Allerorten wird vermessen, bewertet, zertifiziert. PISA ist überall. Doch was geschieht hier überhaupt? Leiden immer mehr Schüler und Lehrer am Bildungssystem weil dasselbe bisher einfach zu wenig vermessen worden ist? Ganz sicher nicht, meint der emeritierte Professor für Erziehungswissenschaft **Hans Brügelmann** im Gespräch mit **Jens Wernicke** und skizziert die allerorten zu beobachtende Zunahme des von anderer Seite bereits als „Menschenmesserideologie“ kritisierten Testwahns als Indikator für den Einzug neoliberaler Paradigmen ins Bildungssystem, die großen Schaden anrichteten.

Dieser Beitrag ist auch als Audio-Podcast verfügbar.

http://www.nachdenkseiten.de/upload/podcast/160114_Die_Standardisierung_NDS.mp3

Podcast: Play in new window | Download

Herr Brügelmann, in Ihrem unlängst erschienen Buch „Vermessene Schulen – standardisierte Schüler“ wenden Sie sich gegen die allgemeine Testeritis, die das Bildungssystem seit einiger Zeit heimsucht. Wieso? Was haben Sie gegen derlei Bemühungen um Qualitätssicherung?

Durch Leistungstests die Qualität von Schule sichern? Eine kühne These! Solche Tests erfassen doch nur den sogenannten „Output“ und davon wiederum nur einen kleinen Teil, schon vom Ansatz her eingeschränkt auf wenige Fächer und von diesen wiederum nur Ausschnitte – und das auch noch in sehr oberflächlicher Form.

Wo bleiben die Wirkungen der Schule auf das soziale und das ästhetische Lernen oder gar auf die Persönlichkeitsentwicklung der Kinder und Jugendlichen? Und was ist mit der Qualität des pädagogischen Umgangs: Respekt füreinander, Offenheit für die besonderen Bedürfnisse und Interessen einzelner Kinder, Bemühen um die Überwindung ihrer Schwächen, demokratische Entscheidungsverfahren usw.? Ich finde, die Fixierung auf Punktwerte in Leistungstests führt zu einer Verarmung der Qualitätsdiskussion.

Die allerorts boomenden Tests messen also nicht wirklich Qualität? Warum schießen sie dann aber wie Pilze aus dem Boden und behauptet man das allerorts?

Es ist die Faszination durch scheinbar objektive Zahlen… Das ist mit den Quartalszahlen in der freien Wirtschaft nicht anders als mit den Liegetagen im Krankenhaus. Man hofft, das ja

immer fehleranfällige Urteil beteiligter Menschen ausschalten zu können. Denken Sie nur an die Diskussion über Noten.

Aber das Problem kriegt man mit Tests nicht weg. Der Preis für „Objektivität“ ist Standardisierung. Menschliche Verhaltensweisen, also auch Leistungen in Tests, sind allerdings stets mehrdeutig. Standardisierung bedeutet daher den Verzicht darauf, unter die Oberfläche zu gucken. Denn eben das würde auch eine Interpretation des Sichtbaren erforderlich machen. Man käme um ein persönliches Urteil gar nicht herum.

Der aktuelle Testboom stellt also nur ein „oberflächliches Messen“ dar, das der Vielschichtigkeit und Mehrdeutigkeit menschlicher Lebensäußerungen nicht gerecht zu werden vermag, und daher eine Art „Standardisierung“ forciert? Wie meinen Sie das denn? Im Sinne von: der Tests erzeugt eine Lebensrealität, in der dieses gewollt und jenes verpönt ist und daher passen sich dann die Getesteten der vermeintlich „allgemeinen Wahrheit“ des Testes mehr und mehr an, verlieren selbst ihr Vielschichtig- und Mehrdeutigkeit?

Das Grundproblem: Es zählt nur die „richtige“ Antwort, das heißt, die von den Testautoren vorgegebene Lösung. Die persönliche Lesart eines Textes, die individuelle Deutung einer Aufgabe, untypische Lösungswege bringen keine Punkte.

Damit werden Tests zum „heimlichen Lehrplan“ – angesichts der genannten Schwächen und aktuellen gesellschaftlichen Entwicklungen eine echte Gefahr.

Inwiefern?

Nun, zum einen lernt man am besten, wenn es Bezüge zu eigenen Erfahrungen gibt, wenn das Neue auch an persönlichen Vorstellungen anknüpfen kann, wenn sich eine Brücke zu individuellen Interessen schlagen lässt. Genau das aber wird um der Standardisierung willen ausgeblendet.

Und zum anderen: Es entsteht gerade eine gewaltige Testindustrie weltweit, und wenn es bei der Gläubigkeit der Politik an die Messbarkeit alles Menschlichen bleibt, bestimmen mehr und mehr diese Firmen, was Bildung eigentlich ist. Bildung würde verflachen und sich mehr und mehr nur noch um Aspekte wie „schnell“ und „verwertbar“ drehen. Das ist eine Gefahr nicht zuletzt auch für die Demokratie.

Das wäre dann ja im Wortsinne eine „Kolonialisierung der Lebenswelt“, wie Habermas das einmal nannte: Eine Ökonomisierung aller Lebensbereiche, die sich

dank geeigneter Verfahren so inszeniert, dass sie als Ideologie und größtenteils Bedrohung von den Menschen wahrgenommen wird, sondern als „wissenschaftliche Messung“ vielmehr die Menschen zu jenem Homo Oeconomicus erzieht, den sie braucht…

Zumindest besteht die Gefahr. Dabei ist vor allem die Frage, welchen Status Lehrerinnen und Lehrer, Schülerinnen und Schüler, Eltern und Arbeitsgeber den Tests und ihren Ergebnissen beimessen.

An sich können Tests ja durchaus hilfreich sein, als Warnlampe: Hier muss man genauer hinschauen, da noch was tun. Ähnlich wie in der Medizin eine erhöhte Temperatur Anlass zu einer differenzierteren Diagnose ist – aber eben nicht die Diagnose selbst. Schließlich käme niemand auf die Idee, das professionelle Urteil des Arztes durch ein Thermometer zu ersetzen.

Genau das aber droht in der Pädagogik. Und wenn dann noch auf den Test hin geübt wird, verliert der letztlich sogar noch seine Aussagekraft als Warnlampe. Wie viele Buchstaben Schulanfänger benennen können, war einmal ein brauchbarer Indikator für die vorschulische Schriftspracherfahrung der Kinder. Daraufhin haben Eltern oder Oma und Opa mit den Kleinen das Alphabet geübt, so dass sie es aufsagen können. Jetzt aber weiß man nicht mehr: Ist das nur inhaltsleeres Oberflächenwissen oder steckt mehr dahinter?

> „Wenn Wissenschaft das ist, was die Meinungsführer unter jenen Leuten sagen, die von der Regierung das meiste Geld für etwas bekommen, was ‚Wissenschaft‘ genannt wird, dann ist PISA natürlich Wissenschaft. Im Mittelalter gehörten in diesem Sinne die elaboriertesten Theorien zur Hexenverbrennung ja auch der wissenschaftlichen Speerspitze an. Wenn Sie hingegen an ein wissenschaftliches Messinstrument den Anspruch haben, dass benennbar ist, was überhaupt gemessen wird, dass das Instrumentarium der wissenschaftlichen Debatte zugänglich ist, dass das verwendete mathematische Modell zum Gegenstandsbereich passt und dass die Stichprobe bestimmten Kriterien genügt, dann würde man das PISA-Instrumentarium sicherlich eher in den Grenzbereichen des Wissenschaftlichen verorten.
> Im Vorwort des wissenschaftlichen Sammelbandes ‚PISA zufolge PISA‘ von Stefan Thomas Hopmann und anderen heißt es, die dort publizierten 18 Beiträge von Forscherinnen und Forschern aus 7 europäischen Ländern

> zusammenfassend, denn auch in aller Deutlichkeit, ich zitiere:
> ‚Das PISA-Projekt ist offenkundig mit so vielen Schwachstellen und Fehlerquellen belastet, dass sich zumindest die populärsten Endprodukte, die internationalen Vergleichstabellen sowie die meisten nationalen Zusatzanalysen zu Schulen und Schulstrukturen, Unterricht, Schulleistungen und Problemen wie Migration, sozialer Hintergrund, Geschlecht usw., in den bisher praktizierten Formen wissenschaftlich schlicht nicht aufrecht erhalten lassen.‘
> Genau darum geht’s.“

Wolfram Meyerhöfer: „Macht PISA dumm?“

Könnten Sie das mit den Tests denn bitte einmal konkreter ausführen? Den Prozess des Messens und Fehlmessens, des Standardisierens und Menschenübersehens dabei?

Eine Schwäche von Tests ist stets die Künstlichkeit der Aufgabe. Zum Beispiel müssen mathematische oder naturwissenschaftliche Probleme als Text dargeboten werden. Insofern wird indirekt auch die Lesefähigkeit mit geprüft. Besonders benachteiligt sind damit Kinder mit einer anderen Muttersprache. Bei der Beantwortung von Fragen zu Texten wiederum spielen auch Wortschatz, Weltwissen und Kombinationsfähigkeit eine Rolle.

Auch die Vorgabe von Auswahlantworten oder die Begrenzung der Zeit engen den Lösungsraum künstlich ein. Man kann seine Antworten nicht mehr erläutern, das heißt, die Deutungshoheit für die richtige Lesart der Aufgabe und für die Bewertung der Lösung geht verloren. Sie wird sozusagen an den Test beziehungsweise dessen Macher überstellt.

Zielen Sie mit Ihrer Kritik auch auf die sogenannte „Hattie-Studie“, eine gigantische Metaanalyse, mit der man uns bspw. in Hessen erklärt, endlich gäbe es den unwiderlegbaren wissenschaftlichen Beweis, dass Lernerfolge unabhängig von der Klassengröße zustande kämen, weswegen wir Linken und Gewerkschafter doch endlich die Klappe halten und aufhören sollten, nach besseren Lern- und Lehrbedingungen zu rufen?

Ja, Metaanalysen, in denen die Kennwerte vieler Studien zu einem einzigen Thema miteinander verrechnet werden, stellen in diesem Kontext ein ganz besonderes Problem dar: Bei der Verdichtung von Ergebnissen auf immer höheren Abstraktionsebenen verlieren

die bewerteten Phänomene wie „Hausaufgaben“ oder „offener Unterricht“ immer mehr an inhaltlicher Bedeutung, weil in den zusammengefassten Teilstudien jeweils ganz unterschiedliche Realisierungen untersucht wurden.

Dabei zeigt doch schon die Alltagserfahrung, dass es etwa bei Hausaufgaben darauf ankommt, worauf die sich inhaltlich beziehen, wie man die Aufgaben vorbereitet bzw. stellt und nach Erledigung bespricht oder auswertet, ob und wie die Eltern eingreifen. Weiter macht es für den Sinn von Hausaufgaben einen Unterschied, um welches Fach es sich handelt und welche Kompetenzen man mit den Hausaufgaben fördern will.

In Pauschalurteilen, wie sie bei Hattie herauskommen, geht alles das verloren.
Was vielleicht noch für Allgemeinentscheidungen hilfreich ist, zum Beispiel für die Politik, also Häufigkeiten, Durchschnitte, Korrelationen und ähnliches, wird dem Einzelfall nicht gerecht. Kontextbedingte Abweichungen und das individuell Bedeutsame werden forschungsmethodisch zum „Schmutzeffekt“, zum „Ausreißer“ erklärt.

So etwas geschieht gerade ja allerorts. Im Sozial-, im Gesundheitssystem – offenbar in allen Bereichen der öffentlichen Daseinsvorsorge. Skizzieren wir hier die Ideologie des Sozialabbaus; die Rechtfertigungslehre des Neoliberalismus, mit der er seine Schandtaten zu legitimieren sucht?

Ja, die generelle Tendenz zu quantitativen Kennwerten als Ausweis für die Qualität öffentlicher Versorgung – also die „Leistung“ in Krankenhaus, Verkehr, Theater, Museum, Bildung etc. – muss sehr kritisch betrachtet werden.

Es ist ja richtig, Rechenschaft für die Verwendung von Steuergeldern zu fordern. Aber vom ständigen Wiegen ohne angemessene Fütterung wird die Sau auch nicht fett. Ganz im Gegenteil. Und, um im Bild zu bleiben: Wenn das Ziel gar nicht sein sollte, dass die Sau fetter, das Bildungssystem und anderes also nur im Testsinne „besser“ wird, sondern dass sie sich gesund entwickelt, wenn die Waage also das falsche Instrument ist, überhaupt „Qualität“ zu messen, dann hat das verheerende Folgen.

Dann werden beispielsweise Patienten aus dem Krankenhaus „blutig“ entlassen, um die Statistik zu verbessern, und dann wird auch im Bildungsbereich getäuscht, um bei Vergleichstests wie etwa VerA gut abzuschneiden.

Ihre Kritik also in einem zusammenfassenden Satz?

Standardisierte Messverfahren, die vorgeben, Qualität erst zu erheben und dann zu

erhöhen, wirken wie Nebelkerzen, die die realen Probleme verdecken.

Bereits 1976 hat der Sozialpsychologe Campbell als „Gesetz“ formuliert: „Je häufiger ein quantitativer sozialer Indikator für gesellschaftliche Entscheidungen genutzt wird, desto stärker verführt er zum Täuschen und desto nachhaltiger wird er die sozialen Prozesse stören, die er messen soll.“

Wolfram Meyerhöfer spricht in diesem Zusammenhang von einer „Menschenmesserideolgie“. Damit beschreibt er den Glauben daran, alles Menschliche messen zu können: „PISA behauptet, mit ein paar dümmlichen Ankreuzaufgaben die Qualität von Schulsystemen messen zu können. Die Stiftung Warentest behauptet, den Geschmack von Schokocreme auf einer Skala angeben zu können. Es gibt sogar Skalen für religiöse Kompetenz. Natürlich kann man jedes Element des Seins in eine Skala pressen. Menschenmesserideologie ist aber der Glaube, dass diese Skala wirklich etwas über den Menschen erzählt“. Trifft es das? Sehen Sie es ähnlich?

Ja, der gesellschaftspolitische Kontext verschärft das Problem – ganz unabhängig von den test-immanenten forschungsmethodischen Problemen, auf die ich eingangs hingewiesen habe.

Wie Karl-Heinz Dammer in fundierten Analysen herausgearbeitet hat, werden diese Tests dabei auch als Schmiermittel der neoliberalen Ideologie genutzt: Wo sie unkritisch „geglaubt“ werden, wird nicht nur die Deutungshoheit über den Bildungsbegriff, über Gesundheit usw. privatisiert – es wird auch ein Aktionismus unterstützt, der nicht wirklich hilft, und somit der dringend notwendigen Debatte darüber, welche realen Nöte im Bildungssystem bestehen und was gegen diese getan werden kann, der Saft abgedreht.

Was würden Sie denn vorschlagen, wäre hiergegen zu tun? Welche konkreten Schritte wären wider diese Testerits und ihre Heilsversprechungen aktuell sinnvoll zu tun?

Wir können das persönliche Urteil nicht ausschalten. Das ist auch nicht schlimm, solange es nicht mit zu viel Macht verbunden ist. Nötig ist eine soziale Kontrolle von Bewertungen statt technisch immer perfekterer Methoden. Wie im Gerichtswesen mit unterschiedlich definierten Rollen. Wahrnehmungen müssen ausgetauscht, Deutungen ausgehandelt werden: zwischen Lehrerinnen, Schülerinnen und ihren Eltern. Statt Standardisierung brauchen wir Mehrperspektivität. Auch bei der Beurteilung von Schulen und auf der Systemebene.

Ein Beispiel: Ich war mehrere Jahre Sprecher des seit 1989 bestehenden Schulverbunds „Blick über den Zaun“. Diese 140 Schulen haben sich in 16 Arbeitskreisen zusammengeschlossen. Gemeinsam haben sie Standards für ihre Arbeit beschlossen, die sich vor allem auf die Qualität der pädagogischen Prozesse beziehen. Für deren Evaluation wird versucht, den hilfreichen Außenblick und die notwendige Vertrautheit mit der Alltagspraxis auszutarieren, indem regelmäßige Peer-Reviews verabredet haben. Zweimal im Jahr wird eine Schule von jeweils zwei Vertretern der anderen Schulen des gleichen Arbeitskreises besucht, die zwei bis drei Tage in der Schule verbringen. Die gastgebende Schule kann einen spezifischen Beobachtungsauftrag formulieren, doch unabhängig davon sind die Gäste frei, das zu beobachten, was ihnen wichtig erscheint. Sie nehmen am Unterricht teil, unterhalten sich mit Kollegen, Schülern und Elternvertretern. In einer Schlussrunde spiegeln die Besucher dem Kollegium jeweils ihre individuellen und damit unterschiedlichen Eindrücke zurück.

Die Erträge dieses kollegialen Austauschs sind vielfältig – sowohl für die besuchte Schule als auch für die Besucher, die ihre Beobachtungen als Anregung in die eigene Schule zurückbringen. Dialog der Praxis statt hierarchischer Dominanz der Wissenschaft oder Verwaltung – das sehe ich als Alternative zu den standardisierten Test.

Also Unterstützung und Kooperation statt stetem Messen, Testen und Kontrollieren – und das alles unter dem Damoklesschwert, dass, wenn etwas am „Ergebnis“ nicht stimmt, es Sanktionen gäbe etc.? Dieses Bild wurde mir zuletzt von deutschen Schulen gezeichnet – und es wundert wohl nicht, dass hier wenig besser zu werden vermag und noch weniger an Gutem und Neuem „erblüht“… Manchmal habe ich mir inzwischen sogar gedacht: Diese ganzen Tests, die werden doch überhaupt nur dazu gemacht, um später die Lehrer als unfähig oder Schüler als dumm klassifizieren zu können – und also einen Schuldigen zu haben, der eine gute Ablenkung von den systemischen Missständen, der grundlegenden Unterfinanzierung und anderem ist…

In der Tat besteht die Gefahr, dass gesellschaftliche Probleme „pädagogisiert“ oder gar individualisiert werden und die Bedingungen, unter denen Lehrende sowie Schülerinnen und Schüler ihre „Leistungen“ erbringen, aus dem Blick geraten: Armut in den Familien, Selektionsdruck im Bildungssystem, Finanzierung der Schulen usw. Allein kann man nichts dagegen ausrichten. Da hilft nur, sich in Gewerkschaften und Fachverbänden zu organisieren, um politisch agieren zu können.

Hans Brügelmann *war bis 2012 Professor für Erziehungswissenschaft an der Universität Siegen. Inzwischen arbeitet er als Fachreferent für Qualitätsentwicklung beim Grundschulverband, wo er an qualitativen Evaluationsverfahren einer „pädagogischen Leistungskultur“ mitgearbeitet hat.*

Weiterlesen:

- Telepolis-Interview: „Schluss mit PISA?“
- NachDenkSeiten-Interview: „Macht PISA dumm?“
- NachDenkSeiten-Interview: „PISA beschädigt die Bildung weltweit“
- Wortmeldung beim Neuen Deutschland: „Menschenmesserideologie“
- Interview: „Rechenschwäche gibt es nicht“
- NachDenkSeiten-Interview: „Von wegen Rechtschreipkaterstrofe!“
- NachDenkSeiten-Interview: „Die Kinderkrankmacher“
- BdWi-Studienheft Nr. 10: „Naturalisierung und Individualisierung. Beiträge der Wissenschaft zur Legitimation von Armut und Ausgrenzung“

Weitere Veröffentlichungen von ***Jens Wernicke*** *finden Sie auf seiner Homepage jenswernicke.de. Dort können Sie auch* ***eine automatische E-Mail-Benachrichtigung*** *über neue Texte bestellen.*