

Inhalt

0	Auftrag und Kontext der Expertise	4	2	An welchen Maßstäben sollen Leistungen gemessen werden? (Bezugsnormen)	27
0.1	Ansatz und Aufbau des Gutachtens	4	2.1	Wo steht ein Schüler im Vergleich zu anderen (kollektive Norm/Gruppenorientierung)	28
0.2	Datengrundlage des Gutachtens	6	2.2	Wo steht ein Schüler auf dem Weg zum Lernziel? (Sachnorm/Kriteriumsorientierung)	29
0.3	Historischer Rückblick und gesellschaftlicher Kontext	7	2.3	Welche Fortschritte hat ein Schüler gemacht? (individuelle Norm/Entwicklungsorientierung)	30
0.4	Die Situation in den Bundesländern: ein Überblick	9	2.4	Zwischenbilanz zu »Bezugsnormen«	31
0.5	Blicke über den Zaun: Die internationale Situation	13			
1	Mit welchen Verfahren werden Leistungen erfasst?	15	3	Wie werden verschiedenen Formen der Leistungsbeurteilung umgesetzt, und welche Wirkungen haben sie?	32
1.1	Wie gut erfassen Leistungsbeurteilungen, was sie erfassen sollen? (Validität)	15	3.1	Wie weit werden Ziffernnoten und Verbalgutachten ihren eigenen Ansprüchen gerecht?	32
1.1.1	Wie gut sind die Kriterien für Leistungsbeurteilungen inhaltlich abgesichert?	16	3.2	Welche (Neben-)Wirkungen haben verschiedene Beurteilungsformen?	34
1.1.2	Wie gut stimmen Beurteilungen aus verschiedenen Quellen überein?	17	3.2.1	Gibt es einen Zusammenhang zwischen Unterrichtskonzept und Beurteilungsform?	34
1.1.3	Wie genau lässt sich aus der Beurteilung von Leistungen deren zukünftige Entwicklung vorhersagen (prognostische Validität)	18	3.2.2	Beeinflusst die gewählte Beurteilungsform das Unterrichtsklima?	34
1.1.3.1	Kindergarten > Schulerfolg	18	3.2.3	Beeinflusst die gewählte Beurteilungsform zentrale Merkmale der Persönlichkeitsentwicklung?	35
1.1.3.2	Schule > Fachleistungen über die Schuljahre hinweg	19	3.2.3.1	Beeinträchtigen oder stützen Ziffernnoten bzw. Verbalgutachten die Lernmotivation?	35
1.1.3.3	Schule > Studien-/Ausbildungserfolg	20	3.2.3.2	Verringern oder vergrößern Ziffernnoten bzw. Verbalgutachten die Schul- und Prüfungsangst?	37
1.1.3.4	Studium/Ausbildung > Berufserfolg	21	3.2.3.3	Schädigen oder stärken Ziffernnoten bzw. Verbalgutachten das Selbstkonzept?	38
1.1.4	Zwischenbilanz zu »Validität«	22	3.2.4	Belasten oder fördern Ziffernnoten bzw. Verbalgutachten die Leistungsentwicklung?	39
1.2	Wie unabhängig sind Beurteilungen von persönlichen Einflüssen? (Objektivität)	22	3.2.5	Zwischenbilanz zu »Wirkungen«	40
1.2.1	Objektivität des Lehrerurteils	22	4	Wie gut erfüllen Ziffernnoten und Verbalgutachten wichtige Funktionen aus der Sicht der Betroffenen?	40
1.2.2	Kann der Einsatz standardisierter Tests das Objektivitätsproblem lösen?	24	4.1	Einschätzungen von LehrerInnen	40
1.2.3	Wie weit lässt sich das Lehrerurteil objektivieren?	25	4.2	Einschätzungen von SchülerInnen	42
1.2.4	Zwischenbilanz zu »Objektivität«	26	4.3	Einschätzungen von Eltern	44
1.3	Wie verlässlich sind verschiedene Beurteilungsverfahren? (Reliabilität)	26	4.4	Einschätzungen von Arbeitgebern	47
1.3.1	Die Zuverlässigkeit des Lehrerurteils	26	4.5	Einschätzungen in der Öffentlichkeit	47
1.3.2	Die Zuverlässigkeit von Tests	27	4.6	Zwischenbilanz zu »Einschätzungen«	49
1.3.3	Zwischenbilanz zu »Reliabilität«	27	5	Rechtfertigt der Ertrag aufwändigere Formen der Erhebung und Bewertung von Leistungen?	50
1.4	Fazit	27			

6	Zwischenbilanz und pädagogische Folgerungen	52
6.1	Grundlegende Einwände	52
6.1.1	Genereller Verzicht auf eine Rückmeldung zu Leistungen?	52
6.1.2	Verzicht auf eine Zertifizierung nach außen?	53
6.1.3	Verzicht auf Ziffernnoten als Form der Beurteilung?	53
6.2	Keine Beurteilungsform erfüllt alle Anforderungen - einfache Auswege aus dem Bewertungsdilemma gibt es nicht	53
6.3	Daten aus verschiedenen Erhebungsverfahren sind miteinander zu verbinden	54
6.4	Bewertungen müssen auf unterschiedliche Bezugsnormen bezogen werden	55
6.5	In dialogischer Form sollten Fremd- durch Selbsteinschätzungen ergänzt werden	55
7	Fazit und bildungspolitische Bewertung	58
8	Literaturnachweise, weiterführende Literatur und Abbildungsverzeichnis	60

Wie bei allen pädagogischen Fragen (und sozialen Phänomene generell) ist die Befundlage zu Noten nicht auf einen einfachen Nenner zu bringen. Formen der Leistungsbeurteilung wirken unterschiedlich, je nachdem *wie* und in welchem *Kontext* sie eingesetzt werden. Für Folgerungen aus dem Forschungsstand kommt es deshalb darauf an, von welcher Basisannahme man ausgeht: Wer die Beweislast für Veränderungen bei den Reformern sieht, kann zu einer anderen Einschätzung kommen als jemand, der normativ die Förderung des Einzelnen als zentrale Norm und noch uneingelöste Aufgabe der Schule sieht. Vor diesem Hintergrund lässt sich als Ergebnis unserer Analysen festhalten:

▲ Wer an Ziffernnoten festhalten will, weil sie angeblich objektiv und vergleichbar seien bzw. erforderlich, damit SchülerInnen sich auf die Anstrengungen des Lernens einlassen, findet in der Empirie keine stützenden Belege für seine Position.

▲ Auch diejenigen, die Verbalgutachten ablehnen, weil sie angeblich negative Auswirkungen auf die Lernbereitschaft und den fachlichen Lernerfolg der SchülerInnen haben, können sich auf keine empirischen Daten stützen.

▲ Wer andererseits hofft, ohne zusätzliche Maßnahmen, d.h. allein durch die Verordnung von Verbalgutachten, Lernbereitschaft und Lernerfolg von SchülerInnen verbessern zu können, wird durch die Befunde zur bisherigen Beurteilungspraxis und ihre Wirkungen ernüchtert. Ohne eine pädagogische und didaktische Öffnung des Unterrichts und ohne die Sicherung bestimmter Rahmenbedingungen bleibt eine Veränderung der Bewertung meist erfolglos.

▲ Diejenigen aber, die mit dem Verzicht auf Ziffernnoten pädagogische Ziele verfolgen, können mit einer Verbesserung der Unterrichtssituation und des Lernerfolgs, vor allem der schwächeren SchülerInnen, rechnen - sofern sie bereit sind, als LehrerInnen einen höheren Aufwand zu leisten, als Schulverwaltung mehr in die Fortbildung und Unterstützung der LehrerInnen zu investieren und als BildungspolitikerInnen den Selektionsdruck im System zu verringern.

Notabene

Ein zentrales Problem in der Diskussion ist die unklare Begrifflichkeit.

In dieser Expertise verstehen wir - sofern nicht ausdrücklich etwas anderes gesagt wird - unter »Noten« bzw. »Zensuren« *Ziffernnoten*, die zur *formellen* Beurteilung verwendet werden, z.B. bei Klassenarbeiten oder in Zeugnissen.

»Leistungsbeurteilung« verwenden wir als Obergriff für die *Beschreibung* und die *Bewertung* von Leistungen - zwei unterschiedliche Formen ihrer Rückmeldung, die auch in der Umsetzung sorgfältig zu trennen sind.

¹ Diese Expertise geht zurück auf Vorarbeiten in Seminaren an der Universität Siegen (zum Teil auch publiziert) von Erika Brinkmann (2004; 2006), Hans Brügelmann (1980; 2000a+b; 2003a+b; 2005a, Kap. 27, 29, 56-60) und Barbara Müller-Naendrup (2005). Hilfreich waren auch die aktuellen Überblicke in: Valtin (2002a); Jachmann (2003, Kap.2); Barntitzky/Speck-Hamdan (2004); Beutel (2005).