

Kees Both (Übersetzung aus dem Niederländischen: Hartmut Draeger)

Schulstandards und –tests

Internationale Kritik an unpädagogischen Leistungsmessungen

Aus: KINDERLEBEN. Zeitschrift für Jenaplan-Pädagogik H.21, Juli 05, S.5-18

Vorbemerkungen

Die Bildungspolitik in den Niederlanden (wie auch in Deutschland) und die Diskussion darüber sind oft zu einem Teil ein Echo von dem, was auch anderswo „gespielt wird“. Da werden internationale Trends sichtbar, und internationale vergleichende Studien von schulischen Leistungen haben oft großen Einfluss auf die Entscheidungen, die getroffen werden, auch was die Instrumente für Qualitätssicherung betrifft. Die Tendenz, „Lernstandards“ zu formulieren, und die damit verbundene Einführung eines obligatorischen Systems von Tests ist solch ein internationaler Trend. Merkwürdigerweise hört und liest man bei uns sehr wenig über Kritik an diesen Trends in Ländern, die damit schon viel länger zu tun haben wie die Vereinigten Staaten und Großbritannien. So haben, um bei Letzterem zu bleiben, Wales und Nord-Irland im Jahre 2001 die Veröffentlichung von Testergebnissen in den Zeitungen wegen der negativen Nebenwirkung abgeschafft. Wir haben darüber in den Niederlanden nirgends etwas gehört oder gelesen.

In den Vereinigten Staaten verschlechterte sich die Situation für Schulen, die dem Kind in seiner Ganzheit gerecht werden wollen. Das Zitat aus einer E-Mail der in den Kreisen heutiger „progressiver Pädagogen“ in den USA bekannten Pädagogin Patricia Carini lässt etwas davon wahrnehmen:

‘..... an era of regressiveness in education across the country on a scale never experienced in my lifetime. How hard it is for teachers to make room for children, their play and learning, when even 5 year olds are denied recess and testing has assumed maniacal intensity. So, of course, this is the time when what Prospect (das Netzwerk, in dem Carini aktiv ist, KB) can offer is most needed.’

Sie macht auch klar, was ihrer Meinung nach in dieser Situation zu tun ist:

‘What I know is that we who have other visions must both persist and resist -- and continually commit and recommit ourselves to the hard, recursive work of keeping the door ajar (= die Tür einen Spalt offen lassen) so these visions of a humane education do not slip altogether from view and from practice’ (1)

Der nachfolgende Text (2) gibt einige Fragmente aus der Diskussion und den Untersuchungen im Ausland wieder. Er stellt lediglich einen Einblick dar um zu zeigen, wie kritisch über hier diskutierte und manchmal schon erdrückende Entwicklungen nachgedacht wird. Die Auswahl der Quellen ist einseitig – in dem Sinne, dass es um kritische Stimmen geht, die in den Niederlanden (und wahrscheinlich auch in Deutschland) sehr selten gehört werden. Es geht dabei nicht um unbekannte irgendwie sektiererische Leute, oder um marginale Zeitschriften. Menschen wie Elliott Eisner und Deborah Meier, Linda Darling-

Hammond und Selma Wassermann, Klaus-Jürgen Tillmann und John Elliott, William Spady und Patricia Carini sind bekannte Pädagogen, und es geht hier um renommierte Zeitschriften und Verlage. Die erwähnten Publikationen, die hier verwandt werden, greifen ihrerseits wieder auf Untersuchungen zurück – Alfie Kohn zum Beispiel dokumentiert seine Darstellungen sehr sorgfältig mit Hinweisen auf die Literatur. Die Gedrängtheit des vorliegenden Artikels führte dazu, dass in starkem Maße selektiert werden musste, und eine solche Auswahl ist immer subjektiv. Aber wer es will, kann natürlich jederzeit die genannten Quellen selbst befragen.

Um keine Missverständnisse entstehen zu lassen: Es geht in diesem Artikel nicht darum, Tests als solche zu kritisieren, das wäre Unsinn.(3) Dasselbe gilt auch für Lernstandards als solche. Linda Darling Hammond (4) weist darauf hin, dass Qualitätsstandards auch innerhalb einer professionellen Kultur, z.B. von Ärzten, Architekten, Psychologen, Pädagogen u.a. gelten. Sie beinhalten Kriterien für ‚gute Arbeit‘ – ‚standards of practice‘, ein ‚Berufsethos‘. Solche Standards sind Grundlage für Ausbildung, Diplome, usw. Dieses ‚Professionalisierungsmotiv‘ soll scharf unterschieden werden vom ‚Beherrschungsmotiv‘, das in diesem Artikel kritisiert wird.

Wer denkt, dass es in den Niederlanden und Deutschland nicht so schlimm kommen kann wie anderswo, ist ganz schön naiv. Muss muss noch nicht von Komplottheorien ausgehen oder von großer Angst vor ‚big brother is watching you‘ ergriffen sein, um zu sehen, dass hier starke Kräfte am Werk und große (auch kommerzielle) Interessen im Spiel sind. Solche Systeme haben – einmal entwickelt – eine inhärente Tendenz, raffinierter zu werden und alles und jeden einzuschließen („Systemzwang). Bei uns wird sehr wohl von „der englischen Krankheit“ gesprochen, wenn es um das System der Evaluation in England geht.

Mit Blick auf Skeptiker lenken wir zunächst unsere Aufmerksamkeit auf die Wahrnehmung „holländischer Zustände“ von Seiten eines deutschen Pädagogen. Nach dieser Stimme aus Deutschland werden -in einer Liste - kurz negative Auswirkungen zentralen Testens beschrieben. Es folgen dann einige ausführlicher thematische Interpretationen solcher Auswirkungen, wobei wegen der verschiedenen thematischen Zugänge die eine oder andere Wiederholung nicht zu vermeiden ist; diese ist aber funktional. Zum Schluss wird die Frage besprochen, was man tun kann, wenn man diese Entwicklungen kritisiert.

“Holländische Zustände”

Klaus-Jürgen Tillmann, Pädagogikprofessor in Bielefeld, beschreibt (5) die Tendenz zu internationalen vergleichenden Studien über Schülerleistungen auf Gebieten wie Mathematik, Naturwissenschaften, Lesen und andere Bereiche von Sprachunterricht, fremde Sprachen. Die Ergebnisse dieser Studien führen oft zu großer Aufregung und – wenn ein Land schlecht abschneidet – zu mitunter

eingreifenden politischen Maßnahmen, um eine Verbesserung zu erreichen, siehe die Diskussion in Deutschland über TIMMS und PISA. Solche vergleichenden Studien können – so ist die gängige Argumentation – zur Qualitätssicherung in den Schulen beitragen. Aber dann muss auf jeden Fall eine Verbindung zu den Evaluationsdaten von Konfessionsschulen und selbst von Klassen innerhalb dieser Schulen hergestellt werden. Aber wollen wir das wirklich? Welches sind die (unbeabsichtigten) Nebeneffekte dieses Verfahrens? Welches sind die Fallgruben? Die Entwicklung in den Niederlanden verläuft deutlich in dieser Richtung und soll deshalb einer näheren Betrachtung unterzogen werden. „Die Niederlande sind in Deutschland vor allem wegen ihres hohen Anteils an ‚freien Schulen‘ bekannt und gelten dann auch als die ‚schöne Welt‘ der Schulautonomie. Andere Tendenzen sind viel weniger bekannt, namentlich die Formen zentraler Tests in Verbindung mit der Schulaufsicht und deren Folgen für die Verschärfung der Qualitätskontrolle.“ Tillmann beschreibt dann drei Bestandteile dieses Systems (im Anschluss an eine Publikation der niederländischen Schulaufsicht):

1. Zentrale Tests und ‚ranking‘: der (noch nicht obligatorische) CITO-Grundschulabschlusstest, der auch als Instrument für die Bestimmung der Schulqualität verwendet wird, die obligatorischen zentralen Tests am Ende der ersten und der zweiten Phase der Sekundarstufe, die für die abschließende Beurteilung jedes Fachs zur Hälfte beitragen (die andere Hälfte wird beigetragen von den schulspezifischen Lehrerbewertungen), - mit „Ergebnissen“ in der Zeitung.
2. Die Schulinspektion hat die Aufgabe, jede Schule im Hinblick auf die wichtigsten Qualitätsdimensionen (Lernangebote, Lernergebnisse, Schulklima, usw.) zu beurteilen, und sie fordert, den „Ertrag“ des Unterrichts mittels anerkannter externer Leistungsevaluationen vorzulegen. Die Daten dieser Inspektion sind öffentlich. Die Ergebnisse können Anlass für „Verbesserungspläne“ sein.
3. Zielgerichtete Arbeit an Schulverbesserung, u.a. auf Grund von Empfehlungen der Schulinspektion. „Wenn die Schulen drei Jahre hintereinander ‚Ergebnisse‘ aufweisen, die mehr als eine „halbe Standardabweichung“ unter dem Klassendurchschnitt (ihrer Jahresgruppe) liegen, dann erfolgt in allen Fällen eine vertiefte, ausführliche Inspektion. Dies geschieht mit der Absicht, die Probleme besser zu analysieren und mit der Schule möglichst präzise Vereinbarungen für die weitere Arbeit zu treffen“.

„Dieses Modell wird von seinen niederländischen Befürwortern als gelungene Verknüpfung von objektivierter Leistungsmessung, schulischer Qualitätssicherung und öffentlicher Rechenschaftslegung [gegenüber Behörden, Eltern u.a., KB] angesehen. Aus meiner – einer bundesdeutschen – Perspektive wirkt es hingegen eher wie ein technokratisches Kontrollmodell, dessen Folgen und Nebenfolgen mit äußerster Schärfe in den Blick zu nehmen sind.“(6)

Hier eine kurze Zusammenfassung der Kritik Tillmanns:

1. Die Verkoppelung regelmäßiger zentraler Tests mit der Evaluation individueller Leistungsmessung und Zensierung zwingt die Lehrer dazu, ihren *Unterricht auf den Test hin auszurichten und wahrscheinlich sogar mit spezifischen Drills und Übungen* für den Test zu arbeiten. Es ist eine moderne Variante des „Nürnberger Trichters“ – das „Übertragen“ und „Einfüllen“ von Kenntnissen und Fähigkeiten.
2. Die Produzenten von Tests bekommen eine unglaublich starke Machtposition.
3. ‚Ranking‘ von Schulen – in ihrer großen Unterschiedlichkeit – anhand einer einzelnen statistischen Kennziffer – führt über die Eltern dahin, dass sich soziale Unterschiede verschärfen. „Gute Schulen“ bekommen Kinder mit mehr Chancen, „schlechte“ werden weiter „zurückbleiben“.

Sein Ratschlag: Wir brauchen eine kräftige Diskussion über die Frage, ob wir das (in Deutschland) wirklich wollen (7).

Negative Auswirkungen kurz beschrieben

Hier werden einige Beispiele von Beschreibungen negativer Effekte zentralen Testens gegeben. (8) Ab und zu nehmen wir uns die Freiheit, dem etwas hinzuzufügen.

1. Lehrer werden auf Dauer den Testergebnissen mehr vertrauen als ihrem eigenen professionellen Urteil über Kinder, die auf längeren Beobachtungen und Interaktionen beruhen. Letzteres garantiert noch kein qualitativ angemessenes Urteil, aber ersteres noch viel weniger. Es erinnert an die Kritik des Utrechter Schulforschers Karel Stokking seinerzeit an einem Begleitsystem für die Schülerentwicklung (9), dass nämlich Lehrer dadurch „entprofessionalisiert“ werden. In diesem Zusammenhang hören wir auch regelmäßig das Argument, dass stärkerer Druck von außerhalb der Schule – mit externen Evaluationsmethoden – nötig sei, weil die Qualität der Lehrer bei weitem nicht den Anforderungen entsprächen. Wenn man diesen Weg einschlägt, wird das allerdings eine ‘selffulfilling prophecy’ (KB) – siehe nächster Punkt!
2. Gute – kreative – Lehrer und Direktoren werden demotiviert und verlassen die Schule. In The Times Educational Supplement standen im Jahr 2000 einige Interviews mit Lehrern, die den Schuldienst quittiert hatten und mittlerweile im Betriebsleben arbeiteten. Tenor: „Wir merkten ständig, dass die man die Schule als Betrieb ansah und das Lernen der Schüler als Produkt. Diese Situation stieß uns immer mehr ab; zu diesem Zweck sind wir nicht in den Schuldienst eingetreten. Wir arbeiten da lieber gleich in echten Betrieben.“
3. Lehrer denken immer weniger über ihre Arbeit nach, und sie lernen immer weniger durch Reflektion über die Arbeit und von ihren Kollegen. Stattdessen schirmen sie sich ab und gehen in eine Verteidigungsstellung,

und so werden sie zugleich faktisch Konkurrenten (10). John Elliott, Professor an der University of East Anglia (Norwich, England) fängt seine Beschreibung dieses psychologischen Prozesses (11) an mit der Bemerkung, dass der Arbeit von Lehrern vor 15 Jahre vielleicht zu viel unkritisches Vertrauen entgegengebracht wurde. Jetzt aber ist das Umgekehrte der Fall. Unterrichten ohne Vertrauen ist aber unmöglich. Die externe Kontrolle durch die Inspektion ist seines Erachtens zu stark geworden. Im Zuge der zunehmenden ‚Transparenz‘ der Lehrerarbeit vor dem „Publikum“ verschwindet das Vertrauen in und von Lehrern.(12) Letztere fühlen das und entwickeln Misstrauen gegenüber der Integrität der Inspektion. Daraus wird schnell eine pathologische Angst vor der Inspektion, und die Verzweiflung, die entsteht, ist kontraproduktiv in Hinblick auf eine langfristige Verbesserung der Schulen. Lehrer verlieren ihre Motivation und das Interesse an ihrer Arbeit, vermeiden Risiken (‚playing safe‘), es gibt keine Entwicklung, sondern Erstarrung. In unserer instabilen, dynamischen Gesellschaft sind Experimentierlust und Innovation aber unverzichtbar.

4. Der große Druck und die ernsten Konsequenzen auf Grund geringer Testergebnisse rufen offensichtlich Betrug hervor.
5. Lehrer entwickeln in vielen Fällen eine negative Haltung gegenüber Kindern, die niedrige Testergebnisse produzieren, auch weil sie später dafür die „Rechnung“ präsentiert bekommen.
6. Das Gespräch in der Schule über die Schüler und den Unterricht wird auf gravierende Weise eingeengt: Es geht nicht mehr darum, was man als Schule für wichtig erachtet und was Kinder (oft überraschend, unerwartet) entdecken, sondern immer mehr um Tests und wie man in denen am besten abschneiden kann.
7. An amerikanischen Schulen hat das System der zentral durchgeführten Tests zu radikalen Veränderungen in der Didaktik geführt - mit dem Ergebnis größerer Einförmigkeit, - man konzentriert sich nun vor allem auf das Üben für den Test. In den Wochen vor dem in regelmäßigen Abständen erfolgenden Test werden „nicht relevante“ Fächer vernachlässigt und „ineffiziente“ Aktivitäten, wie Exkursionen, kleinere Experimente, Gespräche darüber, was einen beschäftigt, Vorlesen, Lesen zum eigenen Vergnügen usw., geopfert. Anders gesagt: Die übliche technokratische Weise des Umgehens mit Lernstandards und ihrer Evaluation führt zur Standardisierung von Inhalten und didaktischen Arbeitsformen. (13) Ein Beispiel für Letzteres ist in England die Einführung der obligatorischen ‚literacy-hour‘ mit einem festen Unterrichtsschema (eine Wiedergeburt des Herbartianismus?, KB).
8. Schüler werden durch das System normierter Tests (nach Ziffern) sortiert, statt dass festgestellt wird, was sie können. Dies verstärkt eine pädagogisch unerwünschte Rivalität.

9. Schüler entwickeln eine festgelegte und sehr einseitige, um nicht zu sagen *falsche Vorstellung von „intelligentem Verhalten“*, welches nämlich das Erzielen guter Testergebnisse sein soll. Höhere Testergebnisse lassen aber eher auf die Gewöhnung an Testsituationen, als auf bessere Lernleistungen schließen. Gute Testergebnisse geben darum ganz häufig kein gültiges Bild darüber ab, was Schüler können, sie stellen eher den Reflex von Kindern dar, die für das Bestehen dieser Art von Tests gedrillt worden sind.
10. Linda Darling-Hammond (14) drückt es mit folgendem Bild aus: Stell dir vor, dass beschlossen worden ist, dass Krankenhäuser bessere Leistungen vorweisen und die Normen also erhöht werden müssen. Allen Patienten muss dann auf in einem standardisierten Verfahren – d.h. ungefähr zum selben Zeitpunkt – die Temperatur gemessen werden. Kurz zuvor teilt das Pflegepersonal hohe Dosen Aspirin und kalte Getränke aus. Die Ergebnisse sind in der Tat beachtlich – niemand hat Fieber! Die Qualität dieses Krankenhauses ist hervorragend. Natürlich hat die Realität hinter diesem hohen Abschneiden wenig mit guter medizinischer Versorgung und ihrer genauen Evaluation zu tun, so wie sich auch das Unterrichten für den Test voll und ganz von gutem Unterricht und seiner genauen Evaluation unterscheidet.
11. Alfie Kohn zitiert einen Fachmann für solche Schultests, der vorschlägt, die ‚league-tables‘ mit den durch die Tests ermittelten „Schuldurchschnitten“ in der Zeitung – wenn man sie doch unbedingt veröffentlichen will – dort zu plazieren, wo sie hingehören: auf die Sportseiten.
12. Deborah Meier (15) unterstreicht in ihrer Kritik die zerstörerischen Folgen für das Engagement von Lehrern, Eltern und Schülern sowie andere Menschen und Einrichtungen in der Umgebung der Schule.

Demokratische Kontrolle?

Nach Meier wird im Namen einer „demokratischen Kontrolle“ in Wirklichkeit Demokratie in und um die Schule herum untergraben. Es beginnt damit, dass durch Experten für das ganze Schulsystem eines Landes festgestellt wird, ob es im Verhältnis zu den Investitionen und im Vergleich mit anderen hochentwickelten Ländern genug „einbringt“. Dafür braucht man dann die landesweit gültigen Normen, die nach Schultyp spezifiziert werden, und darüber hinaus nach verschiedenen Entwicklungsstufen der Kinder/ jungen Leute. Dies wird weiter in standardisierten Tests umgesetzt, mit Hilfe derer man Individuen und Schulen ‚objektiv‘ vergleichen kann, - regional, landesweit und international. Außerdem kommt ein System von Sanktionen für versagende Schulen hinzu, und diese müssen auf „unparteiische“ Weise festgestellt werden, also durch Institutionen, die einen gewissen Abstand zu den Schulen haben. Man glaubt dadurch auch mehr Gerechtigkeit erreichen zu können – gleiche Chancen für alle Schüler auf ein

Minimum an gutem Unterricht. Schließlich werden Hinweise für einen „effektiven“ Unterricht gegeben, die „Verbesserungen“ zu Stande bringen können. Die Einzelschritte in einem solchen System, gegebenenfalls auch die Denkweise ergeben sich in einer logischen Folge. Dem Ganzen liegen Annahmen über Demokratie, Schulentwicklung, Qualität und Qualitätssicherung etc. zu Grunde.

Meier plädiert für „*alternative*“ Annahmen:

- Pluralität von Schulen als Reichtum – mehr Definitionen dessen, was eine „gute Schule“ ist. Wo eine klare Haupttendenz im Schulwesen zu erkennen ist, müssen „alternative“ Visionen auch eine Chance innerhalb großzügig definierter allgemeiner Bedingungen bekommen. Dies erzeugt in einem Schulsystem eine gesunde Spannung.
- Die lokale Kontrolle darf nicht untergraben werden. Neben den Urteilen von Experten von außerhalb (der örtlichen Situation) ist demokratisches Verfügungsrecht (*ownership*) der Gemeinschaft der direkt Betroffenen – Lehrer, Eltern, Schüler und „Abnehmer“ der Schule (Nachfolgeschulen, Arbeitgeber u.a.) – von wesentlicher Bedeutung, auch hinsichtlich der Ziele, Inhalte und Arbeitsweisen. Darin ist es auch ein wichtiges Vorbild demokratischen Lebens für die Schüler.
- Bei der Evaluation müssen verschiedene Arten von Quellen benützt werden. Dabei sollte die Betonung auf Standards liegen, die von der Schule selbst zu formulieren und für alle Betroffenen einsichtig (und so auch „stabil“) sind. Diese sollten auch auf lokaler Ebene getestet werden, zunächst durch systematische interne Evaluation. Dabei soll die Betonung darauf liegen zu zeigen, was eine(r) wirklich kann - „Performance-Evaluation“, auch „authentische Evaluation“ genannt. (16) Ergänzt werden sollte dieses Verfahren dadurch, dass eine unabhängige Evaluationsgruppe aus der Region in regelmäßigen Abständen - auf Einladung der Schule hin - eine externe Überprüfung organisiert. Evaluations-Akte vonseiten der Behörden (des Bundesstaats in den USA) haben in solch einer Situation den Charakter einer ergänzenden „second opinion“ im Rahmen von flexibel zu handhabenden nationalen Normen.

Gut für sozial benachteiligte Kinder ?

Meier schreibt darüber hinaus auch über „gesellschaftliche Gleichheit“ bzw. soziale Gerechtigkeit. Sie erhält in dem hier zitierten Buch gleichwohl eine Gegenrede von jemandem, der sie fragt, ob sie vielleicht vergessen hat, dass das Formulieren von Lernstandards gerade darauf zielt, sozial-ökonomisch bedingte Unterschiede in Angriff zu nehmen. Gerade arme Familien sollen von höheren Anforderungen an den „output“ der Schule profitieren. Der Unterricht für

Kinder aus benachteiligten Familien ist heute schlecht bis sehr schlecht. Dies wird durch niedrige Erwartungen hinsichtlich dieser Kinder mit verursacht, infolge derer sie nicht ausreichend stimuliert werden. Die Reaktion auf Meiers Argumente erfolgt nach dem Motto: „Es ist doch so gut gemeint, - wie kannst dann ausgerechnet du, Debbie Meier, dagegen sein?“

Meier glaubt aber nicht, dass Standardisierung und zentrale bürokratische Kontrolle von außerhalb der lokalen Gemeinschaft hierfür das geeignete Mittel ist. „Das letzte, was wir tun müssen, wäre ausgerechnet eine Verstärkung der Zentralisierung und Standardisierung, die ja immer schon in den Schulen für die Armen vorherrschend waren. Diese Verstärkung ist gerade ein Teil des Problems.“ Sie will das, was in den USA das Vorrecht wohlhabender Privatschulen ist, in „normalen“ (staatlichen) Schulen in benachteiligten Stadtteilen einführen: mehr Entscheidungsfreiheit, um diesen Kindern und ihren Gemeinschaften zu dienen. Sie weist auch auf den Erfolg hin, den sie mit den von ihr aufgebauten und geleiteten Schulen in East Harlem (New York) erzielte. Das sind die in den USA und darüber hinaus berühmten „Central Park East Elementary and Secondary Schools“ (17). Diese und die aus ihnen hervorgegangenen Schulen in New York haben folgende besondere Merkmale: starkes Elternengagement, Teamteaching von Lehrern, Teilhabe von Schüler an der Gestaltung des Schullebens und des Unterrichts, „Kleinheit“ und gute Gemeinschaft (Schulen dürfen nicht zu groß werden, - oder: Es müssen „Schulen in der Schule“ gebildet werden), hohe intellektuelle und andere Standards, die für jeden einsichtig sind und über die auch in der Öffentlichkeit Rechenschaft abgelegt wird (s.o.). Eine technokratische Veränderung von oben droht, Experimente wie die in New York/East Harlem oder in Boston (an der Schule, die Meier mittlerweile aufbaut) unmöglich zu machen und zu zerstören. In den vorangehenden Abschnitten habe ich bereits einige Anmerkungen über die negativen Auswirkungen (zentraler Standards und Tests) besonders für sozial-ökonomisch benachteiligte Kinder gemacht. Das hängt auch mit den meist verwendeten landesweit eingesetzten standardisierten Tests zusammen. Es geht hier ja um ein relatives Messsystem – die Punktzahl gibt an, welche Stelle (Zuordnung nach dem Abschneiden in Prozent) ein Kind in der Gesamtheit der Population einnimmt, in der dieser Test durchgeführt wird. Wenn die Kinder im Rechentest viel besser abschneiden, dann werden trotzdem „schlechte Rechner“ übrig bleiben, nämlich die Kinder des niedrigsten Sektors (25 %), so sehr man auch als Schüler, Lehrer oder Schule sein Bestes tut. Es ist immer ein System von Gewinnern und Verlierern. Wer gut abschneidet (auch eine Schule), drückt andere nach unten.

Die negativen Auswirkungen standardisierter Tests, besonders für Kinder aus armen Familien, werden in dem Buch von Linda McNeil sehr detailliert beschrieben.(18) Seine Schlussfolgerung lautet: „*Standardisieren reduziert die Qualität und Quantität dessen, was in Schulen unterrichtet und gelernt wird.* Dieser unmittelbare negative Effekt der Standardisierung ist die vorherrschende Folgerung aus einer wissenschaftlichen Studie von Schulen, in denen der Zwang

standardisierter Kontrollen den Umfang und die Qualität der Unterrichtsinhalte reduzierte, die Rolle der Lehrern schwächte und aktives Lernen der Schüler ziemlich unmöglich machte.“

Die Langzeit-Wirkungen sind noch schädlicher: *„Auf lange Sicht schafft die Standardisierung ungerechte Ungleichheit und verbreitert die Kluft zwischen der Unterrichtsqualität für Jugendliche aus armen Familien und Minderheiten einerseits und die von eher privilegierten Schüler andererseits... ...je länger die standardisierten Kontrollen wirken, desto breiter wird diese Kluft, wenn nämlich das System der Tests und die Vorbereitung auf den Test in Schulen für Kinder aus Minderheiten die Stelle des Curriculums besetzt, das privilegierte Schüler angeboten bekommen. Diese neuen Diskriminierungsstrukturen werden durch Kontrollen hervorgebracht, die in den Schulen begannen, die in dieser Forschung unter die Lupe genommen wurden. Diese Strukturen wurden im Laufe der Jahre das vorherrschende Modell des Schulsystems in einem der größten und qua Bevölkerungszusammensetzung am deutlichsten multiethnischen Bundesstaaten der USA: Texas... ... Sie (die Studie) dokumentiert die wachsende Macht dieser externen Kontrollen, der Ausbildung aller Kinder den Garaus zu machen, insbesondere von Schüler(inne)n mit afro-amerikanischem oder latino-Hintergrund.“*

Die Forschungsarbeit McNeils richtete sich anfänglich auf eine ganz andere Fragestellung. Sie war nämlich als Suche nach Organisationsmodellen gedacht, die auf strukturelle Unterstützung authentischen und engagierten Lernens und Unterrichtens zielten. Solche Arbeitsweisen – in „Magnetschulen“, die vor allem Schülern aus armen Familien, hauptsächlich von Afro-Amerikanern und Latinos, zu Dienste standen, - sollten genau dokumentiert werden. Dies war zumindest die Absicht. Aber während dieser Untersuchung wurden die betreffenden Schulen einer Veränderung unterworfen des Inhalts, dass sie dem „Regime“ eines zentralisierten Systems vorgeschriebener Regeln und standardisierter Prozeduren von Qualitätskontrolle („monitoring“) unterworfen wurden. Und die spezifischen Qualitäten auch und gerade dieser vielversprechenden Schulen konnte man vor seinen eigenen Augen verschwinden sehen. Die Entdeckungen, die solchermaßen in diesem Forschungsprojekt gemacht wurden, ohne dass sie anfänglich intendiert waren, sind auf diese Weise um so überzeugender.

Am Ende des Buches schreibt sie:

„Wir als Menschen im Allgemeinen und als Schulpädagogen im Besonderen haben niemals ein so großes Wissen über die Entwicklung von Kindern besessen wie heute. Wir haben niemals so viel gewusst über den Prozess des Erkennens und das Funktionieren des Gehirns sowie die Wirkungen von Achtsamkeit, Fürsorge und Gemeinschaft auf das Wachstum und die Entwicklung von Kindern. Noch niemals haben wir so viel wie heute wissen und genießen können: unser reiches kulturelles Erbes in den Künsten, der Literatur, Geschichte, die Naturwissenschaften, Mathematik, Astronomie, Sport, die

Sprachen dieser Welt, die Geografie von fernen Orten, das Funktionieren unserer Körperzellen. Niemals zuvor haben wir unseren Kindern so viel bieten können. Wie ist es dann möglich, dass sich Systeme entwickeln, die Lernmöglichkeiten für Kinder bedrohlich einengen, die alte Formen von Diskriminierung und Benachteiligung fortsetzen und neue Formen entwickeln, unsere ärmsten Kinder gegen jenes reiche kulturelle Erbe abzuschirmen? Wie können wir Strukturen als „normal“ akzeptieren, die den Einfluss von Eltern und der übrigen Gemeinschaft um die Schule herum auf die Unterrichtsziele hin minimalisieren? Wir müssen gerade die Arten zu unterrichten und Schulen zu organisieren dankbar begrüßen, welche Lehrkräfte stimulieren, das Beste, was sie wissen und können, in der Klasse einzusetzen. Wir müssen Curricula und Lernumgebungen schaffen, die alle Lernfähigkeiten von Kindern anerkennen und stärken, wozu auch die Anerkennung der Unterschiedlichkeit der Kinder als inhärente Stärke unserer pädagogischen Bemühungen gehört. Und wir müssen die wissenschaftliche Expertise voranbringen, die - mit den Worten von Maxine Greene – ‚ein mehrgestaltiges und mehrperspektivisches Lesen der Welt fördert‘ und die dauernd danach trachtet, nicht nur Unterrichtsprozesse zu interpretieren, sondern letztendlich auch bereit ist, die Kosten zu tragen.“

Den Unterricht beherrschen wollen und die Grenzen dieses Unterfangens

Elliott Eisner, Pädagogikprofessor an der Stanford-Universität, schürft tiefer und sucht nach der Grundhaltung, die die Basis des ganzen Systems der Lernstandards und deren Kontrolle bildet.(19) Er bezeichnet diese Grundhaltung als den Willen, unsere Umgebung einschließlich der Entwicklung von Menschen und der Prozesse zwischen Menschen rational zu beherrschen. Es herrscht die Auffassung, dass “kennen”so viel heißt wie “voraussagen können” , und “voraussagen” so viel wie “beherrschen können”. So will man auch den ‘Ertrag’ der Unterrichtsprozesse vorhersagen können sowie den Beitrag, den Unterricht zur Entwicklung von Kindern beisteuert. Man formuliert so genau wie möglich die beabsichtigten Ergebnisse, und zwar so weitgehend wie möglich in messbarer Form – kurz- und längerfristig (von Unterrichtsstunden, Lehrjahren, einzelnen Jahren, dem ganzen Grundschulunterricht), stellt die Anfangssituation von Kindern fest und kontrolliert hinterher, ob die ins Auge gefassten Ergebnisse erreicht worden sind. Gegebenenfalls stellt man auch fest, was der Unterricht zur Entwicklung der Kinder beigetragen hat: den “zugefügten Wert” oder “Mehrwert”. “Qualität” muss in Quantität ausgedrückt werden – wegen der größeren “Objektivität”, Messen heißt Wissen. Das Streben geht in Richtung möglichst universell geltender, “wissenschaftlich festgestellter” Unterrichtsmethoden (“what works”), die die beabsichtigten Ergebnisse garantieren würden und die dann auch allgemein eingeführt werden müssen. Als “wissenschaftlich” legitimiert gilt in diesem Zusammenhang allein die “harte” empirisch-analytische Vorgehensweise, bei der nur messbare Effekte zählen.(20) Und so scheint sich der Kreis zu schließen:

a. Man geht davon aus, dass Wissen, und Voraussagen-Können Beherrschen heißt; darüber hinaus kann man im Rahmen dieser Vorgehensweise nur durch

Messen sichere Erkenntnisse gewinnen (Quantifizierung). Auf Grund dieser beiden Annahmen wird der beabsichtigte “Gewinn” so genau wie möglich formuliert;

b. Mit Hilfe einer Methode, die die vorausgesagten Ergebnisse so zuverlässig wie möglich garantiert, wird auf diese Ziele hin gearbeitet;

c. Der “Gewinn” wird gemessen (Diagnose);

d. Für die Kinder, die ungenügend abgeschnitten haben, wird die Methodik eventuell zugespitzt.

Diese Philosophie der Beherrschung führt leicht zu einer Reihe von Reduktionen:

- Ziele, die schwer zu messen oder messbar zu machen sind, geraten außerhalb des Gesichtsfeldes. Es ist dann zum Beispiel eine unzureichende Menge Geldes vorhanden, um für komplexe Fähigkeiten die passenden Evaluationsinstrumente zu entwickeln oder man findet die anvisierten Ziele schon bald “zu vage”.
- Die Reduktion der Unterrichtsmethoden zu dem, was einfach messbaren Zielen dient.
- Die Betonung “mundgerechter”, “überschaubarer” Module, mit kurzen Lernzyklen auf Kosten von länger sich hinziehenden Lernprozessen, die per definitionem nicht zu modularisieren sind.

John Elliott analysiert die Diskussion, die um die “effektive Schule” geführt wird, einschließlich der Forschung auf diesem Feld.(21) Seine Befunde bestätigen und verstärken noch die von Eisner. Die Befürworter der “effektiven Schule” verweisen bei kritischen Nachfragen gerne auf “Forschung”, die ihnen Recht geben würde (“what works”). Selten kommt zur Sprache, dass diese Daten im Rahmen einer ganz bestimmten Art über Unterricht zu denken zustande gekommen sind und die Elliott als “social engineering” typisiert – das Entwickeln eines Kontrollsystems über psychologische Techniken, Zwang, Lohn und Strafe.

Dieses System hat dieselbe Struktur wie russische Puppen, - ein Subsystem, das von einem höheren Subsystem kontrolliert wird und selbst wieder die Kontrolle ausübt über Subsysteme innerhalb des eigenen Bereichs. “Qualität” wird primär definiert als technische Effizienz; aber dahinter verbirgt sich eine Vision von Unterricht, in der Leitbilder, wie soziale Kontrolle und Anpassung eine wichtige Rolle spielen. Behörden und Schulleitungen mit einer bürokratischen Einstellung finden hier ein ideales Instrumentarium, um Richtlinien einzuführen und zu kontrollieren – in guter Absicht, wie wir wohl annehmen dürfen. Für diesen Typ von Schulprojekten ist es dann auch ziemlich einfach, Geld zu bekommen. Lehrkräfte werden in diesem Gesamtzusammenhang vor allem als Techniker gesehen, die die Anweisungen so exakt wie möglich ausführen und nicht als “reflektierende Praktiker” (oder, wie Eisner es sagen würde, als Künstler – ‘teaching as an art’). Elliott weist darauf hin, dass eine “effektive Schule” mit einer “guten Schule” per definitionem nicht identisch ist. Es hängt

aber davon ab, was man unter Letzterer versteht. Die „auf Effektivität ausgerichtete Schulbewegung“ gibt sich als „research-based“ aus, aber repräsentiert in erster Linie eine bestimmte Sichtweise von Schule. In einem anderen Schulkontext, der aus einer anderen Vision heraus gestaltet ist, sind durch Forschung andere „Effekte“ von Unterricht zu entdecken, die im Rahmen „effektiver Schulen“ nicht oder kaum zum Vorschein kommen.(22)

Einer der Nebeneffekte des oben beschriebenen Trends ist nach Eisner auch das Phänomen des berechnenden Schülers. Schüler bekommen ständig die „Message“ zu hören, dass das Erzielen hoher Testergebnisse das Wichtigste sei. Sie finden die verschiedensten „Tricks“ heraus, um mit möglichst wenig Einsatz die höchst möglichen Testergebnisse zu erzielen. Nicht der Inhalt, in den man sich vertiefen oder in dem man sich verlieren könnte, ist von Bedeutung, sondern der Test. „Ballast“ muss dementsprechend über Bord geworfen werden, - mit dem Ergebnis der Einschränkung von Zielen und der Verkümmern von Inhalten. Das Merkwürdige ist, dass der Unterricht immer weniger auf ein gutes Zurechtkommen im normalen Leben außerhalb eines schulischen Kontextes ausgerichtet zu sein scheint (und um das Letztere geht es doch?), sondern immer mehr auf das gute Funktionieren an der Schule selbst: gute Testergebnisse erzielen, um später bessere Testergebnisse erzielen zu können. Die Schulatmosphäre richtet sich darüber hinaus immer mehr auf extrinsische Motivation. Eisner behauptet nicht, dass ‚geschlossene‘ Lernformen und ‚direkte Instruktion‘ mit einem messbaren Feedback an sich verwerflich seien, - im Gegenteil. Diese Lernformen können sehr funktional sein, auch als Phase innerhalb des ‚selbstgesteuerten‘ Lernens. Er warnt aber – auf Grund dessen, was er in zahlreichen Schulen wahrnimmt – vor einer weitgehenden inhaltlichen und didaktischen Verarmung des Unterrichts. Gleichzeitig setzen die Unterrichtswissenschaften heute beachtenswerterweise den Akzent stärker darauf, dass Lernen zuallererst heißt (oder heißen sollte), dem zu Lernenden einen persönlichen Sinn, eine Bedeutung „für mich“ zu geben, auch in Interaktion mit anderen. Wo sind das Überraschungsmoment, die Kraft der Phantasie, das Improvisieren geblieben? Bedeutet Unterricht nicht immer auch: Interaktion, mit all der damit gegebenen Unvoraussagbarkeit? (23) Zerstört man mit der starken Betonung auf dem Beherrschen-Wollen nicht wesentliche Dimensionen des Unterrichts?

Man sollte sich mal vorstellen – sagt Eisner – , dass für die kommenden fünf Jahre ein Moratorium hinsichtlich landesweit durchgeführter standardisierter Tests verkündet würde. Man könnte sich dann also bei der Beurteilung von Schulen nicht mehr einfach auf höhere Testergebnisse verlassen. Worauf könnte man (als Eltern, als Schulaufsicht etc.) aber doch achten? Eisner bietet dann eine ganze Liste von Qualitätsindikatoren, von denen wir einige nennen:

- die Arten von Problemen und Aktivitäten, in denen sich Schüler engagieren. Werden sie zum ‚Staunen‘ angeregt, zum Stellen von Fragen zu dem, was sie untersucht haben? Wohlgemerkt: ihre eigenen Fragen und nicht so sehr die Fragen, die wir ihnen stellen?

Die wichtigste intellektuelle Leistung, die man erbringen kann, ist ja nicht so sehr das Lösen, sondern das Formulieren von Problemen, sind die Fragen, die eine(r) stellen kann.

- Werden die Schüler mit mehreren und verschiedenen Perspektiven konfrontiert? Können sie damit umgehen, ohne in Verwirrung zu geraten?
- Kommen sie mit verschiedenen Arten von Wirklichkeitserkenntnis in Berührung, die einander nicht ausschließen, sondern sich gerade gegenseitig ergänzen: Wissenschaft, Kunst, Philosophie u.a.?
- Werden Beziehungen zwischen dem auf der Schule Gelernten und der Welt außerhalb der Schule hergestellt, können die Schüler anwenden, was sie gelernt haben? „Die wirklich wichtigen abhängigen Variablen des Unterrichts sind weder in Klassenzimmern noch in Schulgebäuden lokalisiert, sondern außerhalb der Schule. Unsere Evaluationspraktiken reichen kaum an jene Variablen heran. Der echte Ertrag des Unterrichts ist da zu finden, wo Schüler tun können, was sie selbst tun wollen.“

Am Ende seines Artikels spricht Eisner die Hoffnung aus, dass es tieferes Verstehen dessen geben wird, worum es in der Schule wirklich geht und was da wirklich geschieht. Das soll uns – wie er hofft - eher in die Lage versetzen, den Unterricht zu verbessern.

Die Magie der Zahlen

Auf der Linie Eisners weiterdenkend, kann man sich fragen, warum Menschen so viel Wert legen auf Messen und auf Zahlen. Die Menschheit zählt, misst und berechnet schrecklich viel. Zum Teil liegt das auf der Linie natürlicher Prozesse – so wird innerhalb unseres eigenen Körpers schon eine Menge gezählt, gemessen und gerechnet, und es finden auf Grund der vorgefundenen Werte (z.B. des Hormonspiegels) Rückkopplungen statt. In der Gesellschaft, sicher in unserer modernen Gesellschaft ist das Messen ein weitverbreitetes Phänomen – siehe nur schon die Markierungen auf Verpackungen. Dies ist im höchsten Grade funktionell, macht die Dinge kontrollierbar. Es ist nicht übertrieben zu behaupten, dass sich unsere Gesellschaft „auf Zahlen bewegt“. Wir vertrauen normalerweise darauf, dass die Zahlen auch stimmen. Qualitäten in Zahlen auszudrücken ist übrigens – siehe Eisner – problematisch. Nicht alles lässt sich auf diese Weise „einfangen“. Selma Wassermann, Professorin an der Simon Fraser University in British Columbia (Kanada), stellt ebenfalls die Frage, warum Menschen so viel Wert auf Zahlen legen.(24) Ihr zufolge liegen die Gründe für die Zahlengläubigkeit tiefer, als es in dem Postulat der „intersubjektiven Kontrollierbarkeit“ zum Ausdruck kommt, wie sie in der Wissenschaft an der Tagesordnung ist, und auch tiefer, als es ihrer praktischen Funktion für die Gesellschaft entspricht. Die Wertschätzung der Zahlen hat vor allem eine existentielle Dimension: In einer sich schnell verändernden Welt fühlen sich Menschen unsicher, erleben Stress, müssen aufgrund unzureichender

Sicherheiten zahllose Entscheidungen fällen, deren Folgen sie auch nicht überblicken können. In einer solchen als chaotisch erlebten Wirklichkeit ist dann auf einmal die Zahl da. „Nichts gibt ein sichereres Gefühl oder verleiht mehr Sicherheit als eine Zahl. Es ist dabei unwesentlich, ob die Zahlen oder das Messen präzise sind, wir fühlen uns einfach ruhiger, wenn wir glauben können, dass es etwas gibt, das einen festen Halt bietet, etwas, worauf wir mit Sicherheit rechnen können, etwas woran wir uns klammern können, wenn wir mit großer Geschwindigkeit unsicheren Zeiten entgegen gehen.“

Vielleicht befinden sich die Gründe dafür, dass Eltern und Lehrkräfte so gerne Lernleistungen von Kindern in einer einfachen Zahl ausdrücken wollen, tatsächlich auf dieser existentiellen Ebene. Mit einem Perzentilscore von 151 bei einem Test abzuschneiden „verleiht festen Halt“, so kommt es zumindest vielen vor.

Kürzlich war von einer Schulleiterin zu hören, dass an ihrer Schule alle Jahre wieder viele Eltern an den Oberschulempfehlungen „nörgelten“. „Wir haben damals mit dem CITO-Endtest angefangen und die Nörgeleien waren nahezu vorbei.“ In solchen Fällen wird selten nach der Verlässlichkeit und Validität der Zahl gefragt, - diese sind manchmal schlechthin bedenklich. Wassermann liefert in ihrem Artikel Beispiele von Scheingenauigkeit in Fällen, wo es um Leben oder Tod ging (in der Raumfahrt), von regelrechten Fehlern in der Welt der Naturwissenschaft und Technik, wobei auch Nobelpreisträger Richard Feynman zitiert wird, der vor blindem Zahlenglauben warnt. Wassermann schreibt: „Ich denke dass es besonders notwendig ist, dass Lehrer in ihrer Praxis auf klaren Daten beruhende („informierte“) evaluative Beobachtungen anstellen. Ja, es ist wahr, dass diese unrichtig sein können; ja, es ist wahr, dass sie subjektiv sind; ja es ist wahr, dass die Urteile von Lehrern über die Lernleistungen von Kindern verkehrt, irreführend oder voreingenommen sein können. Aber statt diese Urteile von Lehrern durch die viel weniger brauchbare Alternative der standardisierten Tests zu ersetzen, können wir vielleicht unsere Energie sinnvoller auf die Verbesserung der von den Lehrern durchgeführten Evaluationen richten, indem wir sie praktikabler machen, so dass sie für Schüler, Eltern und für den gesamten Lernprozess unterstützend wirken.“ Im Weiteren gibt sie Hinweise für sehr konkrete „Standards“ für das Lernen und die Entwicklung von Schülern, welche die Lehrer als Bausteine für ihre Klasse verwenden können. Diese „Standards“ beziehen sich auf die Qualität des Denkens, die Kommunikation von Ideen, auf die Fähigkeiten zu untersuchen und zu forschen, zu interpersonalem Austausch, auf die Entwicklung eines gesunden Selbstbildes, auf die Werteentwicklung sowie auf die Fähigkeiten zur Selbstevaluation und die entsprechenden Haltungen. Wassermann sucht Evaluation also in einer ganz anderen Richtung, nämlich in der der ‚authentischen Evaluation‘, die mit Portfolios verbunden ist.(25)

Viele Kinder werden unterfordert

Standardisierte Tests sind stark an die Denk- und Arbeitsweise des klassischen Intelligenzbegriffs angelehnt. Dieses „psychometrische“ Modell hat folgende Voraussetzungen:

- a. Es gibt bei jedem Menschen so etwas wie ein einigermaßen zusammenhängendes und konstatierbares Ganzes an Denkfähigkeiten, das „Intelligenz“ genannt wird;
- b. Menschen verfügen in höherem oder niedrigerem Maße über Intelligenz, und das kann gemessen und in einer einfachen Zahl, dem IQ, zusammengefasst werden;
- c. Dieser IQ ist ziemlich stabil und gibt die Obergrenze der intellektuellen Entwicklung der betreffenden Person an.

Dieses psychometrische Modell wird immer stärker kritisiert (26). Zunächst gibt es nicht so etwas wie „die“ Intelligenz einer Person, sondern Intelligenz wird immer mehr als ein Bündel vielgestaltiger Möglichkeiten angesehen, die bei allen Menschen potenziell vorhanden sind, aber bei verschiedenen Menschen in unterschiedlichem Maße präsent sind bzw. bei verschiedenen Menschen mit unterschiedlichen Akzenten entwickelt werden können. Howard Gardner spricht über „mehrfache Intelligenzmerkmale“ und hat mittlerweile neun davon beschrieben. Neben dem IQ haben wir jetzt beispielsweise auch das EQ als das Maß für die emotionale Intelligenz. Darüber hinaus scheint die Stabilität bzw. Voraussagbarkeit der Lernobergrenze viel geringer zu sein. Siehe u.a. das Werk von Feuerstein, der sich dagegen stellt, von einer im Vorhinein festgelegten Obergrenze von Möglichkeiten auszugehen und zum Beispiel mit Down-Syndrom-Kindern erstaunliche Ergebnisse erreichte.(27) Die heutige Praxis der Tests wirkt sich oft schlecht auf Schüler aus, besonders solche mit sozial-ökonomisch schwachem Hintergrund. Im Rahmen eines Projektes an der Universität Cambridge (England) wurde erforscht, ob neue Hypothesen über Intelligenz und kognitive Entwicklung anders geartete Daten hervorbringen können. Wir folgen der Argumentation, die diesem Projekt zu Grunde liegt (28). In England versucht man im Rahmen der Tests für das nationale Curriculum und der Publikation der Ergebnisse in der Zeitung den Vergleich für Schulen mit vielen „zurückgebliebenen Schülern“ ehrlicher zu machen. Sie schneiden auf Grund ihrer Lebenssituation fast immer niedrig ab, und Schulen mit vielen Schülern dieser Art landen irgendwo unten auf der Liste. Das verantwortungsvolle Korrigieren der rohen Testergebnisse mit Blick auf sozial-ökonomische Faktoren erscheint als unmögliche Aufgabe (29) . Darum werden, um einen gerechteren Vergleich zu bekommen, zu Beginn der Schulzeit in Grund- und Oberschule die kognitiven Fähigkeiten getestet. Dieser Test muss dann die Grundlage für die Voraussage für das letztendlich zu erzielende Niveau am Ende des betreffenden Schulabschnitts (also über x Jahre hin) abgeben – mit der Eingabe und Verarbeitung der zu erwartenden natürlichen Entwicklung des Kindes und dem von der Schule hinzugefügten „Wert“(erbrachten Mehrwert). Es wird davon ausgegangen, dass man das „Endniveau“ voraussagen kann.

Schüler werden auf Grund der Testergebnisse - vor allem im weiterführenden Unterricht, weniger im Unterricht der Primarstufe – nach „ability“ in Gruppen eingeteilt – mit je ergänzender Unterstützung. Kinder werden faktisch auf ein im Voraus bestimmtes Niveau festgelegt. Die Erwartungen von Lehrern hinsichtlich des Könnens der Kinder stellen sich darauf ein, - bewusst oder (öfter) unbewusst. Für den Teil der Kinder, von dem man annehmen kann, dass sie sich unter dem vorausgesagten Niveau befinden, kann das günstig sein, - sie können ausreichend gefordert werden. Aber immer mehr stellt sich heraus, dass ein großer Teil der Kinder in solchen Situationen zu wenig gefordert wird und mehr erreichen könnte. Die Schüler werden also durch diese Herangehensweise faktisch benachteiligt, und zugleich hindert sie in der Praxis auch die Lehrer, das ihnen Mögliche zu tun, die Lernergebnisse zu verbessern. Ein Zitat des Biologen Stephen Jay Gould bildet eine Leitlinie dieses Projektes (30):

„Wir kommen nur einmal auf dieser Welt. Nur wenige Tragödien sind größer, als wenn das Leben verpfuscht wird, wenige Ungerechtigkeiten tiefer, als wenn die Chance, aufzustreben oder auch nur zu hoffen, an einer Grenze zunichte wird, die von außen gesetzt ist, aber fälschlicherweise als im eigenen Innern liegend wahrgenommen wird.“

Von einer *vielgestaltigen* und *nicht-deterministischen* („transformativen“) Auffassung von Intelligenz bzw. von Lernpotential her wird ganz anders gedacht und gearbeitet, wird man den Kindern und ihrem vielgestaltigen Lernpotenzial gerechter. Das Lehrverhalten von 10 Lehrpersonen, von denen erwartet werden konnte, dass sie auf ‚transformative‘ Weise arbeiten, ist ein Jahr lang beobachtet und dokumentiert, von Tiefeninterviews ergänzt worden usw. Und so sind vermutlich maßgebliche Merkmale einer ‚transformativen‘ Didaktik herausgefunden worden, die in ergänzenden Studien weiter erprobt werden sollen.(31)

Wenn sich Einsichten wie die aus diesem Projekt gewonnenen weiter durchsetzen könnten, dann würde sich auch erweisen, dass man mit der heutigen Praxis standardisierter, landesweit normierter Tests nicht mehr viel ausrichten kann, weil deren Grundlagen obsolet sind.

Auf dem Weg zu einer Netzwerkgesellschaft

Wir haben aber in unserem Land und anderswo noch mit einer starken Lobby für die Einführung von Lernstandards zu schaffen, verbunden mit einem technokratischen Kontrollsystem. Aber die Zahl der Menschen wächst, die diesen Typ von Schulpolitik als Nachwehen einer Gesellschaftsform betrachten, die im Verschwinden begriffen ist, - nämlich der von Industrie und Bürokratie dominierten Gesellschaft (32). Metapher für diese Gesellschaft ist das Fließband mit Massenproduktion, einer einheitlichen (einförmigen) Qualitätskontrolle, eine weitgehend durchgeführte Arbeitsteilung und Rationalisierung (Standardisierung), wie auch Eisner sie beschreibt. Auch in der Schule sind diese Merkmale zu finden. Die Schulen können, so dachte man und denkt man

manchmal noch immer, mit diesem Modell genau so effizient Schüler mit spezifizierten Kennzeichen produzieren, wie einst Henry Ford T-Ford-Wägelchen produziert hatte. Siehe die Formen der Unterrichtsplanung, das Lehrstoff-Jahrgangsklassen-System („grading“), das fragmentarisierte Herangehen an den Lehrstoff, die Probleme mit „langsamen Schülern“ in diesem System. Viele Schulen versuchen die Grenzen dieser Herangehensweise zu durchbrechen, auch weil sich die Gesellschaft, einschließlich der Kinder, langsam zu verändern scheint. Auch die im vorigen Abschnitt beschriebene dynamische, nicht deterministische Sicht auf Lernen und Entwicklung passt in ein neues Bild von der Gesellschaft, die als eine *Netzwerkgesellschaft* charakterisiert werden kann. Hierarchische Strukturen werden in zunehmendem Maße zur Diskussion gestellt und die Gesellschaft kann immer weniger von oben gesteuert werden.

Dies kommt mit dem in diesem Artikel beschriebenen Trend überein. Da gilt: Je weiter man von der Schule entfernt wirkt, je höher man in der Hierarchie „angesiedelt“ ist, um so mehr ist man von zentralem Testen begeistert. Und umgekehrt: Je mehr man von Unterricht weiß, das heißt: je näher man der Schulpraxis steht, um so kritischer und nuancierter denkt man darüber.(33)

Lernen in der Netzwerkgesellschaft ist außerdem viel weniger linear, die Kräfte der Phantasie und Kreativität werden stärker beachtet.

Über diese sich entwickelnde Netzwerkgesellschaft lässt sich viel mehr sagen. Eine solche Zukunftsperspektive kann uns helfen, die heutigen Probleme selbstbewusst, mit Humor und Phantasie anzugehen und kreativ gegen falsche Entwicklungen Widerstand zu leisten (34).

Es ist möglich, für diese Ziele Netzwerke von gleichgesinnten Einzelpersonen, Schulen und Organisationen zu bilden. (35)

Was ist zu tun?

Alfie Kohn liefert am Ende seines Buches ‘The Case Against Standardized Testing’ eine Liste von Ratschlägen für Aktionen – für bessere Alternativen und gegen den schädlichen Tests. Der letzte und weitestgehende Rat ist: Boykottiert die Tests! Er gibt dann eine Übersicht über solche Aktionen in den USA, besonders solcher, die von Eltern organisiert worden sind.

Man muss, schreibt er:

- die Unterstützung von Eltern und von anderen Individuen und Gruppen in der Umgebung der Schule suchen;
- politisch unabhängig bleiben, aber zugleich politisch bewusst sein und auf keinen Fall politisch naiv!
- eine klare und relativ einfach zu erklärende Vision und Motivierung haben;
- von der Presse Gebrauch machen.

Konkret soll man danach streben, die verschiedene Funktionen von Tests gut zu unterscheiden. Tests für internationale Vergleiche, wie PISA und TIMMS,

sollen nicht verbunden werden mit Schulvergleichen. Diese Tests können Anlass sein für Maßnahmen zur Schulverbesserung, aber sind gerade nicht darauf orientiert, funktionieren auf einer anderen Ebene. Schulvergleiche haben fast immer negative Auswirkungen und können auf Dauer Verbesserungen des Unterrichts eher verhindern als fördern (s. oben Elliott). Schulen sollen wenigstens genügend Spielraum erkämpfen, um die Qualitätskontrolle ihres Unterrichts in einer Weise zu gestalten, die zu ihrer Konzeption einer 'guten Schule' passt. Und diese 'pädagogischen' Evaluationsverfahren sollen systematisch vorgehen und Verantwortung darüber ablegen, wie die Entwicklung der einzelnen Kinder im Team von den Lehrern systematisch gefördert, besprochen und dokumentiert wird.(36)

Um auf das Zitat von Carini im 'Vorspann' dieses Kapitels zurückzukommen: Was wir tun müssen, ist, gegen den heutigen "Mainstream" Widerstand zu leisten und durchzuhalten sowie weitere Einblicke in "anderen Unterricht" zu ermöglichen, in Schulen, die – um mit Alfie Kohn zu sprechen – unsere Kinder verdienen, auf die sie Anspruch erheben können.(37)

ANMERKUNGEN

- 1) Über die Arbeit von Carini und das Prospect Center: Carini, P.F. (1996), Voortbouwen op de sterke punten van kinderen (... Von den Stärken der Kinder aus), in: Mensen-kinderen, jrg. 11, nr. 3 p. 41-47. Ursprünglich in: Journal of Education, 1986; Carini, P.F. (2001), Starting Strong: A different look at children, schools and standards., New York/ London; Himley, M./ P. Carini (2000), From Another Angle. Children's Strengths and School Standards, New York/ London. Internet: www.prospectcenter.org
- 2) Der vorliegende Aufsatz ist die gestraffte, überarbeitete Version eines Kapitels einer einschlägigen niederländischen Broschüre, die eine landesweite Aktion gegen zentrale Tests mit Fakten und Argumenten unterstützte: Boes, A./ K. Both (2002), Streep door de eindtoets. Toelichting en achtergronden (*Mit dem Endtest soll Schluss sein: Erläuterung und Hintergründe*). Valthermond. S. Dazu auch: Both, K. (2002), Mit dem Endtest soll Schluss sein. In: Frankfurter Rundschau, 31. Januar 2002
- 3) -Boes, A. (2003), Tests – Sinn und Unsinn. In: Päd Forum: Unterrichten erziehen. Jahrg. 33/22, Heft 2, s. 94-97
- 4) Darling-Hammond, L. (1997), The Right to Learn. A Blueprint for Creating Schools that Work, San Francisco
- 5) -Tillmann, K.J. (2001), Leistungsvergleichsstudien und Qualitätssicherung, oder: Auf dem Weg zu holländischen Verhältnissen?, in: Neue Sammlung, Vol. 41, Heft 3, S. 359-368
- 6) Tillmann, 2001, S. 366. Siehe dazu auch die Kritik von Roger Standaert der in der heutigen Diskussion über Schulqualität und Evaluation– wenn man vereinfacht - zwei konträre Ansätze unterscheidet: eine ökonomisch-technische Orientierung und eine pädagogisch-didaktische Orientierung. Die erstere ist im internationalen Diskurs dominant: Standaert, R. (2001), The temptation of simplification, A plea for integration of internal and external quality control in: H. Bachmann, e.a. Turning the Perspective. New Outlooks for Education, Enschede
- 7) S. dazu auch: Both, K. (2003), Schulqualität und innere Schulreform. In: Die Deutsche Schule, 95. Jg., Heft 2, S. 197 - 205
- 8) -Kohn, A. (2000), The Case Against Standardized Tests. Raising the Scores, Ruining the Schools, New York; Kohn, A. (2001), Fighting the Tests. A practical guide to rescuing our schools, in: Phi Delta Kappan, Vol. 82, nr. 1, S. 349-357; Meier, D./ J.Kozol/ J. Cohen (eds.)(2000), Will Standards Save Public Education?, Boston; Ohanian, S. (1999), One Size Fits Few. The Folly of Educational Standards, Portsmouth ; Ohanian, S. (2001), Caught in the Middle. Nonstandard Kids and a Killing Curriculum, Portsmouth. Internet: www.alfiekohn.org und www.fairtest.org
- 9) Niederländische Schulen sollen die Entwicklung der einzelnen Kinder in einem ‚Begleitsystem für die Schülerentwicklung‘ dokumentieren. Siehe dazu: Both, (2001), Jenaplan 21. Schulentwicklung als pädagogisch orientierte Konzeptentwicklung. Baltmannsweiler, S.51 und S. 223. Die Inspektion in den Niederlanden ist eigentlich nur interessiert an jenen Begleitsystemen, die Vergleiche von Jahrgangsklassen (Durchschnittswerte) und Schulen ermöglichen, d.h. an solchen Begleitsystemen, die ‚standardisierte‘ Tests beinhalten.
- 10) S. dazu auch Both, 2003
- 11) Elliott, J. (2002), The impact of ‚Value for Money‘ Performance Auditing in Educational Systems. In: Educational Action Research, Vol. 10/3, S. 499-505

- 12) Gute Lehrer sind in ihrer Arbeit sehr verletzlich, weil sie – allgemein gesprochen – sich selbst stark emotional engagieren in ihrer Fürsorge für die Schüler und sich stark verantwortlich fühlen für das Wohl und die Lernerfolge ‚ihrer‘ Schüler. Siehe dazu die vielen Studien über die Motivation von Lehrern und ihre Verletzlichkeit, u.a. Kelchtermans, G. (1986), *Teacher vulnerability. Understanding its moral and political roots*. In: *Cambridge Journal of Education*, Jhrg. 26, Heft 3, S. 307-323; Hargreaves, A. (1998), *The emotional practice of teaching*. In: *Teaching and Teacher Education*, Jhrg. 14, Heft 8, S. 835-854; Troman, G. (2000), *Teacher Stress in a Low-Trust Society*. In: *British Journal of Sociology of Education*, Vol. 21/3, S. 331-353; Vandenberghe, R. / M. Huberman (1999), *Understanding and preventing teacher burnout. A sourcebook of international research and practice*. Cambridge
- 13) Darling-Hammond, 1997
- 14) Darling-Hammond, 1997, a.a.O.
- 15) Meier, u.a. (2000)
- 16) S. dazu auch: Darling-Hammond, L./J. Ancess/B. Falk (1995), *Authentic Assessment in Action*. New York/London: Teachers College Press; Allen, D. (ed.) 1998), *Assessing Student Learning*. New York/London: Teachers College Press; Blyth, T., u.a. (1999), *Looking Together at Student Work*. New York/London: Teachers College Press.
- 17) S. dazu Meier, D. (1996), *The Power of Their Ideas. Lessons for America from a Small School in Harlem*. Boston
- 18) McNeil, L.M. (2001), *Contradictions of School Reform. Educational Costs of Standardized Testing*. New York
- 19) Eisner, E.W. (2001), *What does it mean to say a school is doing well?*, in: *Phi Delta Kappan*, Vol. 82, nr. 1, p. 367-372
- 20) Vgl. für eine Kritik dieser einseitigen Auffassung von ‘Wissenschaft’: Eisner, E.W. (1991), *The Enlightened Eye. Qualitative Inquiry and the Enhancement of Educational Practice*. New York
- 21) Elliott, J. (1998), *The Curriculum Experiment. Meeting the challenge of social change*, Buckingham/Philadelphia. Kap. 5
- 22) Karin Kleinespel zeigt dies auch in: *Schulpädagogik als Experiment*. Weinheim/ Basel, 1998. In den 70-er Jahren des vorigen Jahrhunderts gab es in Deutschland und den USA schon ähnliche Kritik. Vgl. Rumpf, H. (1971), *Zweifel am Monopol des Zweckrationalen Unterrichtskonzepts*. In: *Neue Sammlung*, Jhrg. 11, Heft 4; Karrier, C. (1972), *Testing for Order and Control in the Corporate Liberal State*. In: *Educational Theory*, Jhrg. 22, S. 159-180; Patton, M.Q. (1975), *Alternative Evaluation Research Paradigm*. Monograph North Dakota Study Group on Evaluation. Grand Forks: University of North Dakota; Williams, J.D. (1975), *Testing and the Testing Industry: A Third View*. Monograph North Dakota Study Group on Evaluation. Grand Forks: University of North Dakota. In 1974 wurde die North Dakota Study Group on Evaluation gegründet von Vito Perrone, Lilian Weber, u.a. um die Entwicklung von ‘statewide testing’ kritisch zu begleiten. Seitdem eine ganze Reihe von Broschüren (‘monographs’) veröffentlicht. Diese netzwerkartige Gruppe – mit Mitglieder aus den ganzen USA - ist noch immer aktiv. Pädagogen wie Carini und Meier sind Mitglied dieses Netzwerks.
- 23) S. auch: Eisner, E.W. (1992), *Educational Reform and the Ecology of Schooling*, in: *Teachers College Record*, Vol. 93, nr. 4, p. 610-627; Eisner, E.W. (1998), *The Kind of Schools We Need. Personal Essays*, Portsmouth
- 24) Wassermann, S. (2001), *Quantum Theory, the Uncertainty Principle and the Alchemy of Standardized Testing*, in: *Phi Delta Kappan*, Sept. 2001, p. 28-40
- 25) S. Anmerkung 16
- 26) Siehe dazu: Gipps, C.V. (1998) *Beyond Testing: Towards a Theory of Educational Assessment*, London/Washington, D.C.; Gould, S. J. (1983), *Der falsch vermessene Mensch*, Basel - Boston - Stuttgart
- 27) Feuerstein, R. (1998), *Don't Accept Me as I Am*. Andover (Ma).
- Andere Beispiele einer anderen Auffassung von Intelligenz (einer ‚Dekonstruktion‘ des konventionellen IQ) sind Robert J. Sternberg (Yale University) und David Perkins (Harvard University). Sternberg spricht auf Grund seiner Erfahrung in der (Erforschung von) Selektionsprozessen über drei Typen der Intelligenz, die wichtig sind für eine gute Lebensführung, Beruf eingeschlossen: analytische Intelligenz, kreative Intelligenz und praktische Intelligenz. Er glaubt dass alle Menschen die Möglichkeit haben, diese Intelligenztypen zu entwickeln, das ganze Leben lang, und ermutigt Erzieher, sie alle zu fordern. Der IQ betrifft nur den ersten Typ und enthält deshalb eine sehr beschränkte Aussage über späteren Erfolg in Schule und Beruf. S. Sternberg, R. (1988), *The triarchic mind: A new theory of human intelligence*. New York; Sternberg, R. (1998), *Thinking styles*. New York
- Perkins stellt dar, dass Intelligenz unterrichtet und gelernt werden kann – ‚the good use of the mind; the artful deployment of our faculties of thinking‘ – Perkins, D. (1995), *Outsmarting IQ. The emerging science of learnable intelligence*. New York.
- Siehe für eine Übersicht von ‚Changing Perspectives About Intelligence‘: Costa, A.L./ B. Kallick (2000), *Discovering and Exploring Habits of Mind*. Alexandria, Kap. 1.
- 28) Hart, S. (1998), *A Sorrytail: Ability, Pedagogy and Educational Reform*, in: *British Journal of Educational Studies*, Vol. 46, nr. 2, p. 153-168.

29) Standaert, 2001

30) Gould, 1983, S. 24. Zitiert in: Dixon, A./M.J. Drummond/ S. Hart/ D. McIntyre (2002), Back where we started? Developing teaching free from ability labeling. In: Forum for promoting 3-13 comprehensive education, Vol. 44, Nr. 1, S. 7-12

31) Dixon, u.a. (2002). Der Bericht dieses Projektes wird in 2004 als Buch veröffentlicht.

32) Spady, W. (2001), Beyond Counterfeit Reforms. Forging an Authentic Future for All Learners, Lanham/ Maryland/ London; Elliott, 1998; Castells, M. (2002), Die Netzwerkgesellschaft. Opladen; Aalst, H. van (2000), Leren in de netwerkmaatschappij (*Lernen in der Netzwerkgesellschaft*), in: Mensen-kinderen. Tijdschrift voor en over Jenaplanonderwijs, jrg. 16, nr. 1

33) Kohn, 2000

34) Taylor, Ch. (1998), Das Unbehagen an der Moderne. Frankfurt a.M.

35) Insbesondere ist eine kritische Analyse der ‚Performativität‘ wichtig. Dieser Begriff wurde vom französischen Philosoph Lyotard ‚erfunden‘ und beschreibt den Alldruck der ‚Effektivität‘ und ‚Effizienz‘, die an konkreten Leistungen ‚demonstriert‘ werden sollen. Im Namen der sozialen Gerechtigkeit und wissenschaftlichen Wahrheit, wird die Legitimation der Macht gegründet auf deren Optimalisierung der ‚performance‘ des Systems – ‚Effektivität‘. Alles wird gemessen und miteinander verglichen nach dem Maß der Effektivität, ohne dass jemand einen Gedanken darauf verschwendet, warum man das so machen soll.

Unterschiede werden heruntergespielt, alles wird an demselben Standard gemessen. Die Modernität parodiert auf diese Weise sich selbst, wobei Antworten gegeben werden, ohne eine Idee der ursprünglichen Fragen‘. Blake, N., u.a. 1998), Thinking Again. Education after postmodernism. Westport/ London, S. 1 – Übersetzung Kees Both. Kap.9 desselben Buches erläutert dieses Vorgehen ausführlich. Siehe dazu auch: Gleeson, D./ C. Husband (eds.)(2001), The performing school. Managing teaching and leaning in a performance culture. London/ New York

36) Siehe Anmerkung 1 und 16

37) Kohn, A. (1999), The Schools Our Children Deserve, Boston/ New York

Dieser Artikel war zuerst in der Pädagogischen Rundschau H. 2 März-April 2004, S. 151-170 veröffentlicht.

Für die Erlaubnis des Nachdrucks dieses Artikels danken wir herzlich Dr. Birgit Ofenbach von der PR.

**Kees Both ist Studienleiter des
Niederländischen Jenaplanverbandes (NJPV)
und Chefredakteur der niederl. Jenaplanzeitschrift ‚Mensen-kinderen‘
sowie Mitarbeiter im europäischen Netzwerk für Schulerneuerung 21**