

Peter Bender
Universität Paderborn EIM-Fak
bender@math.upb.de

Erschienen in: Mitteilungen der Gesellschaft für Didaktik der Mathematik (GDM) 91 (2011), 4–12, ISSN 0722-7817,
und in: Profil (Magazin des DPhV) 2011, Heft 7–8, 16–27

Hinweis insbesondere auf den ersten Absatz von **5. Fazit** (S. 11 unten)

Angeblich wissen es die Eltern besser als die Fachleute für Schule

1. Grundsätzliches zu den Übergangsempfehlungen der Grundschule

Unser (ach, so rückständiges) Bildungssystem hält zahlreiche Hindernisse bereit, mit denen vielen unserer hoffnungsvollen Sprösslinge der direkte (= gymnasiale) Weg zum Abitur (und weiter zu der in gewissen Kreisen unentbehrlichen Promotion) erschwert wird. Da wäre insbesondere der Umstand zu nennen, dass es überhaupt außer dem Gymnasium andere Schulformen gibt und dass manchen Eltern – wenn sie denn nicht von selbst darauf kommen – von Seiten der Lehrerschaft nahe gelegt wird, ihr Kind auf eine solche andere Schulform zu schicken! In einigen der 16 Bundesländer hatte bzw. hat dieses Nahelegen als sog. Grundschulempfehlung (GE) sogar einen (mehr oder weniger) verbindlichen Charakter. Infolge der Änderungen der politischen Verhältnisse in jüngerer Zeit im Saarland, in Baden-Württemberg und in Sachsen-Anhalt wurde bzw. wird jetzt dort die Verbindlichkeit abgeschafft. In Nordrhein-Westfalen und in Hamburg unterblieb die Einführung aus unterschiedlichen Gründen. Und so besteht sie demnächst nur noch in Bayern, Brandenburg, Sachsen und Thüringen. Es fällt der hohe Anteil der neuen und der "PISA-Sieger-" Bundesländer auf.

Was ist eigentlich dagegen einzuwenden, dass die Fachleute für Unterricht, Bildung und gesellschaftliche Erziehung eine fundierte Prognose über die günstigste Laufbahn eines Kindes bei Eintritt in die Sekundarstufe abgeben und dass, bei entgegenstehendem Elternwillen, diese Prognose als Entscheidungsgrundlage genommen wird? Zum einen gibt es allerlei Maßnahmen, die einen solchen Konfliktfall auflösen können, z.B. Probeunterricht sowie spätere Übergangsmöglichkeiten. Zum anderen wird ja der Weg zum Abitur auf den nicht-gymnasialen Schulformen nicht wirklich verbaut. Den wenigsten hauptberuflichen Akademikerinnen & Akademikern ist bekannt, dass z.B. im Jahr 2008 über 44% der Studienberechtigungen in Deutschland nicht auf einem Gymnasium erworben wurden.

Im Gegensatz etwa zu Finnland hat in unserer Gesellschaft, geschürt durch verächtlich machende Urteile der einen & des anderen Angehörigen der politischen Klasse (inklusive politisch agierender Mitglieder des Bildungssystems, z.B. Hochschulpräsidenten), der Lehrerstand kein allzu hohes Ansehen. Auch und gerade in Teilen der zeitgenössischen universitären Pädagogik ist man über die mageren Ergebnisse des alltäglichen Schulunterrichts enttäuscht, bleibt dieser i.A. doch weit hinter den eigenen idealen Vorstellungen und Entwürfen zurück. Gerne hängt man dort einem naiven pädagogisch gewendeten Konstruktivismus an, auf dessen Basis die Lehrerschaft (und damit man selbst als deren Ausbilderinnen & Ausbilder) von der Verantwortung für diese Misserfolge frei gesprochen werden kann und man sich auch der Pflicht zur Prüfung der eigenen Vorstellungen und Ent-

würfe etwa auf Schulrealitätsbezug enthoben fühlen kann. Zugleich stellt man damit allerdings auch das Expertentum der Lehrerschaft für ihre zentralen Aufgaben "Unterricht", "Bildung" und "gesellschaftliche Erziehung" in Frage. Und von da ist es nur noch ein kleiner Schritt, ihr die Eignung für die Abgabe von Schullaufbahnprognosen abzusprechen.

Die Kehrseite dieser Medaille ist ein libertinäres Hochhalten des Elternwillens, bemerkenswerter Weise aus einer politischen Richtung kommend, die sonst eher zum Gängelnn neigt. Aus deren Sicht böte sich die Verbindlichkeit der GE eigentlich sogar in besonderem Maße an, weil – bekanntlich – die "Freigabe ... des Elternwahlrechts ... zwangsläufig zulasten der Verstärkung des ohnehin extremen sozialen Bias" ginge (T&BM, S. 657).

Es ist eben das Wesen der GE als ein Instrument des gegliederten Schulsystems, das diese ins Visier der Kämpferinnen & Kämpfer für die Einheitsschule geraten lässt. Es ist interessant zu beobachten, wie die einer großen Lehrgewerkschaft Nahestehenden dabei mit den beiden Konflikten umgehen, (i) dass sie mit ihrer Kritik an der GE indirekt ihrer eigenen Klientel wesentliche Kompetenzen absprechen und (ii) dass sie bereit sind, den von ihnen bemängelten "ohnehin extremen sozialen Bias" beim Übergang auf die Sekundarstufe noch zu verstärken.

Dabei wird heutzutage auch in Sachen "GE" nichts so heiß gegessen, wie es gekocht wird. Ich möchte nicht wissen, in wie vielen Elterngesprächen die Grundschullehrerin & der Grundschullehrer den massiv vorgetragenen Elternwünschen nachgibt und für den Sprössling eine Empfehlung für eine Schulform ausspricht, für die sie & er ihn nicht für geeignet hält. Ob man dem Kind damit einen Gefallen tut, steht zwar in Zweifel; aber als Lehrerin & Lehrer hat man seine Ruhe bis hin zur Vermeidung von juristischen Auseinandersetzungen. Und, wie gesagt, im (etwa nicht gelösten) Konfliktfall bestehen allerhand Möglichkeiten, wie der Elternwille doch noch durchgesetzt werden kann. Es ist heute wirklich nicht mehr wie vielleicht vor vierzig Jahren, als (wie kolportiert wird) ein Professor Schleicher seinen Einfluss bei der Behörde in Hamburg geltend machen musste, damit Söhnchen Andreas trotz verweigerter Empfehlung auf das Gymnasium gehen konnte.

Selbstverständlich steht "die" sog. Bildungsforschung bereit, ihren Beitrag zur Klärung der Sinnfrage der GE zu leisten, indem sie diese Sinn- auf eine numerisch zu erledigende "Erfolgs"frage zurückführt. — Dazu sind mir kürzlich zwei Arbeiten untergekommen, die mir ein guter Freund zur Beurteilung zur Verfügung gestellt hat, der in einer Partei aktiv ist, die dem gegliederten Schulsystem, der Institution des Sitzenbleibens, der Verbindlichkeit der GE u.ä. eher ablehnend gegenübersteht:

Block, Rainer (2006a) (= Ba); dieses graue Papier wurde nach vordergründiger Beseitigung einiger Fehler noch "ordentlich" publiziert als:

Block, Rainer (2006b) (= Bb),

Tiedemann, Joachim & Elfriede Billmann-Mahecha (2010) (=T&BM).

Ich werde mich neben T&BM fairerweise auf Bb konzentrieren, komme aber nicht umhin, auch Ba wenigstens zu streifen, weil man dadurch die Argumentation in Bb erst richtig versteht und weil ich befürchte, dass es die Version Ba ist, die kursiert und als Munition gegen die GE verwendet wird, nach dem Motto "wie Block (2006) gezeigt hat, ...".

Einig sind sich Block und T&BM in der grundsätzlichen Ablehnung der GE, während sie sich in der konkreten Begründung eigentlich diametral gegenüber stehen. Für Block besteht die Fehlerhaftigkeit vor allem darin, dass zu viele Jugendliche für das Gymnasium empfohlen werden, die dann dort scheitern. Dagegen stützen T&BM ihre Kritik darauf, dass sie einige "erfolgreiche" Gymnasiastinnen & Gymnasiasten (G&G) identifiziert haben,

die trotz verweigerter Empfehlung an diese Schulform gegangen waren. — Der Eine stört sich also am Fehler 1. Art, die Anderen stören sich am Fehler 2. Art.

Etwas kurzschlüssig könnte man folgern: Gebt doch einfach keine GE ab; dann könnt ihr keine Fehler machen, weder der 1., noch der 2. Art! — Natürlich würden dann noch viel mehr Fehler geschehen, und zwar vor allem der 1. Art, halt nicht in Form einer falschen GE, sondern einer falschen Laufbahnentscheidung. — Das nächste Argument lautet dann natürlich (und in dessen Dienst steht ja die ganze Kritik an der GE): Schafft doch das gegliederte Schulsystem überhaupt ab; dann könnt ihr gar keine solche Fehler mehr machen. Ich will diese nicht empirisch zu beantwortende Frage nicht vertiefen, sondern nur feststellen, dass die Einheitsschule prinzipiell der (realistischen, organisierbaren, bezahlbaren) individuellen Förderung entgegensteht und dass dort Fehler der o.a. 1. und 2. Art bei der nach wie vor erforderlichen Differenzierung der Schülerinnen & Schüler (S&S) an der Tagesordnung wären.

2. Die Studie von Block

Gestützt auf die Daten von PISA 2000 betrachtet Block vier Populationen von 15-jährigen Jugendlichen: "Realschüler, die von einem Gymnasium gewechselt sind" und "Hauptschüler, die einen Bildungsabstieg hinter sich haben" in Deutschland und die beiden entsprechenden Gruppen in Nordrhein-Westfalen. In Ba rechnet er aus: "das relative Risiko für Realschüler, einer falschen (zu hohen) Schulform zugewiesen zu werden, ist aufgrund einer unzutreffenden Grundschulempfehlung rund 24 Mal größer als aufgrund falscher (überhöhter) elterlicher Bildungsansprüche" (Ba, S. 2, S. 7).

Dabei stehen die "falschen elterlichen Bildungsansprüche" für die Fälle, wo Kinder ohne gymnasiale GE auf das Gymnasium gehen und es wieder verlassen. Diese Identifizierung ist nur scheinbar plausibel, wie ich weiter unten diskutieren werde.

Für die drei anderen Populationen kommt der Autor auf ebenfalls beeindruckende Faktoren 8 bis 9, 20 und 15.

Die Zahl 24 entsteht aus dieser Vierfeldertafel (Ba, S. 7, Bb, S. 154; zum besseren Verständnis habe ich lediglich die Bezeichnungen noch etwas vereinfacht)

15-jährige R&R	GE für Gymn.	keine GE für Gymn.	
aus Gymnasium "abgestiegen"	862 (73,0%)	319 (27,0%)	1 181
von Anfang an in Realschule	637 (10,2%)	5 590 (89,8%)	6 227
	1 499	5 909	7 408

auf folgende Weise: Die angegebenen Prozentzahlen beziehen sich auf die rechte Randverteilung. Nun bildet der Autor die beiden Quotienten $\frac{73,0}{10,2} \approx 7,2$ und $\frac{27,0}{89,8} \approx 0,30$, di-

vidiert diese noch einmal durcheinander und erhält so die Zahl 24, die er wie oben beschrieben interpretiert. Offenbar versteht er den ersten Quotienten als relatives Risiko für Realschülerinnen & Realschüler (R&R), *aufgrund einer GE für das Gymnasium* dieses zu besuchen und es dann wieder (aufgrund von Überforderung) zu verlassen, und den zweiten als relatives Risiko für R&R, *trotz einer fehlenden GE für das Gymnasium* dieses zu besuchen und es dann wieder (aufgrund von Überforderung) zu verlassen.

Die Zahl 24 lässt sich durchaus sinnvoll interpretieren, ist dann allerdings völlig belanglos: Sie entsteht ja auch als Quotient aus den waagrecht gebildeten Quotienten $\frac{73,0}{27,0} \approx$

2,7 sowie $\frac{10,2}{89,8} \approx 0,11$. Sie besagt dann: Unter den "Drop-Outs" sind 2,7 Mal so viele

mit gymnasialer GE als ohne; unter den Jugendlichen, die von Anfang an zur Realschule gehen, sind 9 Mal so viele ohne gymnasiale GE wie mit gymnasialer GE; die Zahl 24 ist das Produkt dieser beiden Faktoren, und insgesamt hat man berechnet: unter den "Drop-Outs" ist das Verhältnis derer mit gymnasialer GE gegenüber denen ohne gymnasiale GE 24 Mal so hoch wie dieses Verhältnis bei den Jugendlichen, die von Anfang an zur Realschule gingen. — Das ist aber nichts Bemerkenswertes, insbesondere nicht der große inverse Quotient 9 der zweiten Zeile.

Allerdings ist der Quotient der ersten Zeile sinnlos, weil schon die Prozentzahlen in dieser Zeile sinnlos sind. Indem man sie so bildet und mit ihnen rechnet, unterliegt man einem *elementaren stochastischen Gedankenfehler, der base-rate-fallacy*. Was man braucht, ist

- (i) die Population *aller* Jugendlichen mit GE für das Gymnasium, die dann auch hingehen, und davon den Anteil der "Drop-Outs" sowie entsprechend
- (ii) die Population *aller* Jugendlichen ohne GE für das Gymnasium, die es trotzdem besuchen, und davon ebenfalls den Anteil der "Drop-Outs"

(in T&BM werden jeweils die richtigen Populationen betrachtet und in Beziehung gesetzt). Bei den Tabellen in Ba und Bb müsste die zweite Zeile komplett ersetzt werden durch die entsprechenden Zahlen der sich am Gymnasium befindenden 15-Jährigen, und die Prozentzahlen müssten bezüglich der unteren Randverteilung gebildet werden:

15-Jährige	GE für Gymn.	keine GE für Gymn.
aus Gymn. in Realsch. 'abgestiegen'	862 (5%)	319 (30%)
immer am Gymnasium geblieben	ca. 16 000 (95%)	ca. 750 (70%)
<i>kursive Zahlen: fiktive Annahmen!</i>		

Ich kenne die Zahlen für die zweite Zeile nicht, sondern habe sie unter Anlehnung an T&BM grob geschätzt. Außerdem kommen mir die Zahlen in Ba und Bb insgesamt erheblich zu klein vor. Möglicherweise handelt es sich bei dem Datensatz nur um eine kleine Teilpopulation, vielleicht derjenigen Jugendlichen, bei denen bekannt ist, welche GE sie hatten. Mir geht es jetzt aber nur um (relative) Größenordnungen.

Jedenfalls erhält man nun sinnvolle Verhältnisse: der erste Quotient $\frac{5}{95} \approx 0,053$ ist das relative Risiko, *aufgrund einer GE für das Gymnasium* dieses zu besuchen und es dann wieder (aufgrund von Überforderung) zu verlassen, und der zweite $\frac{30}{70} \approx 0,457$ ist das re-

lative Risiko, *trotz einer fehlenden GE für das Gymnasium* dieses zu besuchen und es dann wieder (aufgrund von Überforderung) zu verlassen. Der Quotient der beiden Quotienten lautet schließlich $0,1228 \approx \frac{1}{8}$. Das ist der Wert, den der Autor ermitteln wollte,

und er ist eigentlich genau das Gegenteil von dem Wert $\frac{24}{1}$, den er ermittelt hat und auf den er seine zentrale Aussage stützt. Tatsächlich ist das Risiko des "Abstiegs" für Jemanden, die & der mit gymnasialer GE auf das Gymnasium geht, nicht 24 Mal so hoch wie für Jemanden, die & der das ohne gymnasiale GE tut, sondern *im Gegenteil*: Wer ohne gymnasiale GE auf das Gymnasium geht, hat ein (größenordnungsmäßig!) 8 Mal so hohes "Abstiegs"-Risiko, wie wer das mit einer gymnasialen GE tut.

Wie gesagt, in der endgültigen Veröffentlichung hat der Autor seine fehlerhaften Berechnungen und die sehr weit gehenden falschen Aussagen eliminiert. Stattdessen schreibt er: "Nicht beantwortet wird hingegen die Frage – und das soll in aller Deutlichkeit vorangeschickt werden, um möglichen Missverständnissen vorzubeugen – ob die Eltern oder die Grundschullehrer letztlich besser wissen, welche weiterführende Schulform die geeignete für ein Kind ist" (Bb, S. 151).

Doch, die Frage wird beantwortet! Wenn man die Zahlen der G&G hinzunimmt, wird deutlich, dass – statistisch! – die Grundschullehrerinnen & -lehrer es viel besser wissen als die Eltern.

Aber außerdem kann Block tendenziell trotzdem nicht von der Beantwortung dieser Frage gemäß seiner einmal gefassten Überzeugung lassen: "Die Schulformabstiege sind weniger ein Problem übersteigter elterlicher Bildungsansprüche als vielmehr ein Ausdruck der mangelnden prognostischen Zuverlässigkeit von Grundschulempfehlungen" (Bb, S. 155) und "Die Schulformabsteiger weisen zum überwiegenden Teil Grundschulempfehlungen für die Schulformen auf, an denen sie letztlich gescheitert sind" (Bb, S. 160), fast wortgleich noch einmal von dem zitierten Satz von S. 155 gefolgt. Offenbar hat Block seine base-rate-fallacy nicht überwunden. Dies zeigt sich auch daran, dass er nach wie vor seine Tabellen ohne die Grundgesamtheit aller G&G darstellt, die wenig sinnvollen Prozentzahlen zur rechten Randverteilung bildet und diese Prozentzahlen dann von oben nach unten vergleicht (jetzt lediglich, ohne den Doppelquotienten auszurechnen).

Block möchte die Zuverlässigkeit der GE quantifizieren. Dazu verweist er u.a. auf "Forderungen ... in der medizinischen Forschung", wonach "im Zusammenhang mit diagnostischen Tests für Krankheiten ... erst dann von einem guten Diagnoseverfahren" gesprochen werden soll, wenn bestimmte Parameter bestimmte Werte nicht unterschreiten, und zitiert dazu eine Arbeit von 1956 (Bb, S. 157). Er und wir wissen jedoch nicht, wie diese Forderungen zustande gekommen sind, ob sie sich in "der" Medizin oder in Teilen durchgesetzt haben, ob sie heute noch gültig sind bzw. haltbar wären, vor allem aber, ob sie dem Instrument der GE gerecht werden. Immerhin: "Die Validität der Übergangsempfehlungen bewegt sich rein rechnerisch an der Akzeptanzgrenze" (Bb, S. 159); aber "unserer Meinung nach ist es ... letztlich nicht möglich, das ... Instrument Grundschulempfehlung insgesamt auf seine Validität hin präzise zu prüfen" (Bb, S. 160). – In meiner **Skepsis gegenüber der Numerokratie** stimme ich Block hier aus vollem Herzen zu.

Der Vollständigkeit halber sei erwähnt, dass ich Block ebenfalls folge, wenn er einen Großteil der Komplexität des Phänomens der Mobilität zwischen Schulformen (zu denen ja auch die Hauptschule und allerlei Spielarten der Gesamtschule gehören) sowie die Frage des Sitzenbleibens und überhaupt der Messung von Schulerfolg stark reduziert hat (s. aber die u.a. Diskussion von T&BM). Auch halte ich vereinfachte Sprechweisen wie der "'richtige' Realschüler" (Bb, S. 151) usw. für angemessen und hilfreich zum Zwecke der Les- und Verstehbarkeit.

Dem Umfang des Begriffs "übersteigerte Bildungsaspiration der Eltern" (Bb, S. 151), der zentral für die Arbeit ist, liegt m.E. allerdings *ein weiterer schwerer Gedankenfehler* zugrunde. Der Autor identifiziert

(i) die Population der Jugendlichen, die *ohne* gymnasiale GE aufs Gymnasium gehen und wieder zur Realschule zurückkehren, mit der Population derjenigen, die diesen Weg aufgrund einer "übersteigerten Bildungsaspiration der Eltern" gehen, sowie

(ii) die Population der Jugendlichen, die *mit* gymnasialer GE aufs Gymnasium gehen und wieder zur Realschule zurückkehren, mit der Population derjenigen, die diesen Weg aufgrund von "unzutreffenden oder 'falschen' Grundschulempfehlungen" (Bb, S. 151) gehen,

und unterstellt, dass die beiden mit diesen Merkmalen ("übersteigerte elterliche Bildungsaspirationen" vs. "zu hohe GE") gebildeten Populationen disjunkt wären und sein Zahlenwerk dieser Aufteilung entspräche.

Tatsächlich sind die beiden von ihm so beschriebenen Populationen alles andere als disjunkt. Vielmehr hat bei der überwiegenden Mehrzahl der "Drop-Outs" *mit* gymnasialer GE diese GE selbstverständlich den elterlichen Bildungsansprüchen entsprochen, die man folglich ebenfalls als "übersteigert" bezeichnen muss. D.h. so gut wie alle "Drop-Outs" *mit* gymnasialer GE müssten ebenfalls der Population mit "übersteigerten elterliche Bildungsaspirationen" zugeordnet werden. Mehr noch: Nicht wenige GEen für das Gymnasium entstehen auf Druck der Eltern bzw. infolge prophylaktischen Ausweichens vor einem solchen und müssten eigentlich sogar aus der Population mit "zu hoher GE" herausgenommen werden, jedenfalls für den vom Autor angestellten Vergleich. Dieser würde, wenn man denn diesen Sachverhalt in Zahlen fassen könnte, noch ungünstiger für den Elternwillen ausfallen als so schon.

3. Die Studie von Tiedemann & Billmann-Mahecha

Wie Block stellen auch T&BM den Wert der GE in Frage, ebenfalls indem sie den Prognosewert des Elternwillens in einem günstigeren Licht erscheinen lassen (wollen). Während Block relativierend vorgeht und die Fehlerhaftigkeit der GE in den Vordergrund stellt, betonen T&BM, wie "erfolgreich" doch die G&G sind, die keine GE für diese Schulform erhalten hatten. Wegen der relativen Betrachtungsweise ist es bei Block nicht ganz so wesentlich, was er unter "Erfolg" bzw. "Versagen" versteht (Wortwahl nach T&BM, S. 653). Man kann gut mit seinem bewusst stark reduzierten Begriff leben, nämlich dass es lediglich darauf ankommt, ob eine 15-jährige Realschülerin & ein 15-jähriger Realschüler schon einmal auf dem Gymnasium war (eigentlich: ob eine Gymnasiastin & ein Gymnasiast sich mit 15 Jahren immer noch auf dem Gymnasium befindet). – Bei T&BM dagegen, wo ja der absolute Erfolg der G&G ohne gymnasiale GE positiv dargestellt werden soll, kommt es sehr wohl darauf an, was mit "Erfolg" und "Versagen" gemeint sein soll.

Sie stellen die "Schullaufbahnpfehlungen und Schullaufbahn von niedersächsischen Grundschulern des Übertrittjahrgangs 2004/2005 bis zum Übergang in Klassenstufe 7" übersichtlich in einer Tabelle dar, die ich im Folgenden auszugsweise und, um der Kompatibilität mit Blocks Tabelle willen, an der Hauptdiagonalen gespiegelt wiedergebe (S. 653):

nach GS auf Gymnasium	GE für Gymn.	GE für Realsch.	GE für Hauptsch.
"Versagen"	1 339 (5,0%)	1 541 (29,8%)	59 (45,0%)
"Erfolg"	25 660 (95,0%)	3 625 (70,2%)	72 (55,0%)
(nur S&S mit vollst. Daten)	26 999	5 166	131

"Da keine Noten für die Schülerinnen und Schüler vorliegen, wurde als Kriterium für 'Erfolg' die Versetzung in die siebte Klassenstufe des Gymnasiums ohne Klassenwiederholung ... herangezogen" (S. 653). Die Aussage, dass die 131 S&S mit einer Hauptschul-GE "überwiegend erfolgreich" (S. 654) sind, ist dann zwar definitionsgemäß numerisch zutreffend. Aber selbst bei diesem bescheidenen Verständnis von Erfolg haben nach zwei Jahren ("Übergang in Klassenstufe 7") bereits 45% der S&S dieser Kategorie "versagt", indem sie eine Klasse wiederholten oder in eine "niedrigere" Schulform wechselten.

Um einschätzen zu können, ob die 72 verbliebenen S&S mit einer Hauptschul-GE wirklich nennenswert erfolgreich waren, müsste man ihre Noten kennen (ich bin überzeugt, dass sie in ihrer Gesamtheit deutlich unterdurchschnittlich waren). Weiterhin müsste man wissen, wieviel mehr als die Anderen sie für die Schule arbeiten mussten (für Mehrarbeit spricht nicht zuletzt, dass ihre Eltern — mit den Worten von Block — "übersteigerte Bildungsaspirationen" hatten; s.a. T&BM, S. 656), bei wie vielen der "Erfolg" bis zum 10. Schuljahr 2011 anhielt und wie viele letztlich bis 2013 auf dem direkten gymnasialen Weg zum Abitur gekommen sein werden.

Die Studie von Roswitha Urbanek (2006) zeigt, wie eine solche Untersuchung angelegt werden kann, auch wenn sie ebenfalls keine Noten und nicht den Arbeitsaufwand der S&S berücksichtigt. Sie hat für 32 (!) Schülerjahrgänge ihres Gymnasiums in Meinerzhagen Klassenbücher, Kurslisten, Abgangszeugnisse und Personaldaten ausgewertet und ohne Schnickschnack Statistiken zu *langfristigen* "Erfolgen" angefertigt. Da fehlt es natürlich an der Repräsentativität. Anders als bei vielen statistischen Untersuchungen, wo sie wesentlich wäre, aber großzügig ignoriert wird, ist sie hier jedoch keine relevante Kategorie, und der Erkenntnisgewinn kommt aus dem Exemplarischen.

Die Zahlen und Zusammenhänge von T&BM stellen sich etwas freundlicher dar, wenn man R&R mit Hauptschul-GE (38,7% "Versager") sowie G&G mit Realschul-GE (29,8% "Versager") betrachtet. Aber das prinzipielle Dilemma und die genannten Zweifel bestehen auch bei diesen Populationen.

In ihrer Diskussion (S. 657) verlassen T&BM die wissenschaftliche und begeben sich auf die politische Ebene, indem sie für "einen substanziellen Anstieg des Bildungs-Outputs" plädieren, in Frage stellen, ob "in unserem Bildungssystem alle Ressourcen ausgeschöpft" werden und "auf beachtliche Quellen eines ungenutzten Bildungspotenzials" verweisen, in der "Diskrepanz zwischen der GE" und "den elterlichen Bildungsentscheidungen" liegend, "wenngleich nicht alle betroffenen Schülerinnen und Schüler den Abschluss an der leistungsstärkeren Schulform erzielen werden."

Dieses Plädoyer wirft Fragen auf bzw. stößt auf Einwände:

(i) Sollen zusätzlich zu den bereits jetzt durch die abweichende Elternentscheidung über die GE hinaus erschlossenen Bildungsressourcen weitere erschlossen werden, indem sich noch mehr Eltern über die GE hinwegsetzen (womit der "Versager"-Anteil vermutlich überproportional erhöht würde)?

(ii) T&BM begrüßen eine Steigerung der Realschulabschluss- und Abiturquote um 5% durch die von der GE abweichende Elternentscheidung. Sie rechnen nicht dagegen die persönlichen und familiären Katastrophen bei den "Versagerinnen & Versagern" sowie die zusätzlichen privaten und öffentlichen Kosten, die vermieden würden, wenn diese Kinder gleich auf die adäquate Schulform kämen.

(iii) Wenn die Quoten "höherer" Abschlüsse durch eine Erhöhung entsprechender S&S-Zahlen gesteigert werden sollen, warum soll das gerade auf der Basis des bekanntlich sehr fehleranfälligen und sozial besonders stark auslesenden Elternwillens geschehen? — Man könnte doch den Grundschullehrerinnen & -lehrern aufgeben, dass sie zunächst realistische GEen verfassen und dann noch die nächsten 10% der nicht empfohlenen Kinder zusätzlich für die Realschule oder das Gymnasium benennen. Dieses Verfahren wäre weniger fehleranfällig und sozial auslesend und würde ebenfalls den Zweck der Ausschöpfung aller Bildungsressourcen verfolgen. (Ich schlage das nicht ernsthaft vor, sondern bringe nur die Argumentation von T&BM auf den Punkt.)

4. Diskussion

Gewiss gibt es heute absolut und relativ viel mehr Weiße- als Blaue-Kragen-Werktätige als etwa vor 60 Jahren, wird als Voraussetzung für viele Ausbildungsberufe, für die man damals mit einem Volksschulabschluss mit 14 Jahren gut geeignet war, heute der Realschulabschluss oder das Abitur gefordert und hat man mit diesen "höheren" Abschlüssen einfach mehr Chancen als ohne sie. Aber auch ein ordentlicher Hauptschulabschluss wird in vielen Ausbildungsbetrieben nach wie vor estimiert, etwa weil die & der Auszubildende besser "passt", hoffentlich keine "höheren" Ambitionen hat und nicht das Klima beeinträchtigt bzw. nicht nach kurzer Zeit wieder weg ist.

In der ganzen Debatte um unser Schulsystem gewinne ich immer wieder den Eindruck, dass für manche Beteiligten der Mensch erst mit dem Abitur beginnt. Ich kann zwar akzeptieren, dass Block sowie T&BM um der besseren Lesbarkeit willen von "Erfolg", "Versagen", "Auf-" und "Absteigen" reden, und ich mache ja bei dieser Rede mit; aber wirklich wohl ist mir bei diesen Wörtern nicht, weil für mich Haupt- und Realschule nicht weniger wert oder "niedriger" als das Gymnasium sind, sondern *andere* Bildungsgänge darstellen, zumal auch sie wirklich alle Laufbahnmöglichkeiten bis zum Studium offen halten.

Es besteht m.E. keine Notwendigkeit, noch die letzten 5% "unserer Bildungsressourcen" in Form einer höheren Abiturquote "auszuschöpfen", auch wenn die OECD und interessierte Kreise in Deutschland diese uns fortwährend einreden. Die zu gewärtigen negativen Begleiterscheinungen inklusive Kosten habe ich bereits angesprochen. Zu ergänzen wäre noch: eine weitere Absenkung des Niveaus und damit einhergehende Entwertung des Abiturs, Ausbreitung der scheinbaren Überqualifizierung für mancherlei Berufe sowie Überfüllung unserer Hochschulen mit Studierunfähigen.

Als studierter Grund- und Hauptschullehrer strebe ich nicht an, das Gymnasium gegen die "Anderen" abzuschotten, sondern ich möchte der seit Jahren stattfindenden Abwertung der nicht-gymnasialen Bildungsgänge durch Schlechtreden und des Gymnasiums durch Niveausenkung entgegenwirken.

Leider ist die Debatte ideologisch überfrachtet, und zwar bis hin zu vielen quantitativen empirischen Arbeiten. Neben den diversen Durchgängen von TIMSS, PISA und IGLU habe ich einige kleinere Untersuchungen studiert und dabei zahlreiche Aussagen gefunden, die sich nicht aus dem analysierten Zahlenmaterial ableiten lassen, sondern der Überzeugung der Autorinnen & Autoren entspringen, aber wirken (sollen), als ob sie "empirisch belegt" seien.

Der Kardinalfehler der **fehlenden Repräsentativität und Unabhängigkeit** der Stichprobe tritt eher in fachdidaktischen Studien zum Erfolg von unterrichtlichen Maßnahmen auf. Wenn man vier Klassen an einer Schule auswertet, hat man keine repräsentative Stichprobe mit $N > 100$ unabhängigen Probandinnen & Probanden, und da nützt es auch wenig, mit SPSS irgendwelche Signifikanzniveaus auszurechnen. — Überhaupt ist **erkenntnistheoretisch kritisch**, wie gut die in der Praxis des produktiven Gewerbes entwickelten **Methoden** des Hypothesentests u.a. **für die Bildungsforschung geeignet** sind, zumal dort die tatsächliche Vorgehensweise infolge der regelmäßig erst nachträglich vorgenommenen Ermittlung des Signifikanzniveaus eine ganz andere ist. Auch die **mathematische und sachliche Korrektheit** so mancher raffinierten Methode steht durchaus **in Frage**. Allerdings entziehen sich die Statistik-Expertinnen & -Experten gerne dieser Befragung, indem sie die verwendete Methode, wenn überhaupt, nur schwer verständlich darstellen und die (relativen) Laiinnen & Laien nach dem Schibboleth-Prinzip (Jahnke 2010) aus dem Kreis der Wissenden ausschließen.

Entsprechend trifft man immer wieder auf Wörter und, tiefer gehend, Begriffe, die in gewissen Kreisen (Ökonometrie, Soziometrie, Psychometrie, "Bildungsforschung") in einer gewissen Bedeutung gebraucht werden oder von Autorinnen & Autoren für einen gewissen Zweck definiert werden, aber von (mehr oder weniger) gebildeten Laiinnen & Laien insbesondere in der (Bildungs-) Politik in einer natürlichen (oder auch missbräuchlichen) Weise anders verstanden werden. Dabei tun die Autorinnen & Autoren oft zu wenig, um dem falschen Verständnis entgegenzuwirken, sei es, dass sie die Problematik gar nicht erkennen, sei es, dass ihnen eine solche Popularisierung ihrer Ausführungen (aus Eitelkeit, zur Sicherung von Stellen, zur Stützung politischer Überzeugungen) durchaus recht ist, sei es, dass sie dem falschen Verständnis selbst unterliegen, sei dies das Ergebnis einer gesteuerten oder einer unkontrollierten Autosuggestion. Welche dieser Motive jeweils wirksam sind, lässt sich kaum feststellen. Aus wissenschaftstheoretischer Sicht sind sie alle nicht in Ordnung. Beispiele:

"**Chance**" bedeutet im Englischen einfach "Wahrscheinlichkeit". Indem er wörtlich ins Deutsche übernommen wird, erhält dieser Begriff eine soziale Konnotation. Es geht z.B. auf einmal nicht mehr nur um die Wahrscheinlichkeit, das Abitur zu erreichen, sondern um Chancen, die Jemand (die Gesellschaft, die bürgerliche Klasse) Jemandem (aus bildungsfernen Familien) gewährt oder vorenthält.

Besonders extrem kam mir der statistische (!) Nachweis der **kausalen Abhängigkeit** der Kriminalität von fehlendem Hauptschulabschluss durch Entorf & Sieger (2010) vor, der darin bestand auszurechnen, dass ein gewisser Parameter einen gewissen Schwellenwert überschreitet (s. dazu Bender 2011). In gewissen Zirkeln mag man solche Wörter abkürzend verwenden (vielleicht weil man mit der Kraft der Sprache etwas erreichen oder suggerieren will, was man aus der Sache selbst nicht herausholen kann), und ein Teil der Angehörigen dieser Zirkel versteht die Wörter bestimmt "richtig". Aber spätestens wenn man an die breite (Medien-) Öffentlichkeit geht, ist die Verwendung solcher Wörter nicht mehr angebracht, auch dann nicht, wenn man sie irgendwo im Text "erklärt", weil diese Erklärungen nicht nur regelmäßig nicht gelesen werden (selbst wenn sie – selten genug – verständlich sind), sondern gegen knappe Überschriften und reißerische Pressemeldungen überhaupt keine Chance zur Relativierung haben.

Ebenfalls missbraucht wird gerne die Rede von **Erfolg**. Natürlich kann man, völlig neutral, von den für ein Ereignis "günstigen" Ausfällen reden. Aber mit den Studien, die mir untergekommen sind, soll Politik gemacht werden, sei es von den Autorinnen & Autoren selbst, sei es von ihren Auftraggeberinnen & -gebern, sei es von Nutzerinnen & Nutzern in der Politik. Da sollte man schon genau hinschauen, was da unter Erfolg verstanden wird. Z.B. habe ich genau die Quellen, die Klemm (und vorher Tillmann) für seine Behauptung der Unwirksamkeit des Sitzenbleibens herangezogen hat, umgekehrt gelesen: Durch das Sitzenbleiben kommen viele S&S dauerhaft auf ausreichende Leistungen und erwerben Abschlüsse, die ihnen sonst vorenthalten geblieben wären (s. Bender 2010, S. 15).

Mit dem Erfolgsbegriff von T&BM habe ich mich oben auseinandergesetzt.

Neben (möglichen) ideologischen Gründen gibt es ganz simple ökonomische, praktische Gründe, warum Autorinnen & Autoren ihre Begrifflichkeit in einer ganz bestimmten Weise wählen, z.B. damit sie auf ihre Stichprobe passt. Es ist ja verständlich, dass man oft nicht den Aufwand leisten will (oder kann), für relevante Fragestellungen ein geeignetes Stichproben-Design zu entwickeln und dann die Daten zu erheben. Da bietet es sich an, Datensätze von statistischen Ämtern oder von PISA auszuwerten; und diese sind dann vielleicht nicht detailliert genug oder führen zu **unpassenden Stichproben**.

Block übersieht, dass auch bei Eltern mit Kindern *mit* gymnasialer GE zumeist eine hohe Bildungsaspiration vorliegt, die man bei "Misserfolg" der Kinder konsequenter Weise "überhöht" nennen müsste. D.h. die entsprechende Population der Familien mit "überhöhter Bildungsaspiration" ist viel, viel größer als die von Block als solche angenommene und endet keineswegs da, wo die gymnasiale GE beginnt. Das war aber die Grenze, die anscheinend der PISA-Datensatz geliefert hat, und nun hätte Block bei allen Kindern *mit* gymnasialer GE die Bildungsaspiration der Eltern sowie das Zustandekommen dieser GE erheben müssen. Zugegeben, dieses Vorgehen ist praktisch nicht leistbar und würde natürlich auch wieder Probleme der Begrifflichkeit aufwerfen, aber es allein wäre der Fragestellung angemessen gewesen.

Zur **Datengewinnung** muss man oft auf **Selbstauskünfte** zurückgreifen. Wir alle wissen um deren Schwächen, und trotzdem werden sie, wenn sie einmal erhoben sind, gerne als "harte" Daten behandelt. Anzuerkennen ist hier auf jeden Fall die Primärforschungsleistung. Aber aus PISA weiß man z.B., dass nur 40% aller 15-Jährigen in Deutschland den genauen Beruf ihres Vaters nennen können (und auch nur 70% zutreffend antworten, wenn sie den Beruf *ungefähr* bezeichnen sollen). Mit Hilfe solcher weichen Daten werden dann Parameter wie "sozialer Gradient" oder "Wahrscheinlichkeit des Gymnasialbesuchs" berechnet, wo Deutschland bzw. Bayern (zunächst) schlecht abschneiden, die aber so empfindlich sind, dass einige Jahre später und/oder bei leichter Abwandlung der Definition die Reihenfolge ganz anders aussieht (was aber zu interessierten Kreisen, die sich einmal – manchmal hämisch – ihre Meinung gebildet haben, dann oft nicht mehr durchdringt).

Bei Ökonometrikerinnen & -metrikern sind **Zeitreihen** beliebt. So haben (s. Bender 2010) Wößmann & Koautoren 36 internationale Vergleichstests aus den Jahren 1964 bis 2003 ausgewertet und "störungsfrei" bis 2090 (und darüber hinaus) fortgeschrieben, nach dem Motto "das geht immer so weiter". Die stochastische Qualität der vergangenen Studien wurde ganz gewiss nicht berücksichtigt. Wir kennen sie in den wenigsten Fällen (was gegen sie spricht). Z.B. weiß man jedenfalls von FIMSS und SIMSS, dass sie, nicht nur aus fachdidaktischer Sicht, wahrlich keine großen Würfe waren.

Wie absurd diese Fortschreibung von Zeitreihen sein kann, sieht man schön an folgendem Beispiel.

Viele Menschen in Deutschland haben Sorge vor der Überalterung unserer Gesellschaft, wie sie in folgenden Zahlen zum Ausdruck kommt: Während heute 100 Erwerbstätigen 34 Menschen über 65 Jahre gegenüberstehen, wird im Jahr 2050 dieses Verhältnis 100 : 65 lauten. – In ihrem Buch "**Lügen mit Zahlen**" (2010) (Besprechung im Wirtschaftsteil der Frankfurter Rundschau (FR) am 24.01.2011) "entkräften" die Autoren Bosbach & Korff diese Sorge mit der Begründung, dass im Jahre 1950 dieses Verhältnis 100 : 17 betragen hatte und dass bis heute ja auch eine Verdopplung stattgefunden hat, "ohne eine Katastrophe auszulösen".

Dieses Argument erinnert mich an den See, von dem ein Teil der Oberfläche mit Algen bedeckt ist. Die bedeckte Fläche verdoppelt sich jeden Tag. Wenn dann der See eines Tages halb bedeckt ist, ist das angeblich kein Anlass zur Sorge, denn die tägliche Verdopplung hat ja auch in der Vergangenheit stattgefunden, "ohne eine Katastrophe auszulösen". – Es mag ja sein, dass unsere Gesellschaft auch im Jahr 2050 mit der dann bestehenden Altenquote "fertig" wird; – aber bestimmt nicht deswegen, weil wir bis 2010 schon einmal eine Verdopplung der Quote hatten.

Zum Schluss noch eine besonders schöne und stinkende Sumpflüte von **mangelhafter Datenerhebung**, durchgeführt von einer seriösen Institution mit politischem Einfluss, dem

Deutschen Institut für Wirtschaftsforschung (DIW) zum Thema "Kinderarmut in Deutschland" (FR vom 07.05.2011: "Statistisches Schrumpfen").

Menschen "gelten als arm, wenn sie in Haushalten leben, die mit weniger als 50% des mittleren Einkommens zurechtkommen müssen". Diese Definition ist natürlich fragwürdig; denn bei einer Verzehnfachung oder auch Zehntelung des Einkommens aller Mitglieder einer Gesellschaft gelten genau dieselben Menschen wie vorher immer noch als arm. Und in der Tat kann man Kinderarmut in Deutschland nicht wirklich vergleichen mit Kinderarmut z.B. in Bangla Desh. Die OECD nimmt solche Vergleiche aber dennoch vor (mit den OECD-Ländern), u.a. gestützt auf die DIW-Daten, und stellt Deutschland seit vielen Jahren an den Pranger, weil seine Kinderarmutsquote immer ein paar Prozentpunkte über dem OECD-Durchschnitt von etwa 12 bis 13% lag, insbesondere tat sie das auch im Wahlkampf kurz vor der Bundestagswahl in Deutschland 2009.

Inzwischen hat das DIW seine Daten revidiert (und zwar i.W. halbiert). 2004 betrug die Kinderarmutsquote in Deutschland tatsächlich 10%, 2008 sogar nur 8%. Die Änderung kam folgendermaßen zustande: Jedes Jahr werden 20 000 Menschen über ihre Lebenssituation befragt, und zwar immer dieselben (sozioökonomisches Panel SOEP). Wenn in einem Haushalt mit mehreren Erwachsenen eine Person keine Auskunft gibt, hat man früher ihr Einkommen einfach auf Null gesetzt. Inzwischen wird in solchen Fällen das Einkommen aufgrund gewisser Anhaltspunkte geschätzt, und schon haben viele Familien ein um so viel höheres Gesamteinkommen, dass sie nicht mehr arm sind, und die Kinderarmutsquote nimmt entsprechend ab.

Wann das DIW seine Erhebungsmethode geändert hat, geht aus dem Artikel in der FR nicht hervor. Bei der OECD ist diese Änderung erst nach 2009 angekommen. Dieses SOEP ist angeblich eine der besten Datengrundlagen für einschlägige Studien, z.B. auch zur Bildungsforschung, auch im internationalen Maßstab. Obwohl das DIW behauptet, dass die beschriebene Änderung der Datenerhebung außer auf die Kinderarmutsquote keine nennenswerten Auswirkungen hätte, so müssten m.E. dennoch zahlreiche Studien – vermutlich auch PISA – daraufhin überprüft werden, wie weit sich die Parameter ändern und ob so manche Interpretation und politische Aussage heute noch haltbar ist bzw. jemals zugetroffen hat.

Dieser Vorfall erinnert mich an die Erzählung von Heinrich Böll "Die Waage der Baleks". Da hat eine hochherrschaftliche Familie ein Messinstrument (eine Waage) gefälscht, über viele Jahre hinweg das Vertrauen der einfachen Menschen missbraucht, sie regelmäßig um einen Teil des ihnen zustehenden Entgelts betrogen und sich auch noch dazu berechtigt gesehen. Das DIW hat natürlich andere Motive (Ökonomie der Datenerhebung) als diese mit Arroganz gepaarte Geldgier der Baleks. – Für einen **handfesten Skandal** halte ich diesen Vorfall dennoch, und man wundert sich, warum er in den Medien nicht hochgespielt wird. Vermutlich ist, unabhängig vom Wahrheitsgehalt, eine Aussage wie "die Kinderarmut in Deutschland ist relativ hoch" politisch korrekter als eine Aussage wie "die Kinderarmut in Deutschland ist relativ niedrig".

5. Fazit

Den politischen Überzeugungen "für die Einheitsschule", "gegen das Sitzenbleiben", "gegen die Verbindlichkeit der Grundschulempfehlung" und weiteren in der etablierten Erziehungswissenschaft herrschenden Meinungen (z.B. die Pädagogisierung des erkenntnistheoretischen Konstruktivismus) lässt sich im Grundsätzlichen ja durchaus etwas abgewinnen. Anzuerkennen ist darüber hinaus, dass die Vertreterinnen & Vertreter sich immer wieder um eine empirische Fundierung (in irgendeinem Sinn) dieser ihrer Überzeugungen

bemühen. Allerdings erfüllen die zu diesem Zweck angestellten und veröffentlichten Untersuchungen wegen allerlei methodologischer, methodischer und erkenntnistheoretischer Mängel diesen Anspruch oft nicht. Solche Mängel würde man vermutlich jedes Jahr bei vielen tausenden "empirischen" Arbeiten (nicht nur) in der sog. Bildungsforschung finden. Man kann (und sollte) diese auf sich beruhen lassen. Wenn jedoch mit solchen Arbeiten Politik gemacht werden soll, ob von den Autorinnen & Autoren selbst oder von Anderen, dann ist eine kritische methodologische, methodische und erkenntnistheoretische Prüfung angezeigt.

Überhaupt gilt: Eigentlich darf man keiner Statistik glauben, die man nicht selbst gemacht (gefälscht) oder wenigstens auf Herz und Nieren geprüft hat. Schon gar nicht darf man sich auf Sätze verlassen wie "wie X&Y gezeigt haben, ...".

Literatur

Bender, Peter (2010): Kosten- und Leistungsrechnung der Bertelsmann-Stiftung und der OECD für das deutsche Bildungssystem. In: Mitteilungen der GDM 89, 13–21, und in: Profil 2010, Heft 6, 10–21

Bender, Peter (2011): Endlich wissenschaftlich nachgewiesen: Die Hauptschule ist an Allem schuld. In: Mitteilungen der GDM 90, 4–6, und in: Profil 2011, Heft 3, 12–15

Block, Rainer (2006a) (= Ba): Schulrecht vor Elternrecht? Neue empirische Befunde zur Zuverlässigkeit von Übergangsempfehlungen der Grundschule. Universität Essen, Campus Essen, Fachbereich Bildungswissenschaften, Arbeitsgruppe Bildungsforschung/-planung: Manuskript

Block, Rainer (2006b) (= Bb): Grundschulempfehlung, elterliche Bildungsaspiration und Schullaufbahn. In: Die deutsche Schule 98, 149–161

Bosbach, Gerd & Jens Korff (2011): Lügen mit Zahlen. München: Heyne

Entorf, Horst & Philip Sieger (2010): Unzureichende Bildung: Folgekosten durch Kriminalität. Gütersloh: Bertelsmann-Stiftung

Jahnke, Thomas (2010): Kritik empirischer Unvernunft – zur sogenannten quantitativen empirischen Bildungsforschung in der Mathematikdidaktik. Potsdam: Manuskript, 10.05. 2011: http://www.math.uni-potsdam.de/prof/o_didaktik/aa/Publ/Jahnke_Empirikkritik.pdf

Tiedemann, Joachim & Elfriede Billmann-Mahecha (= T&BM) (2010): Wie erfolgreich sind Gymnasiasten ohne Gymnasialempfehlung? Die Kluft zwischen Schullaufbahneempfehlung und Schulformwahl der Eltern. In: Zeitschrift für Erziehungswissenschaft 13, 649–660

Urbanek, Roswitha (2006): Grundschulurteil – besser als sein Ruf. In: Landeselternschaft der Gymnasien in NRW 184, 22–24