

Warum quantitativ empirische Forschung nicht neutral ist

MESSBARKEIT ALS NORM

Lutz Koch

Quantitative Forschung hat als Steuerungsgrundlage Hochkonjunktur. Aber die Zahlen sind nur scheinbar neutral und unbestechlich. Der Autor zeigt theoretisch wie auch am Beispiel der PISA-Studien, wo Empirie Normativität mitführt, mittelbar und unmittelbar.

1. Zählen und Messen spielt in den sogenannten exakten Wissenschaften eine wesentliche Rolle, vor allem natürlich in der Physik und den übrigen Naturwissenschaften. Es ist dort aber nicht üblich, darüber nachzudenken, wie es möglich ist, dass die Natur in mathematischen Lettern geschrieben ist. Das liegt außerhalb ihres Faches. Sie gehen davon aus, dass die Naturwissenschaften einen empirischen »Stoff« vorfinden, den sie faktisch in mathematische »Form« bringen können. Man kann daher sagen, »Natur« sei alles empirisch Vorfindliche und zugleich mathematisch Erfassbare.

2. In diesem Sinne sind auch solche empirischen Disziplinen wie die Psychologie und die Soziologie, die sich mit dem Menschen und der menschlichen Gesellschaft befassen, Naturwissenschaften, und zwar Naturwissenschaften im weiten Sinne, wie man vielleicht sagen kann. Die Soziologen werden sich gegen diese Einordnung vermutlich verwahren, die empirischen Psychologen, die Psychometrie betreiben und mit Lernstandserhebungen gegenwärtig Konjunktur haben, eher nicht. Viele von ihnen, vielleicht sogar die Mehrheit, sind davon überzeugt, alle menschlichen Leistungen und Verhaltensweisen messen zu können. Einwände, die sich auf das Unmessbare wie den Traum, die Phantasie oder die Kreativität berufen, werden leicht mit dem Hinweis aus dem Felde geschlagen, dass man das Unmessbare eben messbar machen müsse. Das ist kein Witz, sondern die Herstellung der Messbarkeit ist die Hauptaufgabe des Verfahrens im Vorfeld

der Messungen, etwa bei den internationalen Vergleichsmessungen wie PISA. Hier wird ja, abgesehen von Schülerinnen und Schülern gleichen Alters, sehr Verschiedenes gemessen: Lernleistungen in verschiedenen Ländern mit verschiedenen Traditionen, verschiedenen Schulen, unterschiedlichen Didaktiken, differenten sozialen Milieus usw. Bei so vielen Verschiedenheiten entsteht ein Problem. Zählen kann man nur Gleiches, zum Beispiel Münzen, und Messen ist eigentlich nur ein Zählen des Gleichen (der Maßeinheit, z.B. der Längeneinheit Zentimeter). Wo aber alles verschieden ist, kann man nicht zählen und daher auch nicht messen, es sei denn, dass man sich mit dem Allgemeinen, dem Zählen der vorhandenen Gegenstände und der stattfindenden Ereignisse überhaupt zufrieden geben will. Aber daran ist niemand interessiert. Also muss das Verschiedene gleich und damit messbar gemacht werden.

3. Als erstes muss dafür eine Normgröße gefunden, wenn nicht erfunden werden. Diese Normgrößen sind bei PISA die Mindeststandards und Basiskompetenzen. Sie stellen das Maß dar, nach dem gemessen wird. Zum Messen gehören auch Messinstrumente. Das sind in der Psychometrie die Tests. Testentwicklung ist nach der Definition der Standards die zweite Aufgabe auf dem Weg der »Messbarmachung«. Nun ist es leicht zu sehen, dass schon die Definition der Normgrößen, etwa der sog. Bildungsstandards, von Anfang an ebenso abhängig ist von der Messabsicht wie die Tests. Die Definitionen müssen so ausfallen, dass Messung möglich wird

und entsprechende Tests entwickelt werden können. Das geschieht durch die operationale Definition, die das *definendum* mit Hilfe eines beobachtbaren und dann auch durch Tests kontrollierbaren Verhaltens bestimmt. Die der Normierung dienende Definition des Beobachtungsobjekts bzw. -subjekts ist daher selbst schon »normiert« bzw. »normativ« bestimmt. Der Gegenstand soll nicht so beschrieben und definiert werden, wie er ist, sondern so, dass er gemessen werden kann. Das Interesse an Messbarkeit dirigiert das wissenschaftliche Tun bis ins Einzelne. Es drückt sich in der erwähnten Maxime aus, das Unmessbare messbar zu machen. Ein treffendes Beispiel aus einem anerkannten Handbuch lautet: »Ist ein Phänomen (z.B. ›Faulheit‹) als Forschungsgegenstand ausgewählt worden, muss es messbar gemacht werden, d.h., der Forscher muss eine Vorschrift definieren, wie das jeweilige Ausmaß von Faulheit zu erfassen ist (z.B. durch die Punkte, die Schüler im [...] Anstrengungs-Vermeidungs-Test erhalten)« (Rost 2005, S. 41).

»Konstruktivismus der Forscher«

4. Das hier vielleicht ein wenig grob geschilderte Vor-Verfahren der empirischen Messung im Bildungsbereich sollte einen Punkt genau getroffen haben: die in die Definition der zu messenden Sachverhalte eingehende Bedingung ihrer Messbarkeit, welche die Verwandlung der Sachverhalte in Messkonstrukte zur Folge hat. Die Forscher wissen in der Regel um ihren »Konstruktivismus« und um die »Normativität« ihres Messinteresses. Unaufgeregt und nüchtern sehen sie darin notwendige Elemente der Forschungsmethodik. Die Normativität, die Bestandteil ihrer Forschung ist, betrachten sie als eine Bedingung ihres Vorgehens; sie ist für sie gewissermaßen von theoretischer Art. In der Tat ist der Ausdruck »Normativität« doppeldeutig, denn er kann eine theoretisch-technische Richtschnur (lat. *norma*) bedeuten, nach deren Anweisung man u.a. ein Experiment oder eben eine empirische

Erhebung ausführt, er kann aber auch ein weit über technische Zusammenhänge hinausgehendes rechtliches oder moralisches »Sollen« bedeuten, und zwar ein Sollen, welches auch andere Personen, unter Umständen sogar jedermann und nicht bloß den empirischen Forscher betrifft. Präskriptive Aussagen aus empirischen Daten abzuleiten, das bezweckt wohl kaum ein Empiriker mit seiner Forschung, jedenfalls nicht unmittelbar. Empirie hat für ihn ganz im Gegenteil die Bedeutung eines Abgrenzungskriteriums gegen die nichttechnische Normativität eines die Lebensführung im weitesten Sinne betreffenden Sollens. Unmittelbar wird er nicht vom Sein aufs Sollen schließen. Das ist nun für die abschließenden Bemerkungen das Stichwort: Kann es sein, dass die normativ abstinente bzw. dem Sollen gegenüber neutrale Forschung dennoch *mittelbar* das Handeln und Werten beeinflusst? Kann es sein, dass sie *mittelbar* normativ wird? Darauf weist manches hin (vgl. Koch 2004).

5. Mit einer derartig mittelbaren Normativität laden sich die Standards der vergleichenden Leistungsmessung auf. Unmittelbar stellen sie lediglich Messnormen bzw. Maßeinheiten für Schulleistungstests dar. Nun ist die zyklische Wiederholung der Tests, wie das bei PISA der Fall ist, nur unter der Voraussetzung sinnvoll, dass in den Zwischenzeiten auch nach den Standards unterrichtet wurde. Anders lässt sich eine positive Entwicklung der Messresultate kaum zuordnen und praktikabel deuten. So werden aus Messbedingungen Handlungsvorschriften, aus Leistungsindikatoren Leistungspräskriptoren, und zwar schon bevor sie von den Kultusministerien verbindlich gemacht werden. Im Übrigen kann die empirische Forschung unmittelbare Normativität gewissermaßen »legal« voraussetzen und ihre Erhebungen in den Dienst ihrer normativen Prämissen stellen, sofern deren Normativität als solche gekennzeichnet wird. Auch dafür bietet PISA 2000 ein gutes Beispiel. Nicht nur lässt sich die Erhebung von einem »funktionalistisch orientierten Grundbildungsverständnis« leiten (Deutsches PISA-Konsortium 2001,

S. 17), sondern es wird auch bewusst darauf hingewiesen, dass die PISA-Tests einem didaktischen und bildungstheoretischen Konzept folgen, »das normativ ist« (ebd., S. 19). Ein unmittelbares Normativ stellt drittens auch die ziemlich umständlich beschriebene Absicht der Untersuchung dar, »den Regierungen der teilnehmenden Länder auf periodischer Grundlage Prozess- und Ertragsindikatoren zur Verfügung zu stellen, die für politisch-administrative Entscheidungen zur Verbesserung der nationalen Bildungssysteme brauchbar sind« (ebd., S. 15). Auch in dieser Absichtserklärung äußert sich eine allem Anschein nach unmittelbare Normativität: der Wille, Steuerungswissen bereitzustellen. Von dieser direkten Normativität unterscheidet sich jene mittelbare Normativität, von der hier thematisch die Rede war. Sie hängt mit den methodischen Bedingungen der Messung zusammen. Im Vorfeld der Messung sind die (theoretischen) Messnormen oder Messkriterien operational zu konstruieren, die schon bald den Charakter von Handlungskriterien annehmen, d.h. praktisch normativ werden. Wenn im Feld der menschlichen Angelegenheiten Messung und Messbarmachung zu forschungsleitenden Normen werden, dann werden die Normen der Messung (die Messkriterien) im praktischen Sinne mittelbar normativ. Trotz aller Objektivität und normativen Zurückhaltung tendiert die Messung dahin, dem Gemessenen die Norm zu diktieren und die Sicht auf die Dinge, in diesem Falle auf die Bildung der Fünfzehnjährigen, zu prägen. Die Frage ist daher nicht aus der Luft gegriffen, ob die messbar gemachte und dann auch gemessene Bildung wirklich das ist, was wir eigentlich unter Bildung verstehen (detailliert dazu: Bender 2010, bes. S. 46). Und wenn man die Geschichte der Anthropometrie vor Augen hat, wie sie um die Mitte des 19. Jahrhunderts mit Schädelmessungen begann und von da zur Intelligenzmessung fortschritt, so kann man mit Blick auf die aktuelle Vermessung der Bildung fragen, ob sie nicht Teil einer schon seit langem in Gang gekommenen Fehlmes-

Literatur

Bender, P. (2010): Problematik der Messinstrumente am Beispiel jüngerer Schulstudien. In: Gauger, J.-D./Kraus, J. (Hg.): Empirische Bildungsforschung. Notwendigkeit und Risiko. St. Augustin/Berlin

Deutsches PISA-Konsortium (Hg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen, S. 17

Koch, L. (2004): Normative Empirie. In: Böhm, W. u.a. (Hg.): Kritik der Evaluation von Schulen und Universitäten. Erziehung, Schule, Gesellschaft. Bd. 31. Würzburg, S. 39–55

Gould, St. J. (1988): Der falsch vermessene Mensch. Frankfurt a.M. (The Mismeasure of Man. New York 1981)

Rost, D.H. (2005): Interpretation und Bewertung pädagogisch-psychologischer Studien. Weinheim, S. 41

Abstract

Der Beitrag erschließt die der scheinbar neutralen empirischen Forschung inhärente Normativität. Am Beispiel der PISA-Studien werden unmittelbare und mittelbare normative Einflussnahmen empirischer Forschung aufgezeigt: (1) ihr Wille, steuerungsrelevantes Wissen bereitzustellen, (2) die Normativität ihres didaktischen und bildungstheoretischen Konzepts, (3) die aus der Untersuchungsabfolge abgeleitete Idee, den Unterricht in den Phasen zwischen den Erhebungen dem Konzept anzupassen und (4) die im Vorfeld der Messung erforderliche Operationalisierung theoretischer Messnormen im Sinne ihrer quantitativen Messbarkeit, die den Charakter von Handlungskriterien annehmen können.



Prof. em. Dr. Lutz Koch lehrte Allgemeine Pädagogik an der Universität Bayreuth.

Kontakt: lutz.koch@uni-bayreuth.de