

Selbst ist die Schule!

Fremd- versus Selbstevaluation

VON GEORG LIND

Lind, G. (2011). Selbst ist die Schule. Grundschule, April 2011, Heft 4, 24-26.

Die Fremdevaluation von Schulleistungen durch Vergleichstests wie VERA kostet viel Geld, Aufmerksamkeit und Arbeitskräfte. Welche Art von Evaluation brauchen wir wirklich?



Verstaubte Klassenzimmer in Schulmuseen zeugen von einer Schulwirklichkeit, die geprägt war von militärischer Disziplin, Lernen im Gleichschritt, festem Zeitakt, Noten. Leider haben diese Merkmale Schulen auch heute noch im Griff – das zeigt sich auch an der Tendenz zur Fremdevaluation.

Seit etwa zehn Jahren werden in Deutschland im großen Stil Vergleichsuntersuchungen durchgeführt, um dadurch Hinweise für die Verbesserung des Lernens an unseren Schulen zu gewinnen. Schulleistungsvergleiche wie VERA, IGLU und PISA kosten viel Geld. Sie ernähren inzwischen eine global operierende Testindustrie (vgl. Bracey 2005; Flitner 2006).

Die Landesämter für Schulentwicklung sind zum großen Teil zu Evaluationsagenturen umgewidmet worden, was sich vielerorts auch in einem neuen Namen zeigt. In Berlin wurde ein zentrales Institut für Qualität im Bildungssystem (IQB) finan-

ziert, und viele Lehrstühle für Pädagogik und Erziehungswissenschaft an deutschen Hochschulen wurden und werden noch immer zu Stellen für empirische Bildungsforschung umgeformt, meist besetzt mit Mitgliedern oder Beratern von Vergleichsstudien. Testleistungsvergleiche binden auch auf allen Ebenen unseres Schulsystems – angefangen von der Lehrperson bis in die Ministerien hinein – viel Geld, Aufmerksamkeit und Arbeitskraft, die für andere Aufgaben in unseren Schulen fehlen.

Lernzeit für andere Fächer und höherwertige Kompetenzen (wie Schreiben, Textproduktion, mathematisches Verständnis, kreative Fertig-

keiten, Kunst, Musik, Sport, freies Lernen etc.) geht verloren. Einsparungen treffen auch die Lehrerbezahlung, die Lehrerbildung und -fortbildung, die Unterrichtsentwicklung, die Schulbuchverbesserung, die Schulleiterentwicklung und anderes mehr.

MUSS EVALUIERT WERDEN?

Brauchen wir überhaupt Evaluation und wenn ja, welche Art von Evaluation? Jeder, der sich mit unseren Schulen beschäftigt hat, weiß: Die meisten Beteiligten geben ihr Bestes und viele machen oft mehr als von ihnen verlangt wird. Dennoch gibt es einen breiten Konsens, dass es an verschie-

denen Stellen klemmt und Schülerinnen und Schüler besser lernen könnten, wenn bestimmte Dinge verbessert werden würden. Die Struktur unseres Bildungswesens geht großteils auf den Beginn des 19. Jahrhunderts zurück, als nach Ende des preußisch-französischen Krieges in Deutschland zum ersten Mal ein staatliches Bildungssystem aus dem Boden gestampft wurde.

Es lag nahe, dass sich die Bildungspolitik damals am Aufbau des Militärs orientierte. Die Lehrer rekrutierte man meist aus den nun arbeitslosen Unteroffizieren und Feldwebeln und mit ihnen die militärische Ordnung: Lernen im Gleichschritt, Zeittakt des Unterrichts, Ausrichtung der Sitzreihen, Disziplin und Unterordnung, Verwaltungshierarchien und die Noten – also das Paradebeispiel für vergleichende Fremdevaluation.

All diese Strukturmerkmale unserer Schulen, die das Schulleben auch heute noch im Griff haben, sind Bedürfnissen der staatlichen Verwaltung geschuldet, aber nicht pädagogisch begründet. Seit Jahrzehnten zeigt auch die pädagogisch-psychologische Forschung, dass Kinder wesentlich schneller und mehr lernen, wenn sie angeleitet und unterstützt von kompetenten Lehrkräften mit anderen Kindern zusammen und ohne ständigen Zeitdruck, aber frei von Fremdvergleichen (Noten), kleinschrittigen Lernanweisungen, Angst vor Strafen, und extrinsischen Belohnungen lernen dürfen (vgl. Kohn 1999; Czerny 2010; Leppert 2010).

Was bringen Vergleichstests?

PISA, IGLU, VERA etc., die selbst ganz in dieser militärisch-bürokratischen Tradition stehen, haben in dieser Hinsicht keine neuen Erkenntnisse bringen können. Aber selbst, wenn man sie nur nach ihren eigenen Versprechungen misst, haben Vergleichsstudien die Erwartungen bislang meist enttäuscht. Obwohl immer wieder so getan wird, als sagten Vergleichstests etwas über die Lernleistung der Schülerinnen und Schüler aus, sind die getesteten Fähigkeiten noch nicht einmal annäherungsweise repräsentativ für das, was Schüler heute lernen oder lernen müssen, um im Leben zu bestehen.

Viele wichtige Fähigkeiten wie musische, physische, künstlerische,

praktische, sprachliche, soziale, demokratische und moralische Kompetenzen kommen in den Tests nicht vor – nicht weil es nicht möglich ist, sondern weil es zu aufwändig ist, sie zu testen und weil viele Bildungspolitiker einen pädagogischen Tunnelblick haben. Sie begnügen sich damit, Lesefähigkeit testen zu lassen, aber nicht kritische Lesefähigkeit. Man muss nur Anweisungen und Tabellen lesen und wiedergeben können.

Daran ändern auch Mathematik- und Naturwissenschaftstests wenig. Bei PISA korrelieren diese Tests so hoch mit den Ergebnissen der Lesetests, dass man auch bei diesen von Tests der Lesefähigkeit sprechen muss. Schwierig sind sie meist nur wegen Aspekten, die mit Mathematik und Naturwissenschaft wenig zu tun haben: nämlich

- wegen überlanger Lesetexten,
- wegen vieler verwirrender Informationen, die teils absichtlich, teil unabsichtlich in die Aufgaben eingebaut werden;
- wegen einer für viele Kinder ungewohnten Aufgabendarbietung;
- wegen des Hintergrundwissens, das zur Lösung vieler Aufgaben benötigt, aber in der Schule nicht gelehrt wird;
- wegen des hohen Zeitdrucks, der Stressfähigkeit und soziale Dominanz auf die Probe stellt (vgl. Jahnke/Meyerhöfer 2007). Ganz offensichtlich leiden unter diesen Hindernissen besonders Kinder aus ungünstigen sozialen Milieus, die wenig Unterstützung außerhalb der Schule bekommen, um diese Testaufgaben erfolgreich bewältigen zu können.

TEACHING TO THE TEST

Der heimliche Lehrplan der Vergleichstests ist also Lesen von Vorschriften, Gehorsam gegenüber Testanweisungen und Stressfestigkeit. Kaum etwas anderes wird durch Testpunktwerte belohnt. Kompetenzen, die für das Leben in einer modernen Gesellschaft unverzichtbar sind, wie das Verstehen, das Anwenden und die Verantwortung von Vorschriften und eigenen Entscheidungen, werden in den meisten Vergleichsstudien völlig außer Acht gelassen. In dem Maß, in dem sich die Schule diesem

SCHNELL GELESEN

Vergleichstests wie VERA kosten nicht nur viel Geld und Arbeitskraft, sondern sagen auch nichts über die tatsächliche Leistung der Schülerinnen und Schüler aus. Um das Lernen in den Schulen wirklich zu fördern, muss sich die Evaluation auf Methoden und Maßnahmen richten und von den Handelnden selbst konzipiert und verantwortet werden. Durch die Fähigkeit zur realistischen Selbsteinschätzung verschiedener Kompetenzen anhand objektiver Tests wird so auch das Selbstbewusstsein der Kinder gestärkt.

Lehrplan anpasst (Teaching to the Test), um „gute Evaluationswerte“ zu erzielen, versagt sie als Motor unseres Wohlstandes und der Demokratie. Vergleichsstudien können von ihrer Anlage her auch kaum Hinweise auf Ursachen geben.

Der vielleicht schlimmste Geburtsfehler der meisten Vergleichstests liegt darin, dass sie als personenbezogene Fremdevaluation konzipiert werden, und daher auf allen Ebenen Korruption auslösen. Eingeklemmt zwischen drohenden Strafen bei schlechten Evaluationsergebnissen und einem Mangel an Ressourcen und Fähigkeiten bleibt vielen Beteiligten nichts anderes übrig, als zu tricksen und zu schwindeln, und zwar, wie wir wissen, auf allen Ebenen, vom Schüler bis zum Minister.

- Damit die Statistik stimmt, werden leistungsschwache Schüler am Testtag nach Hause geschickt.
- Sie werden aus der Statistik gelöscht.
- Sie werden bei der Versetzung in die Vergleichsstudien-Klassen erst zurück gestellt und dann in die übernächste Klasse versetzt.
- Oder sie werden gleich ganz aus der Schule gedrängt und in Sonderschulen verwiesen, wo sie von der Vergleichsstatistik nicht erfasst werden (es ist auffällig, dass die Bundesländer mit den besten PISA-Werten auch die höchsten Schülerzahlen in Sonderschulen haben).

Es kommt auch vor, dass Testantworten von den Lehrern nachkorrigiert und Tests so konzipiert werden, dass die gewünschte Zahl von Schülern sie bestehen. Audrey Amrein und David Berliner (2002) zeigen, dass die Verfälschung der Evaluationsstudien durch diese Betrugereien so massiv sind, dass ihre Daten wertlos werden und zu falschen Schlüssen verleiten,

wie bereits vor 30 Jahren der renommierte Evaluationsforscher Donald T. Campbell (1969) prognostiziert hatte. Die Testleistungen der amerikanischen Schüler sind auch in 45 Jahren kaum gestiegen; stattdessen gibt es viele unerwünschte Nebenwirkungen (vgl. Lind 2009a). Dieses Versagen der Fremdevaluation sollte uns eine Warnung sein. Wir sollten daraus lernen.

UND SELBSTEVALUATION?

Wenn die Evaluation das Lernen in unseren Schulen befördern soll, muss sie sich auf Methoden und Maßnahmen richten (also nicht auf Personen) und von den Handelnden selbst konzipiert und verantwortet (und nicht nur von ihnen bezahlt und durchgeführt) werden. Eine methodenorientierte Selbstevaluation ist nicht nur effektiver, sondern entspricht auch unserem demokratischen Selbstverständnis (vgl. Brügelmann 2004). Demokratie ist kein beliebiges Ideal, sondern eine Lebensform, die besser mit Problemen umzugehen weiß als der autoritäre Ordnungsstaat der Vergangenheit.

Selbstevaluation ist nichts Neues; sie gibt es schon lange. Sie hat, bevor sich die Politik des Themas Evaluation bemächtigte, die wissenschaftliche Diskussion bestimmt.

Selbstevaluation muss Bestandteil von gezielten Bemühungen um Fehlerbehebung und Qualitätsverbesserung sein. Nur wenn klare Ziele formuliert werden, kann es Tests geben, die zielvalid sind, also das messen, was sie messen sollen, und nur dann können Erhebungsdesigns definiert werden, die Aussagen über Wirkungszusammenhänge zulassen, und deren Kosten in einem vernünftigen Verhältnis zu dem erwarteten Nutzen stehen.

Selbstevaluation kann jeder Mensch praktizieren, um seine Fähigkeiten realistisch einzuschätzen. Schülerinnen und Schüler können damit ihre Lernleistungen besser einschätzen, und zwar nicht mit einem problematischen Bezug auf die Mitschüler, sondern in Bezug auf die Lernanforderungen, denen sie sich stellen. Kann ich die Zahlen bis 100? Kann ich das Alphabet herunterzählen und schreiben? Kann ich eine Katze malen? Die Fähigkeit zur realistischen

Selbsteinschätzung verschiedener Kompetenzen durch geeignete, objektive Tests stärkt das Selbstbewusstsein der Kinder und reduziert ihre Abhängigkeit von (falschen) Autoritäten.

Dasselbe gilt für ihre Lehrerinnen und Lehrer. Wenn sie ihre Lehrwirksamkeit selbst objektiv zu analysieren vermögen, können sie falsche Beurteilungen durch Vorgesetzte oder Eltern korrigieren. Dasselbe gilt auch für Vorgesetzte und Bildungsminister. Schulleiterinnen und Schulleiter, die die Qualität ihrer Leitungsarbeit regelmäßig gezielt überprüfen, gewinnen an Selbstbewusstsein und Handlungskompetenz. Bildungsminister können besser erklären, warum bestimmte Reformmaßnahmen notwendig sind und in welchem Verhältnis der Investitionsaufwand zu den erwarteten Gewinnen steht. Auch Selbstevaluation stützt sich zum Teil auf Testdaten von Schülern und Lehrern; aber diese sind anonymisiert und bieten daher kaum Grund für Tricksereien. Man würde dadurch nur sich selbst belügen und den eigenen Einsatz an Zeit und Geld ad absurdum führen.

Richtig gemachte Selbstevaluation führt zur kontinuierlichen Verbesserung des Unterrichts, wie ich das in einem langjährigen Selbstversuch gezeigt habe (vgl. Lind 2009b). Damit sie dies tun kann, muss sie jedoch professionell gemacht werden (vgl. Lind 2004). Selbstevaluation im Bereich der Bildung ist schwerer als zum Beispiel im Handwerk und in der Technik, wo die Evaluation der eigenen Arbeit zum Alltag gehört, aber sie ist nicht unmöglich. Es muss mehr Forschungs- und Entwicklungsarbeit auf diesem Gebiet geleistet werden, statt alle verfügbaren Mittel in Vergleichsstudien zu stecken. Selbstevaluation muss zum integralen Bestandteil der Schulbildung und der Lehrerbildung gemacht werden. Das alles steckt hierzulande noch in den Anfängen. In Ländern wie Schweden und Finnland ist man mit Selbstevaluation in der Schule bereits weiter fortgeschritten als bei uns, weil erkannt wurde, dass „jede Evaluierung, die nicht auf einer Vertrauensbasis, sondern auf einer Kontrollbasis steht, ... vergebens“ ist (Domisch 2007, S. 33).

DER AUTOR

Dr. Georg Lind

ist Professor am Fachbereich für Psychologie der Universität Konstanz.

LITERATUR

- Amrein, A./Berliner, D.:** Highstakes testing, uncertainty, and student learning. 2002. Unter: <http://epaa.asu.edu/epaa/v10n18/>
- Bracey, G. W.:** No child left behind: Where does the money go? In: Education Policy Studies Laboratory. June 2005. Unter: <http://edpolicylab.org>
- Brügelmann, H.:** Kerncurricula, Bildungsstandards und Leistungstests: Zur unvergänglichen Hoffnung auf die Entwicklung der guten Schule. In: Vierteljahresschrift für wissenschaftliche Pädagogik, Heft 4/2004, S. 415-441
- Campbell, D.T.:** Reforms as experiments. In: American Psychologist, Heft 24/1969, S. 409-429
- Czerny, S.:** Was wir unseren Kindern in der Schule antun ... und wie wir das ändern können. München 2010
- Domisch, R.:** Schulentwicklung durch Qualitätsanalyse. In: forum bildung, didacta - die Bildungsmesse 2007 Köln, 27.2.-3.3.07
- Flitner, E.:** Pädagogische Wertschöpfung. Zur Rationalisierung von Schulsystemen durch public-private-partnerships am Beispiel von PISA. In: Oelkers, J./Horlacher, R./Casale, R. (Hrsg.): Rationalität und Bildung. Studien im Umkreis Max Webers. Zürich 2006, S. 245-266
- Jahnke, T./Meyerhöfer, W. (Hrsg.):** PISA & Co. Kritik eines Programms. Hildesheim 2007
- Kohn, A.:** Punished by Rewards. The Trouble with Gold Stars, Incentive Plans, A's, Praise, and other Bribes. Boston 1999
- Leppert, U.:** Ich habe eine Eins! Und du? Von der Notenlücke zur Praxis einer besseren Lernkultur. München 2010
- Lind, G.:** Jenseits von PISA - Für eine neue Evaluationskultur. In: Institut für Schulentwicklung PH Schwäbisch Gmünd (Hrsg.): Standards, Evaluation und neue Methoden. Reaktionen auf die PISA-Studie. Baltmannsweiler 2004, S. 1-7
- Lind, G.:** Amerika als Vorbild? Erwünschte und unerwünschte Folgen aus Evaluationen. In: Bohl, T./Kiper, H. (Hrsg.): Lernen aus Evaluationsergebnissen - Verbesserungen planen und implementieren. Bad Heilbrunn 2009a, S. 61-79
- Lind, G.:** Favorable learning environments for moral development - A multiple intervention study with nearly 3.000 students in a higher education context. Paper presented at the annual meeting of AERA in San Diego, April 13-17, 2009b

IM INTERNET

www.uni-konstanz.de/ag-moral/evaluation/itse_home.html: Informationen über das Projekt Improvement of Teaching through Self-Evaluation (ITSE)