Falsch abgebogen

Category: Blog
geschrieben von Gastautor | 19. März 2024

Eine Sonderauswertung mit einem Vergleich der Bundesländer hatte bei der ersten PISA-Studie im Jahr 2000 ergeben, dass Bayern und Baden-Württemberg über dem OECD-Durchschnitt lagen, und zwar recht deutlich, und Sachsen nur wenig darunter. Diese drei Bundesländer hatten dabei die stringentesten institutionellen Regelungen: die Verbindlichkeit der Schullaufbahnempfehlungen, eine stärkere organisatorische Kontrolle der Schulen, etwa über die Standardisierung des Stoffs, zentrale Prüfungen, regelmäßige Tests und eine Rechenschaftspflicht der Schulen und ihrer Lehrer. Die schlechtesten Leistungen und eine stärkere soziale Ungleichheit zeigten sich dagegen in Bremen, Berlin, Hessen und Nordrhein-Westfalen. Das aber waren ausgerechnet jene Bundesländer mit der größten Öffnung und der stärksten Lockerung. Es gab oder gibt dort keine Verbindlichkeit und eine nur geringe Kontrolle.

Baden-Württemberg und Bayern mit ihren besonders stringenten Regelungen

lagen auch schon in den Grundschulen weit über dem internationalen Durchschnitt.

Hartmut Esser, deutscher Migrationswissenschaftler sowie Professor für Soziologie und Wissenschaftstheorie an der Universität Mannheim.

Die Zusammenhänge lassen sich gut nachvollziehen: Verbindlichkeit und Kontrolle sind darauf gerichtet, dass die Kinder tatsächlich nach ihren Fähigkeiten und Leistungen auf die unterschiedlichen Bildungswege verteilt werden und dass die damit verbundenen institutionellen Änderungen, die Anpassung der Curricula und des Unterrichts an die Unterschiede, auch wirklich implementiert werden. Die Öffnung und Aufteilung auf die Bildungswege nach Belieben ist dagegen mit Nachlässigkeiten und Ungerechtigkeiten aller Art, sowie mangelnder Effizienz verbunden. Der im einfachen Vergleich der Nationalstaaten bei PISA 2000 entstandene Eindruck, dass erst mit Öffnung, Lockerung und Integration, wie in Schweden und Finnland, die Leistungen besser und die Bildungsungleichheit geringer würden, hatte also gleich zu Beginn getrogen, und zwar massiv.

Auch IGLU 2003 für die Grundschulen belegte früh das Gegenteil der seitdem gängigen Auffassung: Baden-Württemberg und Bayern mit ihren besonders stringenten Regelungen lagen auch schon in den Grundschulen weit über dem internationalen Durchschnitt und kamen fast an Schweden heran, Nordrhein-

Westfalen und Bremen mit ihren Öffnungen blieben deutlich darunter. Die anderen Bundesländer hatten nicht teilgenommen oder lagen auch im unteren Bereich, Hessen etwa.

Im Vergleich der Sekundarstufen berichteten die PISA-Vergleiche im Jahr 2003 und dann 2006 von ersten Anzeichen der Besserung in Niveau und sozialer Durchlässigkeit – und von Verschlechterungen in Schweden wie auch in Finnland. Das verstärkte sich 2009, dann noch einmal 2012 und schließlich auch 2015, sodass sich die Länder im internationalen Vergleich mit und ohne Differenzierung kaum noch unterschieden. Im Vergleich der deutschen Bundesländer waren Sachsen und Bayern 2018, also vor Corona und damit noch in der Zeit mit einem regulären Unterricht, in den Leistungen wieder die besten in Deutschland – und international besser als Schweden und sogar Finnland. In der sozialen Durchlässigkeit unterschieden sich die Länder ohnehin nicht mehr.

Breiter Konsens

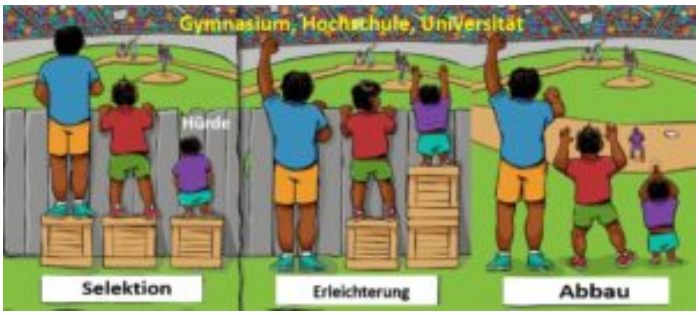
Gleichwohl verstummten die gewohnten Vorhaltungen nicht, in Deutschland mit seiner frühen und rigiden Trennung der Kinder nach der Grundschule seien die Probleme besonders groß und, so mindestens der Subtext, nur durch den Übergang zur vollen Integration beziehungsweise zu einem „längeren gemeinsamen Lernen“ über die ganze Pflichtschulzeit hinweg zu lösen. Es war nach PISA 2000 ein breiter Konsens, und man könnte schon von einer nicht weiter bezweifelten Standardposition zu der Frage nach den Effekten von Differenzierung und Integration in den systematischen Übersichten und Metaanalysen der Bildungsforscher, mancher Institute und Stiftungen und in weiten Teilen der akademischen Bildungsforschung sprechen – bis in die letzte Zeit hinein, obwohl es so noch nie gestimmt hatte.

Theresa Schopper, Kultusministerin, Grüne: 2011 begann der Abstieg.

Auf dem Gipfelpunkt der weithin unbeachteten Verbesserungen und Konvergenzen gab es nun ein wissenschaftlich interessantes, wenngleich politisch umstrittenes Feldexperiment. In Baden-Württemberg war 2011/12 mit dem politischen Wechsel zu Grün-Rot die Verbindlichkeit der Grundschulempfehlungen abgeschafft worden. Das könnte negative Folgen haben, so hatte zum Beispiel der Bildungsforscher Jörg Dollmann früh genug gewarnt, der den damals wohl einzigen Datensatz zur Verfügung hatte, der einigermaßen brauchbar war. Ohne Noten und verbindliche Regelungen entscheiden nicht die Fähigkeiten und Leistungen der Kinder über den weiteren Bildungsweg, sondern vor allem der Ehrgeiz der Eltern, insbesondere in den oberen sozialen Schichten. Zugleich wagen die bildungsferneren Eltern der unteren sozialen Schichten nicht den unbekannten Schritt nach oben und schicken auch begabte Kinder nicht aufs Gymnasium. Das aber führe dazu, so konnte man vermuten, dass mehr Kinder aus den sozial privilegierten Schichten auf dem Gymnasium landeten, die die Leistungsanforderungen nicht erfüllen könnten, aber auch, dass gerade die talentierten Kinder aus den unteren Schichten ihre Chancen nach oben nicht bekommen würden. Hinzu kommt, dass es mit der Verbindlichkeit und dem Wissen über die damit verbundenen Hürden einen besonderen Anreiz gibt, sich schon vor dem Übergang besonders anzustrengen, der bei der freien Wahl kleiner ist oder ganz entfällt.

> *Baden-Württemberg war 2015 und 2018 beim Aufschwung der stringent gebliebenen Bundesländer Bayern und Sachsen nicht mehr dabei.*

So kam es auch. Baden-Württemberg war 2015 und 2018 beim Aufschwung der stringent gebliebenen Bundesländer Bayern und Sachsen nicht mehr dabei. Die waren ungerührt auf dem eingeschlagenen Erfolgspfad geblieben und weiter nach oben davongezogen. Baden-Württemberg sackte zwar nicht unbedingt ab, es blieb jedoch stecken und lag bald nur noch wenig vor Hessen und Nordrhein-Westfalen im unteren Bereich. Und sozial durchlässiger war es auch nicht geworden.



Weiche Kriterien für Selektionsabbau

Das Experiment mit der Abschaffung der Verbindlichkeit der Übergangsempfehlungen war also offensichtlich gescheitert. Dafür gab es noch einen ganz besonderen Grund, den man später fand: Die Kinder hatten nach der Umstellung offenbar noch in der Grundschule und vor dem Übergang die Griffel sofort fallen gelassen als klar wurde, dass sie sich die weiterführende Schule ohnehin aussuchen konnten. Inzwischen haben mehrere Studien diese und weitere Effekte für den Bildungserfolg bestätigt, für den Fall der Reform in Baden-Württemberg speziell, aber auch allgemein: Die Verbindlichkeit steht für die Bedeutung, die der Schule, dem Lehrpersonal und den Leistungen beigemessen wird, und lässt offenbar ein eigenes, übergreifendes „Bildungsklima“ entstehen, das nicht auf Einzelaspekte sehen muss, sondern sich wie selbstverständlich von alleine trägt.

Bis 2012 ungefähr war man in Deutschland und seiner so rigiden Differenzierung nicht mehr ungebrochen der Prügelknabe der PISA-Berichte und der öffentlichen Debatten wie bis dahin. Woran der Aufschwung zwischen 2000 und 2012 lag, ist jedoch nicht geklärt. Wahrscheinlich hatte er mit den besonderen Anstrengungen um die Kinder in den unteren Leistungsbereichen und einer doch deutlich erhöhten öffentlichen Aufmerksamkeit und Anstrengung nach dem PISA-Schock von 2000 zu tun. So hätte es eigentlich weiter gehen können: schrittweise Verbesserungen an kritischen Stellen und die unaufgeregte, regelmäßige Überprüfung, ob es voran geht oder auch nicht.

> *Sie werden auch mit deutlichen Empfehlungen eher von den Eltern zurückgehalten, auf die regulären Gymnasien zu gehen, wo sie erheblich bessere Leistungen erbracht hätten. Für die oberen Schichten sieht es anders aus: Sie haben immer ihre Optionen, darunter die Privatschulen oder andere Zufluchtsorte, wenn es kognitiv nicht reicht.*

Nach 2012 gingen über die Jahre so ab 2012 bis 2018 die Leistungen zurück, zuerst allmählich und dann immer deutlicher. Offenbar hatten die vielen mehr oder weniger sichtbaren Öffnungen und Lockerungen in den anderen Bundesländern ihren Anteil daran: bei den Noten, beim Unterricht, bei den Klassenarbeiten und manchem anderen noch. Möglicherweise hat auch die zunehmende Zusammenlegung der unteren Bildungswege eine Rolle gespielt, die

Schaffung von „Gesamtschulen“ oder „Stadtteilschulen“, parallel zu den differenzierten Schultypen. Es gibt bislang keine wirklich überzeugenden Hinweise, dass sie etwas bringen. Sie wirken eher wie Mobilitätsfallen für gerade die talentierten Kinder der unteren Schichten. Sie werden auch mit deutlichen Empfehlungen eher von den Eltern zurückgehalten, auf die regulären Gymnasien zu gehen, wo sie erheblich bessere Leistungen erbracht hätten. Für die oberen Schichten sieht es anders aus: Sie haben immer ihre Optionen, darunter die Privatschulen oder andere Zufluchtsorte, wenn es kognitiv nicht reicht.

Auch beim Abschwung erreichen die Bundesländer mit der größten Stringenz die besten Durchschnittsleitungen. Außerdem schaffen dort die meisten Kinder die Mindeststandards – gerade in den unteren Leistungsbereichen. Baden-Württemberg verliert nach 2012 den Anschluss nach oben und rückt nahe an die schwachen Flächenländer Hessen und Nordrhein-Westfalen heran. Über Berlin und Bremen müssen wir nicht mehr reden: nur die Katastrophe.

Für 2022 zeigt sich der Verfall nach der coronabedingten Schließung der Schulen endgültig. Aber wieder halten sich die stringenten Länder auf hohem Niveau besser. Baden-Württemberg verliert weiter und bei den schwachen Bundesländern rutschen manche deutlich in den negativen Bereich im Vergleich zu 2013. In der Abbildung ist das zusammenfassend dargestellt.

> *Sachsen und Bayern, legen im Vergleich von 2009 bis 2018, dem letzten Jahr vor den Krisen und Schulschließungen, weiter zu, deutlich sogar. Alle anderen verlieren.*

Der allgemeine Aufschwung bis 2012 und die Unterschiede der Bundesländer nach Stringenz und Leistungsniveau sind gut erkennbar. Ab da bröckelt es, aber nicht überall. Die nach 2012 noch stringenten Bundesländer, Sachsen und Bayern, legen im Vergleich von 2009 bis 2018, dem letzten Jahr vor den Krisen und Schulschließungen, weiter zu, deutlich sogar. Alle anderen verlieren – oder stagnieren, wie auch Baden-Württemberg gleich nach 2012, der Zeit nach der Reform. Deutschland insgesamt ist seit 2009 (und eigentlich schon früher) nicht mehr unterdurchschnittlich, aber die offenen Bundesländer verderben immer den Schnitt, vorher schon und jetzt erst recht. An ihnen hängt es, dass Deutschland und die Differenzierung bei PISA nicht besser zu sein scheinen.

Ganz ähnlich war es bei den Grundschulen. In der IGLU-Studie 2022 zeigte sich,

dass es in Deutschland zwar einen deutlichen Rückgang in den Leseleistungen gegeben hat und nur einen Platz im Mittelfeld hinter Finnland und Schweden. Wieder gab es nach den Angaben der jeweiligen IQB-Berichte die schon bekannten Unterschiede für die deutschen Bundesländer: Die stringenten Länder sind auch schon in der Grundschule besser, und das auch noch in den Krisenzeiten, und die offenen fallen besonders stark ab.

Der einstige PISA-Sieger Finnland erfährt einen totalen Absturz.

Soweit die Sachlage. Sie ist alles andere als das, was man bis dahin zu lesen bekommen hatte. Aber warum dann jetzt auf einmal erst diese ganz andere „Wirklichkeit“? Gab es die deskriptiven Berichte denn nicht schon, bei PISA wie beim IQB-Bildungstrend? Und was ist dann aus den vielen, mehr oder weniger komplexeren statistischen Analysen geworden, die alles Mögliche einbezogen haben, was die deskriptiven Unterschiede auch (kausal) erklären hätte können: weitere Eigenschaften der Länder, der Schulen und Schulklassen, der Familien und der Kinder? Natürlich gab es die, es ist eine ganze Industrie daraus entstanden, aber nahezu immer zeigten sie den gleichen Befund: Eher geringere Leistungen bei Differenzierung, besonders bei den Kindern aus den schwierigeren Verhältnissen und, insbesondere, eine Verstärkung des Effektes der sozialen Herkunft, und das stets in Deutschland „wie kaum in einem anderen Land“. Aber konnte das angesichts der Unterschiede unter den Bundesländern stimmen, die das Gegenteil anzeigten?

Es gibt zwei Antworten. Die eine kennen wir schon: In den PISA-Vergleichen

finden sich nur die nationalen Unterschiede insgesamt und ohne die regionalen Unterschiede. Die andere ist dagegen schon überraschend: In den meisten Untersuchungen wurden zentral wichtige Einflussgrößen auf die Leistungen nicht berücksichtigt, sei es, weil sie schon in den Daten fehlten oder in den Analysen keine Rolle spielten.

Das kann fatale Folgen haben. Fragt man nämlich nach den Effekten bestimmter Bedingungen als Ursache für einen zu erklärenden Sachverhalt, dann müssen in den betreffenden Untersuchungen die jeweils bedeutsamen Bedingungen möglichst vollständig erfasst und in ihrer jeweils eigenen Wirksamkeit bestimmt, also statistisch voneinander getrennt werden. Für das Lernen und die Leistungen sind das die kognitiven Fähigkeiten der Kinder, also die Intelligenz. Sie ist der mit Abstand stärkste Faktor. Aber ausgerechnet der ist in den PISA-Daten nicht enthalten und in den IQB-Berichten, in denen es ihn gab, nicht analysiert worden. Es wird immer nur die soziale Herkunft ausgewiesen – neben einigen weiteren Faktoren wie Geschlecht, Migrationshintergrund oder der Vorschulbesuch. Das sind weit weniger wichtigere Indikatoren. Wenn aber der mit Abstand bedeutsamste Faktor in den Analysen nicht berücksichtigt wird, kann das nur zu Verzerrungen und Fehlurteilen führen. Das gilt analog für die in den Begründungen der Differenzierung so wichtigen Effekte der kognitiven Zusammensetzung der Schulklassen.

> *Die Herkunftseffekte werden generell überschätzt und die für ein steigendes Leistungsniveau verantwortlichen kognitiven Fähigkeiten werden fälschlicherweise als Verstärkung des Einflusses der sozialen Herkunft beziehungsweise der sozialen Segregation in den Schulklassen interpretiert.*

Gary Marks, PISA-Forscher: Keine

Verbesserung der Chancengerechtigkeit

Die Folgen: Die Herkunftseffekte werden generell überschätzt und die für ein steigendes Leistungsniveau verantwortlichen kognitiven Fähigkeiten werden fälschlicherweise als Verstärkung des Einflusses der sozialen Herkunft beziehungsweise der sozialen Segregation in den Schulklassen interpretiert. Es handelt sich also um eine glatte Fehlzuschreibung der Systemeffekte, um die es in der Auseinandersetzung geht. Und das in einem so aufwendigen und wichtigen Projekt wie den internationalen Vergleichsstudien. Gary Marks, ein international bekannter PISA-Forscher von der Universität Melbourne, hat dazu 2014 ein aufschlussreiches und in den zuständigen Fachkreisen anerkanntes Buch geschrieben. Nicht alle hat das offenbar erreicht.

Warum aber hat das kaum jemand vorher gemerkt oder gar ausgesprochen? Es ist eine schwierige und äußerst heikle Angelegenheit. Vielleicht war es eine Art von institutionalisierter Blindheit in Kombination mit gut verankerten politisch-ideologischen Einstellungen gewesen: Der PISA-Schock von 2000 schien mit einem Schlag die ewige Kontroverse um Differenzierung und Integration erledigt zu haben. Und dann kamen die vielen Programme und Institutionen der Bildungsforschung mit ihren Festlegungen in Ansatz und Untersuchungsanlage, die kaum eine andere Perspektive erlaubten. Man hätte sich sehr in die Nesseln setzen können, wenn man an etwas anderes auch nur gedacht hätte.

Aber es hatte noch einen anderen, ganz einfachen Grund: Erst ab 2006 gab es ein Projekt, die „National Educational Panel Study“, abgekürzt NEPS, das die nötigen Daten auch im Zeitverlauf und für die unterschiedlichen Stadien des Bildungsverlaufs erfasst hatte, allerdings nur für die deutschen Bundesländer, nicht international. Es war wohl der Versuch, es endlich einmal richtig zu machen, und sei es auch nur regional begrenzt. Das dauerte natürlich. So kam es, dass die eigentlich immer schon erforderlichen Analysen erst viel später möglich wurden. Ab 2018 zeigte sich so plötzlich das ganz andere Bild. Da aber war man schon längst falsch abgebogen.

Nur wenig später hatte die Bildungsadministration ganz andere Sorgen, „Corona“ vor allem und die damit bedingten Schulschließungen, dann 2022 der so plötzliche Lehrermangel und die versäumte Digitalisierung. An die „Systemfrage“

hat niemand mehr gedacht. Vielleicht war man auch ganz froh. Es wären politisch und auch wissenschaftlich höchst riskante Fragen gewesen, an das Bildungsministerium und die Kultusministerkonferenz (KMK) zuerst, dann aber auch an die nach PISA so breit ausgebaute Bildungsforschung mit ihren vielen Sonderprojekten und PhD-Programmen, die den kommenden Absturz angesichts der Öffnungen wohl nicht für möglich halten konnten.

Das Thema ist dann auch von anderen Problemen überlagert worden: Miserable bauliche Zustände, Lehrkräftemangel, Probleme bei der Digitalisierung, immer mehr Unruhe in der Gesellschaft, Flüchtlingskinder in einem vorher und anderswo kaum bekannten Maß. Derzeit scheint die Stimmung in eine noch andere Richtung umzuschlagen: Alles nicht so tragisch, wahrscheinlich ohnehin nicht zu klären, Reformen brauchten ihre Zeit und der aktuelle Schulfriede reicht uns schon. Im Moment brennt uns sowieso ganz was anderes auf den Nägeln, wie die Lehrerschelte von Alexander Schleicher als seine Erklärung für den Verfall – nachdem er vorher kaum etwas anderes kannte als die Forderung nach einer Aufgabe der Leistungsdifferenzierung und der Schaffung von integrierten Einheitsschulen. Die Forderung nach einer Aussetzung der PISA-Vergleiche liegt dann nahe. Aber das wäre der schlechteste Fall. Es hieße, die schwierige Suche nach den tieferen Ursachen abzubrechen und sich von den aktuellen Ereignissen überrollen zu lassen.

Der Autor hat Soziologie und Wissenschaftstheorie an der Universität Mannheim gelehrt.