Anna Makles, Bergische Universität Wuppertal, Schumpeter School of Business and Economics, Lehrstuhl für Wirtschaftsstatistik, Gaußstraße 20, 42119 Wuppertal, E-Mail: makles@statistik.uni-wuppertal.de*