

Rittelmeyer, Christian

## **Bildungsreform und Evaluation. Ein Beitrag zur aktuellen Diskussion über "Leistung" und ihre "Messung"**

*Die Deutsche Schule 92 (2000) 1, S. 10-12*



Quellenangabe/ Reference:

Rittelmeyer, Christian: Bildungsreform und Evaluation. Ein Beitrag zur aktuellen Diskussion über "Leistung" und ihre "Messung" - In: Die Deutsche Schule 92 (2000) 1, S. 10-12 - URN: urn:nbn:de:0111-pedocs-275917 - DOI: 10.25656/01:27591

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-275917>

<https://doi.org/10.25656/01:27591>

### **Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### **Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

### **Kontakt / Contact:**

**peDOCS**  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Digitalisiert

Mitglied der

  
Leibniz-Gemeinschaft

---

Christian Rittelmeyer

## Bildungsreform und Evaluation

Ein Beitrag zur aktuellen Diskussion über „Leistung“ und ihre „Messung“

---

Es ist eigenartig: Obgleich die Beiträge von Hans-Georg *Herrlitz* und Dieter *Weiland*<sup>1</sup> kontroverse Positionen schildern, habe ich sie doch gleichermaßen mit sehr viel Zustimmung gelesen. Über Gründe dafür nachdenkend, warum ich nicht für die eine oder andere Position Partei ergreifen konnte und überdies auch – trotz aller Zustimmung – ein ungutes Gefühl verspürte, wurde mir ein Problem der zwei Texte deutlich.

Die von Hans-Georg Herrlitz vorgetragene Position, eine Evaluation der Leistungsfähigkeit einzelner Schulen zum Zweck ihres Vergleichs sei unerlässlich, hat meine uneingeschränkte *Zustimmung*. Bei Herrlitz wird jedoch nicht ganz deutlich, ob er diese Evaluation grundsätzlich, oder nur aus Gründen der bildungspolitischen Opportunität für unerlässlich hält. Ich gehe davon aus, dass aus Steuergeldern finanzierte Institutionen die Verpflichtung haben, *Rechenschaft* über den effizienten Einsatz der öffentlichen Mittel abzugeben – und zwar nicht in Form je individueller Bekundungen der eigenen Leistungsfähigkeit, sondern in Gestalt von Prüfungsverfahren, die einen gerechten Vergleich verschiedener Bildungseinrichtungen zulassen.

Die Frage ist freilich: Worauf bezieht sich der *Effizienzbegriff*, und wie kann eine Evaluation der verschiedenen Qualitäten, die eine Schule an Schülerinnen und Schüler vermitteln soll, beschaffen sein?

Es mag sinnvoll sein, in diesem Zusammenhang an eine Diskussion zu erinnern, die *in den 60er Jahren* sowohl über die Struktur intellektueller Leistungen als auch über deren Diagnostik geführt wurde. Das damals von J. P. *Guilford* (u. a. in seinem Buch „*Persönlichkeit*“) vorgestellte komplexe Intelligenzmodell machte deutlich, dass die in den Schulen favorisierten Intelligenzbegriffe des „*konvergenten Denkens*“ nur einen kleinen Teil der für Alltagshandlungen wichtigen intellektuellen Leistungen ausmachen. So wird beispielsweise den Fähigkeiten des flüssigen, anpassungsfähigen Denkens, der Produktion verschiedenartiger Lösungsmöglichkeiten für ein gegebenes Problem, der Assoziationsflüssigkeit oder der Ausdrucksflexibilität, der Identifikation bildlicher Bedeutungen oder symbolischer Zeichen unter dem Titel des „*divergenten Denkens*“ Aufmerksamkeit zuteil. Es sei, so Guilford, ein wesentlicher Fehler bisheriger Intelligenzuntersuchungen, ein weitaus zu großes Vertrauen in *verbales* Testmaterial zu setzen.

---

<sup>1</sup> Vgl. „*Offensive Pädagogik*“ in Heft 4/99 der „*Deutschen Schule*“

Ein weiterer Gedanke muss hier hinzukommen: Es war damals insbesondere das Verdienst von Klaus *Holz*kamp, eines kritischen Psychologen, in einer gründlichen wissenschaftstheoretischen Studie mit dem Titel „Theorie und Experiment in der Psychologie“ (1964) die wichtige Frage präzise diskutiert zu haben, unter welchen Bedingungen eigentlich bestimmte Tests (z. B. Schulleistungs- oder Intelligenztests) *repräsentativ für Alltagssituationen* genannt werden können, auf deren Bewältigung jene getesteten Fähigkeiten vorbereiten sollen. Holz kamp diskutierte das Problem als Frage nach der *Repräsentanz* psychologischer Experimente (zu denen auch Tests gehören), die Testtheorie benutzt hier Begriffe wie „externe Validität“ oder „Alltagsrelevanz“ psychologischer Tests.

Es ist eigentümlich, dass diese Einsichten *bis heute in die Evaluationsverfahren keinen Eingang gefunden* haben – es dominieren nach wie vor „Paper-and-pencil“-Tests, also mündliche oder schriftliche Befragungen. Genau genommen besitzen diese Verfahren im günstigen Fall aber nur Repräsentanz für Situationen des gleichen Typs: in der Regel also für Konstellationen, in denen mündlich oder schriftlich im Stil des konvergenten Denkens reagiert werden muss.

Es ist bemerkenswert, dass mit Blick auf diese Diskussionen neben den „zweidimensionalen Tests“ (d. h. der schriftlichen Befragung) keine „dreidimensionalen Tests“ entworfen worden sind, in deren Arrangement Schülerinnen und Schüler nicht nur verbal, sondern *als ganze Menschen* agieren müssen. Dass es zu solchen Entwicklungen nicht gekommen ist, mag einerseits an den relativ höheren Kosten solcher Verfahren liegen, auch an ihrer aufwendigen Durchführung, andererseits aber auch an der Phantasielosigkeit und an einem tradierten Denktypus vieler Testkonstrukteure.

Solche Verfahren sind aber m. E. die einzige Möglichkeit, um einerseits über eine unter Umständen problematische bloße Selbstevaluation der einzelnen Schule hinauszukommen und (im testtheoretischen Sinn) *objektive* Maßstäbe zu gewinnen, andererseits aber auch dem Problem der *externen Validität* gerecht zu werden.

Ich stimme Dieter Weilands Urteil (These 5) zu, dass neuere Studien wie die *TIMSS* oder *BIJU* einem erweiterten Leistungsbegriff im zuvor genannten Sinn allein nicht gerecht werden; wertvoll sind sie m. E. dennoch als *ergänzende* Studien (die *BIJU*-Studie hat immerhin zunächst einmal mit Gründen das Vorurteil in Frage gestellt, die Gesamtschulen seien auf dem Gebiet des sozialen Lernens besonders erfolgreich).

Ich könnte mir ein *Drei-Säulen-Modell der Evaluation* vorstellen, das den genannten Anforderungen gerechter wird als die bisher üblichen Papier-und-Bleistift-Tests:

- Die *üblichen Fragebogen und Tests* könnten eine erste Säule bilden, weil sie zweifellos wichtige, aber nicht die allein wichtigen Kompetenzen erfassen (diese Aussage gilt mit den üblichen testtheoretischen Einschränkungen, über die im Einzelfall diskutiert werden muss: Wie sind die Gütekriterien des Verfahrens beschaffen, an welcher Stichprobe wurde wie validiert etc.);
- Eine zweite Säule könnte durch die noch zu entwickelnden *Handlungstests* gebildet werden. Hier wären z.B. Mathematik-Leistungen im Teilbereich Tri-

gonometrie nicht nur in Fragebogenbeantwortungen, sondern in konkreten Handlungsaufgaben z.B. einer Feldvermessung zu lösen. Ich wähle dieses Beispiel mit Blick auf eine nicht erfundene, heute an der Universität Statistik unterrichtende Person, die in den Mathematik-Klausuren der Schule stets schlechte Noten erhielt – bis sie in einem Feldmesspraktikum ihre entsprechenden Fähigkeiten erstmals entdeckte; bis heute bemerkt sie ihre Leistungsfähigkeit besonders in solchen Handlungssituationen. (Der Lehrer hatte damals auch den schönen Einfall, eines Nachts Mädchen und Jungen zu trennen, die Mädchen – unter ihnen die Angebetete des betreffenden Jungen – starteten um Mitternacht eine Leuchtrakete aus entfernten Waldgebieten; wie die Jungen da mit Theodoliten und anderen Messgeräten Winkel und Steigungen berechneten, um den Standort der Mädchen möglichst präzise bestimmen und sodann aufsuchen zu können!). Mir ist bewusst, dass die testtheoretisch befriedigende Konstruktion solcher Handlungstests schwierig ist – aber man könnte sich nun endlich doch einmal um entsprechende Planungen und Möglichkeiten kümmern, um die Evaluations-Diskussion voranzubringen.

- Eine dritte Säule schließlich müsste durch *schulspezifische Prüfverfahren* gebildet werden: diese Evaluation wäre kasuistisch angelegt und könnte sich so auch dem besonderen Profil der einzelnen Schule oder auch einer bestimmten Schulform wie der Integrierten Gesamtschule oder der Waldorfschule anpassen. Eine solche dritte Säule ist, wie wohl kaum erläutert werden muss, für jede ernst zu nehmende Evaluation unerlässlich, die Leistungsphänomene wirklich aus dem Schulkontext heraus verständlich machen soll.

Es wäre eine sinnvolle Aufgabe für die zuständigen Gremien in der DGfE (oder auch in der GEW?!), eine derartige *Evaluations-Diskussion* unter Mitarbeit kompetenter Schulpraktikerinnen und -praktiker zu institutionalisieren und entsprechende Vorschläge für die Schulevaluation zu erarbeiten. Verzichtet man in diesem Bereich auf produktive Ideen zugunsten von Abwehrgefechten gegen die vermeintlich unbrauchbare tradierte Evaluationstechnik, dann könnte man in der gegenwärtigen Lage auch sehr rasch aus der Mitgestaltung von Evaluations-Verfahren ausgeschlossen werden.

*Christian Rittelmeyer*, geboren 1940, Dr. phil., Professor für Erziehungswissenschaft am Pädagogischen Seminar der Universität Göttingen. Arbeitsschwerpunkte: Pädagogische Anthropologie, Methoden der Erziehungswissenschaft, Pädagogische Psychologie, Neue Formen der Didaktik;  
Anschrift: Pädagogisches Seminar der Universität, Baurat-Gerber-Str. 4, 37073 Göttingen;  
Email: CRittel@gwdg.de