

Hartmut Esser (Universität Mannheim)

Julian Seuring (Leibniz-Institut für Bildungsforschung, Bamberg)

„Wenn einer, der mit Mühe kaum ...“! Zur Kritik des Kultusministeriums Baden-Württemberg an der Studie „Kognitive Homogenisierung, schulische Leistungen und soziale Bildungsungleichheit“ von Esser und Seuring (2020)

Nach der Landtagswahl 2011 wurde von der neu gewählten grün-roten Regierung in Baden-Württemberg (BW) die Verbindlichkeit der Empfehlungen für den Übergang in die gymnasialen Bildungswege abgeschafft. Die Begründung: Es gebe weniger leistungshemmenden Stress, der Elternwille werde gestärkt und den sozial benachteiligten Kindern der Weg nach oben erleichtert. Seitdem sind mehr als zehn Jahre vergangen und es hat sich einiges getan: BW ist bei den Leistungen aus der Spitzengruppe, die es bis dahin zusammen mit Bayern und Sachsen gebildet hatte, ausgeschieden; der Abfall war gerade in den unteren Leistungsbereichen zu sehen, und die soziale Durchlässigkeit ist eher gesunken. In aktuelleren Studien zeigte sich dann auch, dass die Verbindlichkeit, anders als stets gedacht, die soziale Bildungsungleichheit keineswegs verstärkt, wohl aber die Leistungsgerechtigkeit des Übergangs erhöht, weil die Eltern gerade aus den oberen Schichten nicht mehr machen können, was sie wollen (Esser und Hoenig 2018), dass es stärkere familiäre Aktivitäten, ein größeres Interesse an schulischen Angelegenheiten und höhere Leistungen in der Grundschule schon gibt (Bach und Fischer 2020, Grewenig 2021). Mit der Verbindlichkeit gibt es offenbar besondere Anreize, sich schon früh anzustrengen und um die schulischen Belange zu kümmern.

Unabhängig davon, zusätzlich und auf der gleichen Linie zeigte sich in einer Untersuchung von Esser und Seuring (2020), dass es in den strikter nach den kognitiven Fähigkeiten differenzierenden Bundesländern in der Sekundarstufe weitere Leistungssteigerungen in den Schulklassen mit einer höheren kognitiven Homogenität gibt und dass davon gerade die Kinder in den unteren Bereichen profitieren, dass es also die so oft beklagten Schereneffekte nach oben nicht gibt und es dabei zu keinerlei Verstärkung des Einflusses der sozialen Herkunft kommt, eher im Gegenteil. Über diese Befunde gab es, nicht verwunderlich, alsbald auch eine politische Diskussion und in diesem Zusammenhang Einwände aus der etablierten Bildungsforschung gegen die Befunde bei Esser und Seuring (2020), im Folgenden: ES, insbesondere in einer Replik von Matthewes und Heisig (2022), im Folgenden: MH, auf die sich eine Stellungnahme des Kultusministeriums BW bezieht, nämlich:

„ ... dass die festgestellten Effekte der Leistungszunahme verschwinden, wenn weitere Ausgangsvoraussetzungen der Schülerinnen und Schüler – wie etwa deren kognitive Eingangsvoraussetzungen – berücksichtigt werden.“ (Landtagsdrucksache 17/1586)

Es ist nicht der einzige Einwand von MH gegen ES, aber der wohl zentrale.¹ Er ist in der Stellungnahme etwas unglücklich formuliert. Gemeint ist wohl das, was MH in der Tat einwenden: Die Effekte der kognitiven Homogenität der Schulklassen in der Sekundarstufe verschwinden, wenn neben den *kognitiven Fähigkeiten* auch die *Leistungen vorher* in der Grundschule statistisch kontrolliert werden. Aus Platzgründen konzentriert sich die Antwort hier auf diesen einen Aspekt. Er durchzieht, das nur nebenbei, die ganze Replik von MH und spielt implizit auch bei Klemm (2020) eine Rolle.

Worum geht es bei MH? Nach dem Modell der Leistungsdifferenzierung, dem theoretischen Kern der Analysen bei ES, werden die Leistungen in der Sekundarstufe direkt von Leistungen vorher in der Grundschule mitbestimmt, und der Effekt ist, überall wo man das geprüft hat, empirisch auch mit Abstand der stärkste. Das legt die Vorstellung nahe, die MH im Abschnitt 3.4 ihrer Replik aufgreifen: Die Leistungen *vorher* tragen die Systemeffekte und damit eben *nicht* auch noch das, was in der *Sekundarstufe* in den Schulklassen und über die kognitive Homogenität geschieht. Die Leistungen vorher müssten daher immer mit den kognitiven Fähigkeiten *zusammen* kontrolliert werden. Und wenn man das mache, verschwänden alle Systemeffekte der Stringenz (Tabelle 2 bei MH: 14). Alles, was es an Systemeffekten gebe, sei demnach schon in der Grundschule vorhanden, in der Sekundarstufe, wo es ja erst die Differenzierung gebe, passiere weiter nichts mehr. Was ist dazu zu sagen? Es sind fünf Punkte.

1. Der erste Punkt ist gleich schon der entscheidende: Kognitive Fähigkeiten und Leistungen sind *nicht* das Gleiche: Die *kognitiven Fähigkeiten* beinhalten relativ stabile, latente Eigenschaften der schnelleren Informationsverarbeitung, die im Prinzip in *allen* Umgebungen

¹ Aus Platzgründen konzentriert sich die Antwort auf diesen einen Aspekt. Zu den anderen Punkten und den in Landtagsdrucksache auch erwähnten Beiträgen von Klemm (2021) und Lorenz et al. (2022) gibt es eine ausführliche Besprechung bei Esser (2022). Die als Relativierung der Befunde bei ES benannten Studien von Decristan et al. (2017) und Gröhlich et al. (2009) sind dafür nicht relevant: Bei Decristan et al. (2017) geht es um die *Grundschule*, Gröhlich et al. (2009) finden zwar keine Homogenitätseffekte, dafür aber starke positive Effekte des Gymnasiums, was dem Effekt bei ES entspricht. Beide Studien beziehen sich zudem ohnehin *nicht* auf Systemeffekte, sondern auf die allgemeinen Zusammenhänge (vgl. dazu Esser 2021, Kapitel 8). Und der Hinweis darauf, dass bei einem höheren kognitiven Niveau eher höhere Effekte der kognitiven Heterogenität auftreten ist eine *Widerlegung* der üblichen Vorstellung, wonach es bei Homogenität stets die beklagten Schereneffekte gibt, wonach die Differenzierung die Ungleichheit nach *oben* vergrößere. Tut sie nämlich *nicht*.

und für *alle* Inhalte von Bedeutung sind und sich über die *Effizienz* des Lernens auswirken: Der gleiche Input erzeugt, *ceteris paribus*, einen höheren Output. Die jeweils erreichten *Leistungen* spiegeln dagegen jeweils gegebene *Bedingungen*, von denen das Lernen bei einem weiteren Input ausgeht, der aber nicht effizienter verarbeitet werden kann. Das aber heißt: Wenn in der Sekundarstufe die Anforderungen steigen, dann haben die Kinder mit einem höheren Niveau in den *Leistungen vorher* zwar einen Vorsprung im Wissen, was aber *nur* denen auf dem weiteren und schwieriger werdenden Weg hilft, die auch die dazu nötigen Potentiale an *kognitiven Fähigkeiten* haben.

2. Daher sind die kognitiven Fähigkeiten und die Leistungen (*vorher*) auch *keine Äquivalente*, wenn es um die Effekte in den Schulklassen geht. Es heißt nicht ohne Grund *Ability-Tracking* – und eben *nicht* „Achievement“-Tracking. Es ist der Unterschied zwischen *Fähigkeiten* und angelernten *Kompetenzen*, Vorwissen und bloße Test-Performanzen nach einem entsprechenden Einpauken in der Panik (vgl. Esser 2021: Abschnitt 4.1) und entsprechend den Schulklassen, in denen dann viele Kinder mit einem womöglich nur sehr mühsam erworbenem Vorwissen sitzen, das ihnen dann aber nicht viel mehr hilft: „Wenn einer, der mit Mühe kaum „! Das Problem wird von einer anderen Seite her besonders augenfällig: Wenn man an Migrantenkinder mit hohem *kognitivem* Potential denkt, die aufgrund von akzidentellen Sprachproblemen die Leistungen in der Grundschule nicht erbringen können, es später locker schaffen würden, so aber fehlplatziert werden, was dann auch die mögliche kognitive Homogenisierung und deren Effekte in der Sekundarstufe verwässert.

3. Worauf es also beim Ability-Tracking ankommt, ist also der Effekt der *insgesamt* und *faktisch* entstehenden Verteilung nach den *kognitiven Fähigkeiten* – *gleichgültig* über welche einzelnen Wege das geschieht. Dieser Gesamteffekt wird auch als *totaler* kausaler Effekt der kognitiven Fähigkeiten und Schuleffekte auf die Leistungen in der Sekundarstufe bezeichnet. In ihnen ist *alles* enthalten, was dazu führt, *darunter* auch die Leistungen vorher und die Zusammensetzung der Schulklassen (nur) nach den Kompetenzen.

4. Wenn es aber um den *totalen* Effekt der kognitiven Fähigkeiten und Schuleffekte geht, würde bei der Kontrolle der Leistungen vorher, über die ja ein Teil der kognitiven Effekte vermittelt wird, ein Problem entstehen, das als *Overcontrol* bezeichnet wird (vgl. Elwert und Winship 2014: 35f.; Rohrer 2018: 38f.): Man würde etwas kontrollieren, was in dem angezielten Effekt empirisch schon enthalten ist, daher der Ausdruck von der „Über“-Kontrolle. Und statt des

angezielten totalen Effektes hätte man nur einen *partiellen*, aus dem alles, was auf dem gleichen Kausalpfad liegt, herausgerechnet ist.

5. Die Folge liegt auf der Hand: Der eigentlich interessierende totale Effekt des *Ability-Tracking* wird *künstlich* verringert oder auch ganz wegerklärt, von anderen Verzerrungen ganz abgesehen. Und genau das passiert bei MH ja auch: Sie kontrollieren in ihrer Replikation die Leistungen in der Grundschule vorher *zusammen* mit den kognitiven Fähigkeiten – und die Effekte der kognitiven Strukturen der Schulklassen verschwinden in der Tat weitgehend oder auch ganz. ES hatten auf diese Problematik, das Analyseziel und die Analysestrategie bereits hingewiesen:

„Als erklärende Bedingung für die Leistungen in der Sekundarstufe werden *gezielt* die *kognitiven Fähigkeiten* verwendet, individuell wie als Kontext-Merkmal, und *nicht* die Leistungen vorher in der Grundschule. Denn darum geht es in den Begründungen für die Differenzierung: Dass sich damit die latenten *kognitiven* Fähigkeiten entfalten können, auch unabhängig davon, was bis dahin und aktuell an Leistungen gezeigt worden ist.“ (ES 2020: 290; Hervorhebungen nicht im Original)

Die Effekte der Verbindlichkeit auf die Zunahme des Niveaus schon in der Grundschule gehören dabei, anders als es bei MH dargestellt wird, selbstverständlich auch zu den Systemeffekten der Stringenz: Die Verbindlichkeit bildet neben der Kontrolle den einen zentralen Bestandteil des Konzepts. Dass die Effekte schon vor dem Übergang auftreten, besagt nicht, dass es keine Systemeffekte wären, im Gegenteil: Es zeigt die Reichweite von Regelungen auch für den so wichtigen Aspekt der primären Effekte – bis hinein in die Familien und für die so wichtigen frühen Phasen der kognitiven Entwicklung und Einstellung auf das, was da noch kommt. Die Behauptung, es gäbe wegen diese antizipatorischen Effekte danach in der Sekundarstufe keine Wirkungen mehr, lässt sich weder theoretisch, noch auch empirisch halten. Und genau das zeigt sich ja bei ES auch in der Dreifach-Interaktion von Stringenz, kognitivem Niveau und kognitiver Homogenität. Die Verstärkung des Effektes der kognitiven Homogenität in der Sekundarstufe, und das gerade in den Schulklassen mit einem geringeren kognitiven Niveau, findet sich zudem selbst in Analysen, die die Leistungen vorher kontrollieren, sie werden nur – *artifizuell!* – schwächer. Aus Platzgründen kann das hier leider nicht weiter ausgeführt und dokumentiert werden, wohl aber an anderer Stelle und dann ausführlich (Esser 2022, Kapitel 9).

Bach, M., und M. Fischer. 2021. Understanding the Response to High Stakes Incentives in Primary Education. In: IZA-Discussion Paper Nr. 13845. Bonn: Forschungsinstitut Zukunft der Arbeit. Institute for the Study of Labor.

Decristan, J., B. Fauth, M. Kunter, G. Büttner und E. Klieme. 2017. The interplay between class heterogeneity and teaching quality in primary School. In: International Journal of Educational Research, 86: 109-121.

Elwert, F. und Ch. Winship. 2014. Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable. In: Annual Review of Sociology, 40: 31–53.

Esser, H. 2021. “Wie kaum in einem anderen Land ...“? Die Differenzierung der Bildungswege und ihre Wirkung auf Bildungserfolg, -ungleichheit und -gerechtigkeit. Band 1: Theoretische Grundlagen. 2022. Band 2: Empirische Befunde. Frankfurt/M. und New York: Campus (Band 2 im Erscheinen)

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

Esser, H., und J. Seuring. 2020. Kognitive Homogenisierung, schulische Leistungen und soziale Bildungsungleichheit. Theoretische Modellierung und empirische Analyse der Effekte einer strikten Differenzierung nach den kognitiven Fähigkeiten auf die Leistungen in der Sekundarstufe und den Einfluss der sozialen Herkunft in den deutschen Bundesländern mit den Daten der „National Educational Panel Study“ (NEPS). In: Zeitschrift für Soziologie, 49: 277-301.

Grewenig, E., 2021. School Track Decisions and Teacher Recommendations: Evidence From German State Reforms. In: E. Grewenig. Human Capital and Education Policy: Evidence From Survey Data. Ifo Beiträge zur Wirtschaftsforschung, 96, Kapitel 4: 113-161,

Gröhlich, C., K. Scharenberg und W. Bos. 2009. Wirkt sich die Leistungsheterogenität in Schulklassen auf den individuellen Lernerfolg in der Sekundarstufe aus? Inn: Journal of Educational Resaerch (online), 1: 86-105.

Klemm, K. 2021. Zu den Effekten kognitiver Homogenisierung. Kritische Bemerkungen zu den Befunden von Hartmut Esser und Julian Seuring. In: PeDOCS. DIPF.

Lorenz, G., S. Lenz und C. Rjosk. 2022. Effizienz und soziale Ungleichheit in strikt leistungsdifferenzierenden Bildungssystemen Eine kritische Betrachtung des Model of Ability Tracking (MoAbiT), unv. Manuskript, Berlin

Matthewes, S., und J. P. Heisig. 2021. Keine Belege für leistungsfördernde Effekte von strikter Leistungsdifferenzierung und kognitiver Homogenisierung: Eine kritische Reanalyse von Esser und Seuring (2020). Unv. Manuskript. Wissenschaftszentrum Berlin (WZB). Berlin.

Osikominu, A., G. Pfeifer und K. Strohmaier. 2021. The Effects of Free Secondary Track School Choice: A Disaggregated Synthetic Control Approach. In: IZA-Discussion Paper Nr. 14133. Bonn: Forschungsinstitut Zukunft der Arbeit. Institute for the Study of Labor. Rohrer 2018: 38f.):

Rohrer, J. 2018. Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. In: *Advances in Methods and Practices in Psychological Science*, 1: 27-42.