

Kritik empirischer Unvernunft - zur so genannten quantitativen empirischen Bildungsforschung in der Mathematikdidaktik

Thomas Jahnke

Ich danke, dass ich vor Ihnen hier sprechen darf, und im Voraus für Ihr Gehör, das ich nach Kräften strapazieren werde. Ich will Ihnen allerdings nicht verhehlen, dass ich mit einiger psychischer Anstrengung zu Ihnen spreche gleichsam mit Davidgefühlen. Wie wollte sich ein einzelner schon gegen eine geradezu serielle Produktion von wissenschaftlichen Erkenntnissen behaupten, die wenigstens zum Teil gut finanziert und begutachtet und häufig unter einer Vielzahl von zum Teil bedeutenden Namen erschienen? Das scheint mir unmöglich.

Schon ein Titel wie „Kritik der quantitativen empirischen Bildungsforschung“ verweist auf ein Manko. Sollten nicht wissenschaftliche Erkenntnisse gleich in welchem Metier Ergebnis einer kritischen Auseinandersetzung mit der untersuchten Sache sein, die Kritik also der Untersuchung immanent und nicht etwas, was hier hinterher getragen wird oder werden muss? Es mag wohl einzelne Bedenken oder Gegenargumente gegen die auf solche Weise erarbeiteten Resultate geben, die die Involvierten vielleicht vergaßen oder vorsätzlich unberücksichtigt ließen, aber die ganze Untersuchung selbst sollte – so weit nur möglich – bereits Gegenpositionen durchdacht und widerlegt oder aus guten Gründen verworfen haben; eine Untersuchung wird wohl auch dadurch wissenschaftlich, dass sie sich selbst und ihren eigenen Methoden gegenüber kritisch ist und sich letzteren nicht willfährig oder fraglos erfolgsblind unterwirft. Das scheint mir bei der quantitativen empirischen Bildungsforschung in ganz eigener Weise der Fall zu sein. Es liegt eine eigenartige Verkapselung von Erkenntnis und Interesse vor, die sich gegen jegliches Andersdenken verschlossen hat.

Ganze Bündel von Bedenken lassen sich aufführen:

- erkenntnistheoretische,
- methodologische und methodische,
- solche, die sich auf die Forschungspraktiken beziehen, etwa die Folgen aus der Globalisierung der Forschung bei gleichzeitig arbeitsteiligem Vorgehen der Experten,
- bildungspolitische und schließlich
- solche, die sich aus den gravierenden Folgen der Forschungsergebnisse auf die Lehrerinnen und Lehrer, die Schülerinnen und Schüler und Schulunterricht ergeben.

Ich bin nicht in der Lage, alle diese Dimensionen hier sorgfältig aufzuführen oder ihnen in säuberlicher Trennung nachzugehen, zumal sie sich ja gegenseitig beeinflussen und durchdringen. Ich greife hier nur Einzelnes heraus, dass sich bei aller Unvollständigkeit und Grobheit doch zu einem Ganzen fügen kann.

Die Penetration

Das Eindringen der empirischen Bildungsforschung und deren Methoden in die deutsche Mathematikdidaktik vollzog sich eher schleichend oder mit dem Zeitgeist eilend, jedenfalls ohne einen Auftakt oder den Versuch einer Begründung, welche neuen Erkenntnishorizonte und -momente diese Methoden erschließen. Die ‚neuen‘ Methoden waren vielmehr einfach da, als zöge gleichsam die – von der Forschung offensichtlich bislang vernachlässigte, übersehene oder sogar negierte – Realität in die Mathematikdidaktik ein und nicht eine neue Sichtweise, die einer Legitimation und Charakterisierung ihrer Erkenntnisweise bedürfte. Historisch setzte diese Penetration in der und in die deutsche Mathematikdidaktik mit der Globalisierung der Forschung durch die internationalen Vergleichsuntersuchungen von Schülerleistungen also mit TIMSS und PISA ein. Die Wucht und mediale Wirkung, die diese Untersuchungen und ihre Resultate in Deutschland entfalteten, ließen die neuen Methoden als selbstverständlich (und) wirksam erscheinen, ohne dass sie in irgendeiner Weise erkenntnistheoretisch reflektiert, ja überhaupt als Methoden diskutiert wurden.

In der Einleitung der Schrift ‚TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich – Deskriptive Befunde‘ [Baumert/Lehmann 1997] verkünden die Forscher bereits in ersten Sätzen vollmundig:

Mit dem vorliegenden Buch berichten wir in deskriptiver Absicht über die Ergebnisse der *Dritten Internationalen Mathematik- und Naturwissenschaftsstudie – The Third International Mathematics and Science Study (TIMSS)* [...]. Der Bericht stellt Befunde für die Bundesrepublik Deutschland im internationalen Vergleich dar. Damit werden seit vielen Jahren zum ersten Mal wieder Daten verfügbar, die für die mathematisch-naturwissenschaftlichen Fächer Auskunft über das in ausgewählten Jahrgangsstufen erreichte Wissens- und Fähigkeitsniveau geben.

Mit dem Wort ‚Befund‘ wird hier bewusst eine naturwissenschaftlich-medizinische Assoziation evoziert, die noch durch das Epitheton ‚deskriptiv‘ verstärkt wird; die Formulierung ‚Auskunft geben‘ verweist auf Fakten, die jetzt ‚verfügbar‘ sind wie ‚Daten‘

aus einem neu erstellten Telefonbuch oder solche über Ebbe und Flut oder die Umlaufzeiten der Planeten. Auch die Begriffe ‚Wissens- und Fähigkeitsniveau‘ erscheinen in diesem Zusammenhang nicht als theoretische Konstrukte sondern als bare Münzen, die man nur gezählt, oder wie Luftdruckwerte, die man in verschiedenen Ländern gemessen hat. Schon von der sprachlichen Diktion begegnen wir hier nicht einer diskutierbaren Sichtweise oder Theorie, die einer Begründung bedürfte, sondern schlicht der Realität. Folgerichtig wird auch nicht gedacht oder interpretiert, sondern nur ‚berichtet‘ – wie von einer naturwissenschaftlichen Expedition oder einem Experiment. Im gleichen Jargon heißt es später bei PISA, die Kompetenzstufen seien dazu da, die ‚Skalen zum Sprechen zu bringen‘, als gäben Daten ohne jede Interpretation oder theoretische Rahmung von sich aus ‚Auskunft‘ wie ein Fieberthermometer über den Zustand eines Patienten, wobei den Ärzten nur bleibt, die auftretenden Temperaturen in sehr niedrige, niedrige, intermediäre, hohe und sehr hohe zu klassifizieren, um sich untereinander zu verständigen und die Betroffenen und deren Verwandten und Pfleger über die Schwere der Krankheit informieren zu können.

Der vergessene Streit

Anfang der siebziger Jahre des letzten Jahrhunderts hat in der deutschen Soziologie der so genannte Positivismusstreit stattgefunden, in dem es auch um eine Debatte um den Erkenntniswert empirischer Sozialforschung ging.

Ich zitiere dazu aus einem Aufsatz von Theodor W. Adorno:

Nicht umsonst [...] überwiegen in den Diskussionen der empirischen Sozialforschung Methodenfragen gegenüber den inhaltlichen. Anstelle der Dignität der zu untersuchenden Gegenstände tritt vielfach als Kriterium die Objektivität der mit einer Methode zu ermittelnden Befunde, und im empirischen Wissenschaftsbetrieb richten sich die Auswahl der Forschungsgegenstände und der Ansatz der Untersuchung, wenn nicht nach praktisch-administrativen Derivaten, weit mehr nach den verfügbaren und allenfalls weiterzuentwickelnden Verfahrensweisen als nach der Wesentlichkeit des Untersuchten. Daher die unzweifelhafte Irrelevanz so vieler empirischer Studien. Das in der empirischen Technik allgemein gebräuchliche Verfahren der operationellen oder instrumentellen Definition, das [etwa] Kategorien [wie >Konservatismus<] definiert durch bestimmte Zahlenwerte der Antworten auf Fragen innerhalb der Erhebung selbst, sanktioniert den Primat der Methode über die Sache, schließlich die Willkür der wissenschaftlichen Veranstaltung. Prätendiert wird, eine Sache durch ein Forschungsinstrument zu untersuchen, das durch die eigene Formulierung darüber entscheidet, was die Sache sei: ein schlichter Zirkel. Der Gestus wissenschaftlicher Redlichkeit, der sich weigert, mit anderen Begriffen zu arbeiten als mit

klaren und deutlichen, wird zum Vorwand, den selbstgenügsamen Forschungsbetrieb vors Erforschte zu schieben.

[...] Sobald dann, wie es fast unvermeidlich ist, von den instrumentell definierten Begriffen auch nur auf die konventionell üblichen extrapoliert wird, macht sich die Forschung eben der Unsauberkeit schuldig, die sie mit ihren Definitionen ausrotten wollte.
[Th. W. Adorno 1972]

Nach meiner Ansicht trifft diese kritische Charakterisierung empirischer Forschung mit den angesprochenen Ingredienzen auf fast alle empirischen Untersuchungen in der Mathematikdidaktik zu, und ich bezweifele sogar, dass bei ihnen der empirischen Methode selbst mit dem Ernst und der Sorgfalt nachgegangen wurde und wird, mit denen sie hier kritisiert wird. Solcher Kritik gegenüber ist die empirische Bildungsforschung wissenschaftlich taub und geschäftig ignorant.

Der vergessene Sinn

Sinnfragen sind empirisch nicht entscheidbar, sie lassen sich nur in einem theoretischen Rahmen überhaupt aufwerfen und bearbeiten. Das wissen zumindest implizit auch die empirischen Bildungsforscher, wenn sie sie in einem ersten Schritt, der in ihrem Jargon als Konzeptualisierung bezeichnet wird und der zumeist durch einen Verweis auf irgendeinen undurchsichtigen und kaum erläuterten Begriffsapparat erledigt wird, kurz abtun. Bei PISA wird zum Beispiel Hans Freudenthal herangezogen, obwohl der sich solchen Untersuchungen gegenüber bereits vor vielen Jahrzehnten mehr als kritisch geäußert hat.

In anderen Arbeiten wird ohne jeden Seitenbezug auf kürzere und längere Theorieversatzstücke aus verschiedenen Arbeiten verwiesen, deren Kern weder referiert noch wenigstens paraphrasiert wird, aber dazu dient, irgendeinen Bereich mit wenigen disjunkten Begriffen zu überziehen, die dann in Fragebogenitems, über die man allenfalls beispielhaft fast nichts erfährt, operationalisiert werden.

Bei COACTIV [Krauss/Kunter 2004] liest sich das zum Beispiel so:

Zur Beschreibung der kognitiven Kompetenzen der Mathematiklehrkräfte im engeren Sinn (Professionswissen) greifen wir auf die von Shulman (1986) eingeführte und von Bromme (1997) erweiterte Unterscheidung von Fachwissen (*content knowledge*), fachdidaktischem Wissen (*pedagogical content knowledge*) und allgemeinem pädagogischen Wissen (*pedagogical knowledge*) zurück und ergänzen diese Bereiche des spezifischen Organisations- und Interaktionswissens (Sternberg & Horvath 1995; Hiebert, Gallimore & Stigler 2002; Fried 2003) sowie um das Beratungswissen, das zur Kommunikation mit Laien erforderlich ist

(Bromme, Jucks & Rambow, 2000¹). Professionswissen von Lehrkräften, so unser theoretischer Ausgangspunkt, umfasst diese fünf Kompetenzbereiche (...).

Hier werden Arbeiten, die sich zum Teil gar nicht auf das ‚Professionswissen von Lehrkräften‘ beziehen, aus zwei Jahrzehnten mühe- und kommentarlos zu einem ‚theoretischen Ausgangspunkt‘ versammelt, der überdies gleich wieder eingeschränkt wird, so dass es wieder nur beim Stand von Shulman 1986 bleibt:

Im Rahmen von COACTIV konzentrieren wir uns auf die drei unterrichtsnäheren Bereiche Fachwissen, fachdidaktisches Wissen und allgemein pädagogisches Wissen von Mathematiklehrkräften.

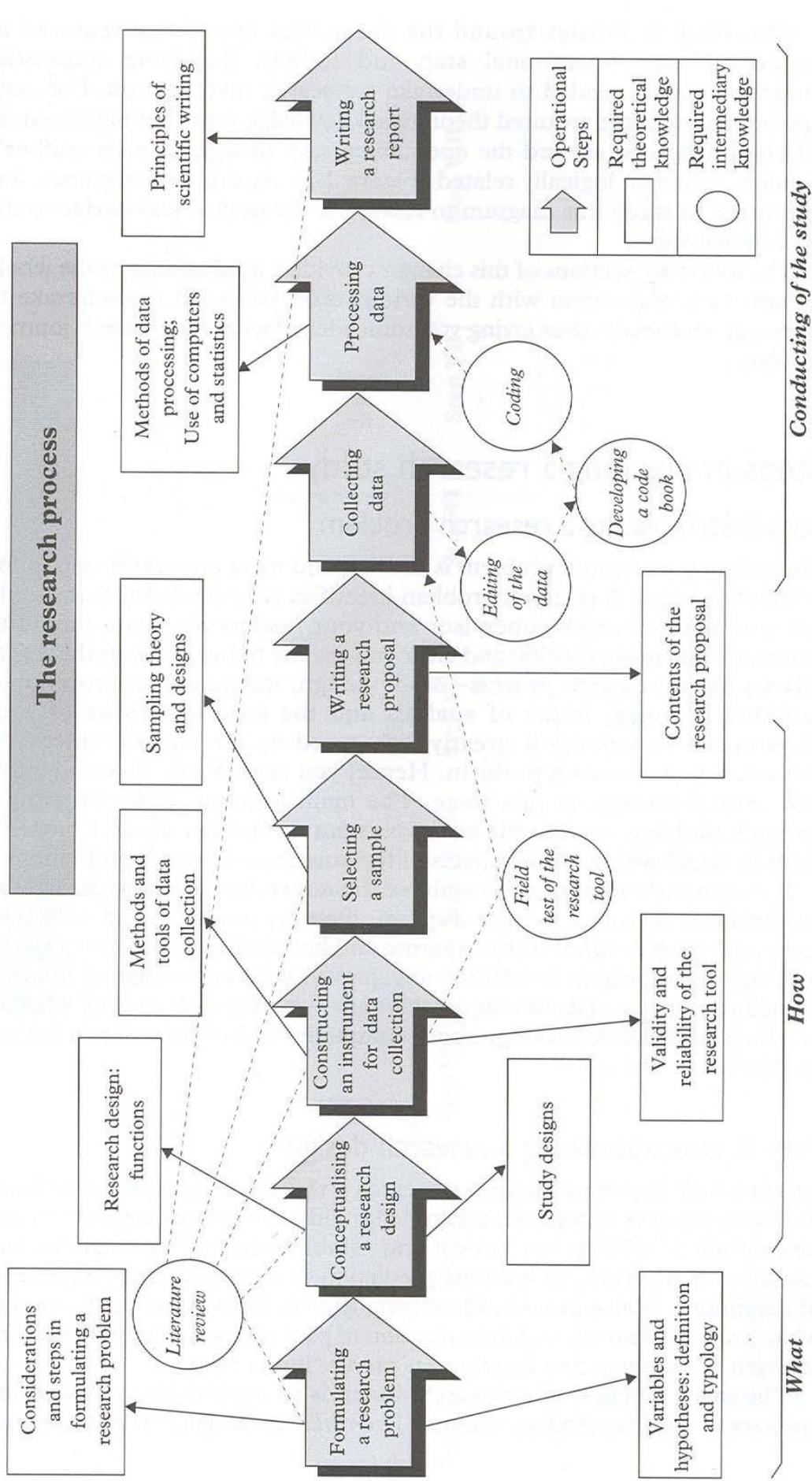
Das Flussdiagramm empirischer Forschung

Ich halte es nicht für meine Aufgabe, der empirischen Forschung eine Logik oder Legitimation zu unterlegen oder ihr Prozedere im Einzelnen zu erläutern oder aus in der Mathematikdidaktik vorliegenden Beispielen zu filtern. Aber da dies in der Regel auch nicht von denen gemacht wird, die sich dieser Methoden bedienen, füge ich hier das Förderband oder Flussdiagramm der Forschung aus [Kumar 2005] auf der nächsten Seite ein und will dabei nur auf zwei Punkte kurz eingehen.

Zum einen fällt die Technizität der Erkenntnisgewinnung auf, zum anderen die unmittelbare Abfolge von dem vorletzten Pfeil *Processing data* zu dem letzten *Writing a research report*. Wo und wann wird denn bei diesem Ablaufschema überhaupt überlegt, gedacht und verworfen, erkannt und durchschaut? Wenn man die Konzeptualisierung – um im Jargon zu bleiben – outsourced hat und die Bewertung der gesammelten Daten nur als Beschreibung der Realität betrachtet, dann gibt es gar keinen Ort des Denkens in der Forschung mehr und auch keinen solchen, an dem die Aussagen im Report Geltung gewinnen könnten. Die Realität ist – zumindest vermeintlich und in messbaren Aspekten – verdoppelt worden, sie liegt nun auch in Ausschnitten in digitaler Form vor, die man nun ihrerseits wieder in Worte fasst und diese als Erkenntnis ausgibt. Eine Auseinandersetzung mit der Sache bleibt aus, man setzt sich allenfalls mit den statistischen Resultaten auseinander.

¹ Die Jahreszahl 2000 ist falsch. Der Artikel ist wohl 2004 erschienen.

Figure 2.2 The research process



Rasterfahndung mit der Wünschelrute

Das Eigenartige an der empirischen Forschung ist, dass sie immer etwas hervorbringt; ob eine noch so kleine Examensarbeit oder einen International Survey, sie kann nicht leer ausgehen. Ihre Förderbänder und Siebe funktionieren zumindest halb- wenn nicht vollautomatisch. Nicht ungern erforscht sie in der Bewusstlosigkeit ihrer Betriebsamkeit auch die Spuren ihres eigenen Grabens, also die Spuren, die ihre Grabwerkzeuge hinterlassen haben. Immer erhält sie Zahlen und damit Skalen, Intervalle, Ranglisten und, was dann der statistische Apparat folglich zur Verfügung stellt, Mittelwerte, Standardabweichungen, Korrelationen und so fort. Man wünscht sich mit der empirischen Rute gleichsam das Ergebnis herbei. Durch das Messen erhält man Daten – eigentlich sogar nur Zahlen, deren Realitätsgehalt (oder genauer Realitätsbeschreibungsvermögen) außer Frage steht, ja dem ganzen Prozess eine Art quasi-naturwissenschaftlicher Dignität verleiht. Die Frage, ob man da tatsächlich etwas misst, was man später ausgraben kann, stellt sich gar nicht, weil der Gegenstand durch den Messprozess hervorgebracht und konstituiert wird. Solche Forschung produziert ihre Begriffe (heute sagt man dann auch *Konzepte*) und Ergebnisse parthenogenetisch, notfalls auch ohne jede Anleihen bei der bemessenen Realität oder Bezügen zu ihr. Und wenn man die Rotzigkeit (Pardon!) des Positivismus zugrunde legt, gilt das dann alles bis zu seiner Widerlegung, die – dessen sind sich die Forscher sicher – faktisch, wenn nicht sogar prinzipiell, ausgeschlossen ist. Die Sache ist versiegelt. Wo Kritik nicht abprallt, freut man sich über Folgeprojekte, die dann ebenso in der beschriebenen Art bearbeitet werden.

In eigener Sache, nämlich als Mathematiker und als Mathematikdidaktiker, möchte ich hier einmal auf die schibboletische Verwendung der Mathematik in der empirischen Bildungsforschung hinweisen. Das Wort *Schibbolet* ist dem Buch *Über die Zeit der Richterinnen und Richter* der Bibel entnommen. Ich zitiere aus der ‚Bibel in gerechter Sprache‘:

Gilead nahm die Furten des Jordans , die nach Efraim führten, ein. Wenn ein Flüchtling Efraims sagte: „Ich will übersetzen“, fragten ihn die Männer Gileads: „Bist du ein Erfraimiter?“ Sagte er: „Nein“, so sagten sie zu ihm: „Sag mal Schibbolet“, das heißt ‚Fruchtweig‘. Sagte er dann „Sibbolet“, weil er es nicht anders aussprechen konnte, ergriffen sie ihn und töteten ihn an den Furten des Jordans. In dieser Zeit fielen aus Efraim 42000 Menschen.

[Bail/Crüsemann 2006]

In übertragenem Sinn wird Schibbolet als Kenn- oder Codewort benutzt. Wer einer bestimmten Sprachgemeinschaft oder Expertengemeinde nicht angehört, der hat – falls er es

nicht kennt – mit schweren Folgen zu rechnen: Sozusagen Aberkennung der Asylrechte mit Todesfolge.

Das Schibboleth der empirischen Bildungsforschung ist die Mathematik in Form gängiger und auch nicht-gängiger statistischer Verfahren, die gern als ‚state of the art‘ apostrophiert werden. Während der Mathematikdidaktiker und Erziehungswissenschaftler Hans Werner Heymann als eines der Ziele eines Allgemeinbildenden Mathematikunterrichts die Rolle der Mathematik als Vernunftverstärker nennt, ihr also eine aufklärerische Funktion im allgemeinbildenden Kontext zuspricht, wird ihr in der empirischen Bildungsforschung eine gegenteilige Rolle zuerkannt: sie dient zur Verschleierung. Nicht nur dass die statistischen Methoden als Werkzeuge in ihrer Eigenart und in Determination und Formation ihrer Ergebnisse nicht reflektiert werden, dass man also – um es lapidar zu sagen – weiß, was man wie und warum und mit welchen Folgen und Wirkungen da macht, wenn man Daten den statistischen Methoden und dem instrumentellen Verstand oder Unverstand unterzieht, nein, das unerklärte und un-erklärende Gegenteil spielt sich ab: Die Generierung der Geltung der Aussagen der empirischen Forschung wird vorsätzlich unkenntlich gemacht, in dem man die eingesetzten mathematischen Verfahren weder erläutert, noch überhaupt ihre Zweckmäßigkeit in dem fraglichen Erkenntniszusammenhang in irgendeiner Weise diskutiert oder rechtfertigt. Ob man sich die – auch für Experten nicht ohne Rest und an einigen Stellen nur widersprüchlich entschlüsselbaren – technischen Handbücher des Großunternehmens PISA hernimmt² oder sich durch einen hohen Wert von Crombachs alpha in einer kleineren Untersuchung beeindrucken lässt, man zieht in der Regel den Kürzeren, und die Sache ist auch darauf angelegt: Man kann Schibboleth nicht aussprechen; dieses Eingeständnis hätte schwere Folgen für die eigene wissenschaftliche Bonität, also liest und denkt man darüber hinweg.

Wenn man noch die vorsätzliche Nicht-Veröffentlichung der so genannten Items, also die Geheimniskrämerei um die Inhalte³ hinzunimmt, kann man sagen, dass in der empirischen Bildungsforschung Items, die keiner kennt, mit Verfahren untersucht und aufbereitet werden, die keiner versteht, um zu Schlüssen zu kommen, die jeder teilt, sofern er auch etwas von dem großen Kuchen oder dem nächsten Projekt abbekommen will. Und die Bildungspolitik ist dankbar, weil hier im Vierteltakt von Wahlperioden Tests und Vergleichsarbeiten immer

² Siehe: Wuttke 2007

³ So wird zum Beispiel in der COACTIV-Studie der Zusammenhang von mathematischem Fachwissen und mathematikdidaktischem Wissen untersucht. Da aber weder die 13 Items zu dem einen Bereich noch die 23 zu dem anderen veröffentlicht wurden, besagen die publizierten Folgerungen wenig. Wenn etwa die mathematikdidaktischen Items im Wesentlichen stoffdidaktischer Natur waren, dann verwundert die konstatierte hohe Korrelation zwischen Fachwissen und fachdidaktischem Wissen kaum.

neue Daten gebären und mediales – meist Schreckens- – Echo finden, was Tatkraft und irgendwie auch Besserung suggeriert.

Der schibbolethische Schutzschirm, der den Einblick in den Forschungsprozess und seine argumentative Rekonstruktion verhindert, wenn nicht gar verhindern soll, wendet sich bei dem gängigen arbeitsteiligen Vorgehen, aber auch gegen die Forscherinnen und Forscher selbst. Die Parzellierung des Sachverstandes etwa in statistische und interpretierende Expertise führt zu unsinnigen oder zumindest fragwürdigen Resultaten. Wer die statistischen Methoden, ihre Charakteristika und ihre Wirkungen nicht durchschaut, der kann die mit ihnen aufbereiteten oder verformten Daten kaum sinnvoll deuten, schon weil er nicht zu unterscheiden weiß, was auf das Verfahren und was auf den untersuchten Gegenstand selbst zurückzuführen ist. So kann etwa eine Faktoranalyse Faktoren liefern, die sich nicht sinnvoll benennen lassen und die es möglicherweise real gar nicht gibt, obschon in der Interpretation dann mit Begriffen wie der „aufgeklärten Varianz“ der nahezu gegenteilige Eindruck erweckt wird, man wisse nun über die Sache mit mathematischer Sicherheit Bescheid. (Zwei einfache Beispiele)

Fragebögen

Die in der empirischen Forschung gängigen Fragebögen, in denen die Antworten in der Regel vorgegeben sind oder man den Grad der Zustimmung oder Ablehnung zu einer Aussage auf einer Skala zum Beispiel von -3 bis 3 ankreuzen soll, geben gar keine Auskunft über die tatsächliche Meinung oder Auffassung der Befragten sondern eben nur darüber, was sie in der vorgegebenen Zeit unter den vorgegebenen Antworten oder auf den entsprechenden Skalen ankreuzen. Aus diesen Kreuzen die Einstellungen der Testandi rekonstruieren zu wollen, ist ein sinnloses und höchst rekurrentes Unterfangen: man erhält die Begrifflichkeiten, die man hineingesteckt hat, nun mit statistischen Werten versehen, wieder heraus und interpretiert sie als Aussagen über die Realität. Auf meine Rückfrage zu der Untersuchung TEDS-M 2008, bei der es sich um eine internationale „Befragung der Lehrenden in der ersten und zweiten Phase der Lehrerbildung“ in 20 Ländern handelt, wie ich meine Auffassung auf dem einundzwanzigseitigen Fragebogen, der nur ein Ankreuzen zulässt, denn zum Ausdruck bringen sollte, erhielt ich von der deutschen Leitung die Antwort, dann sollte ich eben kritisch ankreuzen. Die einzige Möglichkeit, dies zu bewerkstelligen, wäre, chaotisch oder aleatorisch meine Kreuze zu setzen, was nicht nur kindisch wäre, sondern auch in der Mittelung der Ergebnisse ohne jede Wirkung bliebe. So habe ich die Rücklaufquote um ein Epsilon vermindert, was der Auswertung, auf deren Ergebnisse ich nur mit mäßigem Interesse warte, und deren medialer Aufbereitung sicher auch keinen Abbruch tun wird.

Das auktoriale Moment

Forschung wird getragen von Forschern, also von Subjekten, die eine Untersuchung, welcher wissenschaftlichen Qualität auch immer, durchgeführt haben und dann der wissenschaftlichen Öffentlichkeit und interessierten Bildungspolitikern ihre Resultate vorlegen. Das nannte man früher Publizieren, und man konnte den oder die Forscher der Diskussion oder Kritik halber auf ihre Elaborate ansprechen oder auf sie eingehen. Ein Beitrag [Brunner/Kunter 2006] in der Zeitschrift für Erziehungswissenschaften über die nicht uninteressante Titel-Frage „Welche Zusammenhänge bestehen zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung?“ ist von zehn Autorinnen und Autoren gezeichnet, woraus sich im Durchschnitt 2,5 Seiten für jede und jeden ergeben. Als ich einen von ihnen auf einige Unstimmigkeiten in der Untersuchung ansprach, gab er mir zur Antwort, dass der Beitrag keineswegs seine Meinung zum Ausdruck bringe und er auch manche Formulierungen und Schlüsse für fragwürdig halte, sich aber nicht mit seiner Meinung habe durchsetzen können.

Es hat den Anschein, dass solche Papiere zustande kommen wie Kabinettsvorlagen, bei denen die Beteiligten in einer dem Leser nicht zugänglichen und allenfalls erahnbaren Hierarchie zu einem Votum kommen, für das eigentlich keiner so recht die Verantwortung trägt oder übernehmen will. Es ist auch nicht rekonstruierbar⁴, ob die aufgeführten Autorinnen und Autoren überhaupt bei dem Verfassen des von ihnen mit gezeichneten Textes oder möglicherweise an dem Projekt als Codiererinnen bzw. Codierer oder Raterinnen bzw. Rater⁵ oder in anderen Funktionen beteiligt waren⁶. So wird Wissenschaft anonym, während ihre Autorität zugleich schon durch die schiere Zahl der Projektbeteiligten zu steigen scheint. Das hat übrigens auch ad personam ganz merkwürdige Folgen. So sollte ich letztthin über die Rangfolge eines Bewerbers auf einer Berufungsliste entscheiden, dessen Publikationen mit einer Ausnahme solch kollektiven Charakter trugen. Wie entscheidet man über jemanden, von dem man nur weiß, wo er beteiligt war, aber nicht, was er geschrieben hat?

⁴ Unklar ist auch, welche Projektverantwortlichen wie darüber entscheiden, wer als Autorin oder Autor aufgeführt wird. Potentiell entstehen so virtuelle Publikationskarrieren.

⁵ Die Übernahme dieses Begriffs ins Deutsche ist nicht ohne Ironie.

⁶ In dem eingangs aufgeführten Buch [Baumert/Lehmann 1997] nennt die Titelseite und die weiteren Deckblätter ‚Jürgen Baumert, Rainer Lehmann u.a.‘ als Autoren, aber welche der 23 auf der Seite 39 aufgeführten ‚Wissenschaftlichen Mitarbeiter in Deutschland‘ welche Kapitel verfasst und bearbeitet haben und welche von ihnen überhaupt bei der Erstellung des Buches in welcher Funktion mitgewirkt haben, wird nicht angegeben.

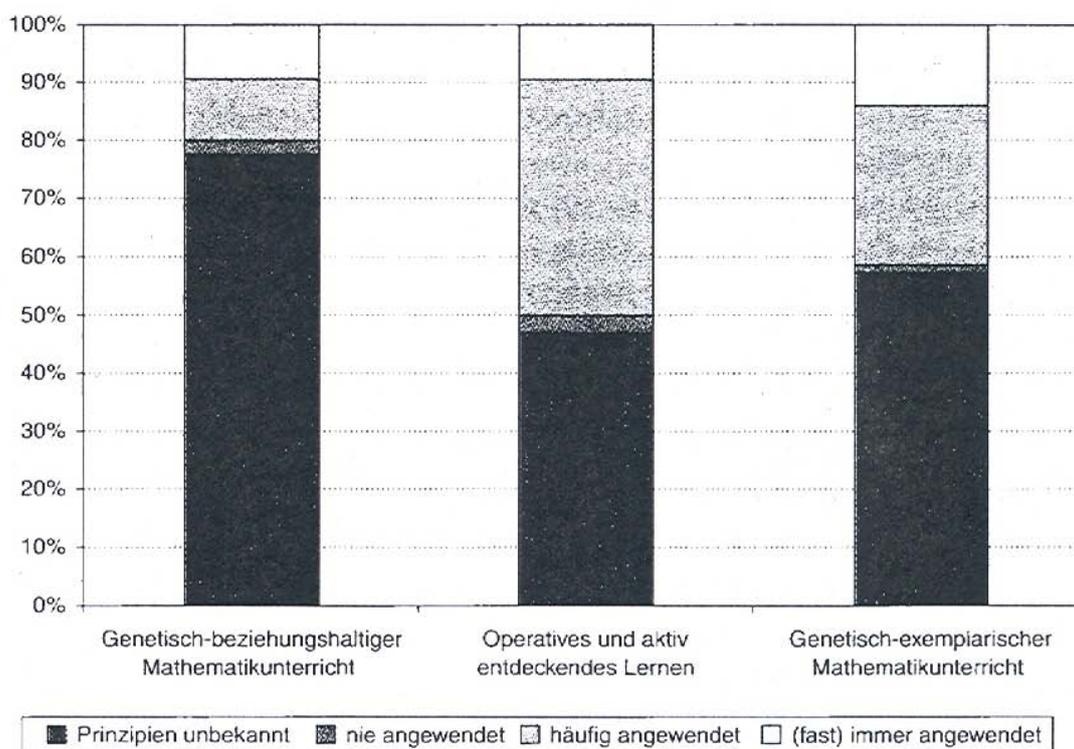
Kollektive Autorenschaft und Gruppenintelligenz - ein Beispiel

In der Arbeit [Kunter/Dubberke 2006, S. 172/3], die von 11 Autorinnen und Autoren gezeichnet ist, wird über die Resultate einer Untersuchung berichtet, in wie weit Lehrkräfte mit *mathematikdidaktischen Instruktionsmodellen* vertraut sind. Es heißt dazu:

Die folgenden Ergebnisse basieren auf Aussagen der Lehrkräfte der untersuchten Klassen, die befragt wurden, ob bestimmte grundlegende fachdidaktische Prinzipien bekannt sind und wie häufig diese eingesetzt werden. Abbildung 7.1 stellt die Ergebnisse dieser Befragung dar.

Die Abhandlung gibt die eingesetzten Fragen nicht wieder, aber die Befragten sollten wohl jeweils ankreuzen, ob die Prinzipien des *genetisch-beziehungshaltigen Mathematikunterrichts* nach Freudenthal, des *operativen und aktiv entdeckenden Lernens* nach Wittmann⁷ und des *genetisch-exemplarischen Mathematikunterrichts* nach Wagenschein ihnen unbekannt sind oder sie diese *nie*, *häufig* oder *(fast) immer* anwenden.

Abbildung 7.1: Ergebnisse aus der Befragung der Lehrkräfte zur Bekanntheit und zum Einsatz mathematikspezifischer Instruktionsmodelle



⁷ Es fällt übrigens auf, dass die Autoren für die Charakterisierung des *mathematikdidaktischen Instruktionsmodell* nach Wittmann auf einen Beitrag dieses Autors in dem Buch *Mit Kindern rechnen* [Müller/Wittmann 1995] verweisen, das vom Grundschulverband veröffentlicht wurde und den befragten Lehrkräften (der Jahrgangsstufe 10!) in der Regel wohl nicht bekannt sein wird.

Nicht etwa dass die befragten Lehrkräfte dafür gelobt würden, dass sie trotz unzureichender und zum Teil sicher weit zurückliegender fachdidaktischer Studienanteile sich alle Mühe geben, die Intentionen der Fragestellungen nahezu übermäßig zu bedienen – es fällt ja auf, dass nahezu jeder, der angibt eines der Prinzipien zu kennen, behauptet, es auch *häufig* oder *(fast) immer* anzuwenden –, sie bekommen vielmehr die wissenschaftliche Klatsche zeitgenössischer Strömungen normativer Mathematikdidaktik gezeigt:

Vor diesem Hintergrund ist es erstaunlich, dass nahezu 80 Prozent der untersuchten Lehrkräfte den Ansatz als nicht bekannt einstufen ...

...wird demnach von über der Hälfte der Lehrkräfte überhaupt nicht angewendet ...

...wurde von ca 60 Prozent der befragten Lehrkräfte mit ‚Prinzipien nicht bekannt‘ beantwortet ...

Zusammenfassend zeigt sich somit, dass theoretisch viel diskutierte grundlegende mathematikdidaktische Unterrichtsprinzipien – folgt man den Aussagen der Lehrkräfte – bisher kaum Einzug in die Unterrichtsrealität gefunden haben.

Betrachtete man dagegen statt der schwarz unterlegten Balkenabschnitte die grau und weiß unterlegten und unterstellte einmal, dass die Befragten in ihrem Unterricht nicht mehrere dieser Prinzipien gleichzeitig anwenden, so käme man zu dem überraschenden Ergebnis, dass 20 % (Freudenthal) + 50 % (Wittmann) + 42 % (Wagenschein), also insgesamt 112 % der untersuchten Lehrkräfte ihren Unterricht *häufig* oder *(fast) immer* nach diesen *grundlegenden mathematikdidaktischen Unterrichtsprinzipien* gestalten. Ein wunderlicher Erfolg – sei er der fachdidaktischen Ausbildung oder der beflissenen Antwortwilligkeit der Befragten zu verdanken.

Der neue Typus des Forschers

Die empirische Forschung bringt den Typus eines neuen Wissenschaftlers hervor, der dem Geist der Zeit entsprechend von vielen Hochschulen heute gesucht wird. Stichworte und vielerorts berufsentscheidende Parameter wie International Visibility, Zahl der bewilligten DFG-Anträge und Drittmittelvolumen kommen hier zur Anwendung und haben diesen Typus gleichsam erzeugt. Er weiß sich geschickt in Forschergruppen und in Antragsgemeinschaften für Großprojekte zu positionieren, er ist, wie das heute heißt, ‚gut aufgestellt‘, publiziert in Englisch häufig mit drei oder mehr Coautoren, wodurch die genuin eigenen Beiträge kaum zu verorten sind; das auktoriale Moment des forschenden Subjekts verschwindet individuell, gedanklich und sprachlich hinter einem Jargon unpersönlicher ‚Objektivität‘, die sich quasi-

naturwissenschaftlich denkt und dünkt. In der Mathematikdidaktik paart sich dabei zunehmende statistische Expertise, die den Methodenkasten dieses Gebiets und seine Software freilich ebenso unhinterfragt, funktional und dinglich nutzend wie Nicht-Mathematiker aller Richtungen (etwa Testpsychologen), mit einer eigentümlichen Unbeholfenheit und Indifferenz gegenüber dem Stoff, der zu Items in ‚normierten‘ und ‚standardisierten‘ Tests gerinnt, die zuweilen nicht einmal ein durchdachtes Schulbuchniveau erreichen, sowie gegenüber der Unterrichtspraxis und den ‚Abnehmern‘ oder ‚Anwendern‘ der Forschungsergebnisse, zu denen vorwiegend die Lehrerinnen und Lehrer gehören sollten. Man misst immer genauer Lehr- und Lernleistungen, von denen man eigentlich inhaltlich nichts weiß, ja wie es scheint, nicht einmal etwas wissen will. In der Lehrerbildung an den Universitäten lehrt man dann diese messende Analytik, der die unterrichtlichen Bildungsinhalte abhanden gekommen sind, und deren Ergebnisse. Soweit dieser Blick auf ‚Glanz und Elend der neuen Kurtisanen‘ und Höflinge.

Die Wirkung

Forschungsergebnisse sollen wirken und meinethalben auch die Welt oder die Schulwelt verbessern. Aber wie sieht das aus, wenn die Forschung auf Wirkung hin angelegt ist? So führt z.B. Manfred Prenzel in dem Buch PISA '06 aus:

Die Auftraggeber, das sind die OECD-Staaten und weitere Staaten (sogenannte Partnerstaaten), erwarten von dem internationalen Vergleich empirisch fundiertes Steuerungswissen. Die teilnehmenden Staaten möchten von Pisa zum Beispiel erfahren, wo sie im internationalen Vergleich stehen, welche Bildungsergebnisse andere Staaten erreichen und unter welchen Bedingungen das geschieht. (...) Besonderes Interesse findet dabei auch die Frage, inwieweit es in den Staaten gelingt, jungen Menschen unabhängig von ihrer sozialen Herkunft vergleichbare Chancen für die Entwicklung ihrer Kompetenzen zu geben.

Es ist also nicht mehr so, dass diese Forschung durch ihre Dignität und die Anerkennung ihrer Bedeutung in der Fachwelt wirkt, sondern sie wirkt umgekehrt auf die Fachwelt durch die Bedeutung ihrer Auftraggeber und ihr fiskalisches Volumen. Wer bedeutend sein will, muss hier mitmachen und sich nicht in schnöder Kritik ergehen. Dabei-sein ist alles und nicht etwa der wissenschaftliche Impetus, den Dingen auf den Grund zu gehen.

Steuerungswissen

Der Begriff Steuerungswissen entstammt nicht dem Wörterbuch der Wissenschaftlerin sondern dem des Funktionärs. Sein epistemologischer Status ist mehr als zweifelhaft. In einer münchhausenhaften Weise behauptet und setzt er sich selbst und suggeriert den Beteiligten und den Auftraggebern diverser Studien, dass es eine Art Stellschraubenwissen gäbe, um die untersuchten und in der Regel als mangelhaft diagnostizierten ‚Werte‘ zu erhöhen, ohne tatsächlich ihre Ursachen zu ergründen. Er setzt schon begrifflich die Erkenntnisse als gesichert, die es eigentlich zu gewinnen gälte, und unterstellt zugleich, dass er die Verhältnisse in seinem Sinne zu beherrschen vermöge.

Während es bei dem bereits oben zitierten Soziologen und Philosophen Theodor W. Adorno heißt: „Denn wahr ist nur, was nicht in diese Welt passt“⁸, lernen wir von dem Begriff Steuerungswissen, dass nur wahr ist, was so zu wirken vorgibt, wie die Auftraggeber einer Studie es sich vorstellen – möge es sich dabei um die der wirtschaftlichen Prosperität der Industriestaaten verpflichteten OECD oder die Kultusministerkonferenz oder Landesbildungsministerien oder andere Geldgeber handeln, die das ‚Drittmittelvolumen‘ und damit das Ansehen der beteiligten Forscherinnen und Forscher steigern. Solche Forschung ist ‚programm‘⁹-gemäß¹⁰ vorsätzlich affirmativ und ihr fehlt, wie ich denke, jeglicher Stachel der Erkenntnis.

Wo dieser Stachel offensichtlich fehlt, also die empirischen Forschungsergebnisse schlicht selbstverständlich sind und sich ebenso aus einfachen Betrachtungen oder Überlegungen ergeben, wird der organisatorische und finanzielle Aufwand der empirischen Forschung gern damit gerechtfertigt, dass man bislang den fraglichen Sachverhalt zwar hätte vermuten können, aber nun erst sei er „wissenschaftlich“ belegt. Dieses Argument ist zwar ein verständlicher Versuch, die empirische Methode als einzige ihrer Erkenntnisse sichere zu verabsolutieren, aber es ist doch so dünn, dass sogar die Auftraggeber solcher Forschung dies auf die Dauer bemerken werden. Milch ist nicht erst dann weiß, wenn man 50.000 Schülerinnen und Schüler befragt hat, welche Farbe sie wohl habe.

⁸ [Adorno 1970, S. 93]

⁹ PISA ist bekanntlich das Akronym für Programme for International Student Assessment.

¹⁰ Auch das ‚Rahmenprogramm zur Empirischen Bildungsforschung‘ des Bundesministeriums für Bildung und Forschung (BMBF) stellt eine Art Zeitgeist-Pantheonisation dieser Richtung dar, die die Antragsteller eher skeptisch stimmen sollte.

Literatur

Adorno, Th. W.: Soziologie und empirische Forschung. In: Th. W. Adorno u.a.: Der Positivismusstreit in der deutschen Soziologie. Darmstadt 1972, S. 86f.

Adorno, Th. W.: Ästhetische Theorie. In: Adorno, G.; Tiedemann, R. (Hrsg.): Theodor W. Adorno. Gesammelte Schriften 7. Suhrkamp Verlag. Frankfurt 1970

Bail/ Crüsemann/Crüsemann et al. (Hrsg.): Bibel in gerechter Sprache. Gütersloher Verlagshaus. Gütersloh 2006, S. 441

Baumert/Lehmann et al: TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich – Deskriptive Befunde. Opladen 1997, S. 17

Brunner/Kunter/Krauss/Baumert/Blum/Dubberke/Jordan/Klusmann/Tsai/Neubrand: Welche Zusammenhänge bestehen zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung? In: Zeitschrift für Erziehungswissenschaften, 9. Jahrg., Heft 4, 2006, S. 521-544

Krauss/Kunter/Brunner/Baumert/Blum/Neubrand/Jordan/Löwen: COACTIV: Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz. In: Doll/Prenzel (Hrsg.): Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung. Waxmann. Münster 2004, S. 31-53

Kumar, Ranjit: Research Methodology. SAGE Publications Ltd. London 2005, S. 19

Kunter/Dubberke/Baumert/Blum/Brunner/Jordan/Klusmann/Krauss/Löwen/Neubrand/Tsai: Mathematikunterricht in den PISA-Klassen 2004: Rahmenbedingungen, Formen und Lehr-Lernprozesse. In: PISA-Konsortium Deutschland (Hrsg.): Pisa 2003: Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahrs. Waxmann. Münster 2006, S. 161 – 194

Müller, G. N.; Wittmann, Ch. E. (Hrsg.): Mit Kindern rechnen. Frankfurt am Main: Arbeitskreis Grundschule

PISA-Konsortium Deutschland (Hrsg.): Buch PISA '06. Die Ergebnisse der dritten internationalen Vergleichsstudie. Waxmann Verlag. Münster 2007, S. 13

Wutke, Joachim: Die Insignifikanz signifikanter Unterschiede: Der Genauigkeitsanspruch von PISA ist illusorisch. In: Jahnke/Meyerhöfer (Hrsg.): Pisa & Co. Kritik eines Programms. 2. erweiterte Auflage. Franzbecker Verlag. Hildesheim 2007, S. 99 – 243

Letzte Durchsicht: Donnerstag, 12. November 2009