

Jürgen Oelkers

Schulentwicklung in Deutschland und die Bedeutung der Selbstevaluation^{)}*

Der Ausdruck „Evaluation“ war vor zehn Jahren ein pädagogisches Fremdwort und löste noch vor fünf Jahren erhebliches Stirnrunzeln aus, weil das Fremdwort inzwischen mit anderen unbekanntem Ausdrücken assoziiert wurde wie etwa „Bildungsstandards“ oder „Outputsteuerung.“ Die schulpädagogische Sprache war bei allem Erfindungsreichtum schon immer etwas verwirlich, man denke nur an die unnachahmlichen Hybriden der Curriculumtheorie, die „Lernzielhierarchien“ vorsahen, als sei der Schulerfolg eine Frage der richtigen Ableitung von oben nach unten oder als könnte der Unterricht durch die Beschwörung seiner Ziele besser werden.

Was dieser Sprache fehlte, war eine Datenbasis. Die Lernziele standen auf dem Papier; ob sie erreicht wurden oder nicht, war Objekt einer vorempirischen Hoffnung, die Lehrplankonstrukteure oft in die Form einer messianischen Erwartung kleiden, wenn sie Auskunft geben müssen, was ihre schwergewichtigen Produkte in der Praxis bewirken werden. Lehrpläne sind eine seltsame Literaturgattung, ihre Wirksamkeit schwindet mit dem Umfang und im gleichen Masse nimmt auch die Leserlichkeit ab. Die Arbeit von Lehrplankommissionen wird bekanntlich nicht daran gemessen, wie viele Lehrkräfte das Produkt tatsächlich benutzen, sondern ob die staatlichen Vorgaben beachtet wurden.

Die Karrieren von Lehrplankommissionen sind nie untersucht worden, aber sie wären sicher ein guter Indikator für die Art und Weise, wie Bürokratie funktioniert. Der Transfer vom Produkt in die Praxis ist nicht überwältigend, um es freundlich zu sagen. Wer über die reale Schulentwicklung in Deutschland etwas sagen will, sollte vor Augen haben, mit welcher Kreativität viele Kommissionen immer neue Bezeichnungen für Fächer, Lernfelder und Leitziele erfinden, aber zugleich auch sehen, dass die eigentliche Steuerungsgröße, die Stundentafel und so die Verteilung der Zeit, faktisch unverändert ist. Es wäre ein Vortrag eigener Art, darüber nachzudenken, welche Folgen die Einführung flexibler Zeitmargen hätte, die die starren Schulstunden ablösen würden. Aber das Thema ist ja nicht Zeit, sondern Evaluation.

Die Lehrplanentwicklung hat vor allem einen Effekt, der als notwendig erachtete Differenzierungsgrad steigert die Unübersichtlichkeit. Ich spreche natürlich nur über Schweizer Lehrpläne, die eine erstaunliche Mutation hinter sich haben. Aus kurzen, für die Praxis gedachten Stoffverteilungen wurden ebenso umfangreiche wie imposante Elaborate, deren Nutzwert mit dem Umfang ständig sank. Das ist leicht erklärbar: Lehrpläne orientieren sich nicht an den Nutzern, sondern an anderen Lehrplänen. Sie gehören zu dem, was die Ökonomen den „Input“ nennen, also zur guten Absicht und nicht zum Effekt. Was letztlich herauskommt, hat wohl einen früheren deutschen Bundeskanzler umgetrieben, aber nicht das Bildungssystem.

^{*)} Vortrag anlässlich der Tagung „Von Daten zu Taten – SEIS als Motor der Schulentwicklung“ am 11. September 2008 in Berlin.

Das Wort „Outputsteuerung“ verliert seine Befremdlichkeit, wenn man es wie Helmut Kohl versteht, parteiübergreifend natürlich. Die Ressourcen der Bildung müssen in sichtbare Erträge umgesetzt werden, möglichst solchen, die zur Zweckbestimmung passen. Heute sind die „Erträge“ weitgehend identisch mit Schulabschlüssen und Berechtigungen, die von der Fiktion ausgehen, hinter ihnen stünde eine angenähert gleiche Qualität. Genauer müsste ich sagen, die Abschlüsse reflektieren traditionell Verwendbarkeit auf dem Arbeitsmarkt und nicht die Schulqualität. Daher ist es eigentlich egal, wo man den Abschluss macht, die Hauptsache ist, dass er angeboten wird. Es ist dem deutschen Bildungssystem eher fremd, von der Qualität auszugehen und dann erst den Abschluss zu betrachten.

Keine Angst, ich werde keine revolutionären Forderungen in die Welt setzen. Das Berechtigungswesen ist eine eiserne Grösse, die den Gegenwert der Schule beschreibt und die für ihren Besuch vermutlich wichtiger als die Bildungserwartung. Dennoch, die Qualität der Abschlüsse, nicht nur ihr formaler Gegenwert, spielt in Zukunft etwa bei Einstellungen und Übertritten in weiterführenden Schulen eine zentrale Rolle. Darauf deuten zum Beispiel veränderte Zeugnisse hin, die von Kompetenzen ausgehen und nicht einfach Fächer benoten. Damit ist zugleich die Frage gestellt, wie diese Qualität zustande kommt und welche Rolle die Schule dabei spielt. Es ist die Frage nach ihrer eigenen Qualität und so nach ihrer Verantwortung für den Ausbildungserfolg.

„Schulqualität“ war ein weiteres Fremdwort, das die deutsche Lehrerschaft neu lernen musste. Lange war dabei Widerwille spürbar; man vermutete eine neuerliche Attacke der Bildungsverwaltung, die nur einen Zweck hatte, nämlich die Existenzberechtigung von Behörden unter Beweis zu stellen. Tatsächlich waren die ersten Versuche zur Qualitätssicherung verbunden mit ziemlich umständlichen Verfahren diskursiver Steuerung, die wohl die Anzahl der Arbeitsgruppen und Sitzungen in die Höhe trieben, nicht jedoch den Ertrag verbesserten, auch weil ganz unklar war, was eigentlich den „Ertrag“ ausmachen sollte. Diskurse sind leicht endlose Kommunikationsschlaufen, die wohl in der Schule stattfinden, aber nicht den Unterricht erreichen. Der aber stellt das Kerngeschäft dar. Der Kanal zum Unterricht sind die Lehrkräfte, und die fragen meistens nicht nach der Philosophie eines Angebots und der damit verbundenen Diskussionsform, sondern nach dem Nutzen.

Evaluationen sind heute Teil der Qualitätssicherung in allen entwickelten Schulsystemen. Ich gebe zu, dass derzeit in der Schweiz auch wieder ein neuer Lehrplan entsteht, der ein wirklicher Anlass zum Stirnrunzeln sein könnte, wenn man von der Dauer des Projekts auf den zu erwartenden Umfang schliesst. Viel schneller und viel effizienter sind Evaluationsverfahren implementiert worden. Heute kennen die Schulen in fast allen Kantonen standardisierte Verfahren der Selbstevaluationen und vom Thurgau über Zürich bis Bern sind inzwischen auch Agenturen tätig, die mit externen Evaluationen beauftragt werden. Die Bertelsmann-Stiftung, will ich sagen, hat mit dem Programm SEIS keine Anomalie in die Welt gesetzt.

Eher muss die Tatsache, dass deutsche und früher auch Schweizer Schulen ohne jede datengestützte Entwicklung auskamen, als Anomalie betrachtet werden. Die historische Linie ist das nicht. Qualitätssicherung in Form von Schulinspektionen kannte die deutsche Schule seit der Reformation, der erste Inspektor war Luther selbst, der vom Juli 1528 an Visitationsreisen unternahm, um die Qualität der religiösen Unterweisung zu überprüfen, mit erschreckenden Resultaten, die, wie man heute sagen würde, unmittelbar „Handlungsbedarf“ nahelegten. Luther hat das etwas drastischer gesagt. Wenn ich einen etwas überzogenen Schluss ziehen darf, dann war das wohl die erste deutsche „Bildungskatastrophe.“

Sie hat enorm zur Schulentwicklung beigetragen, der Schrecken erzieht, wenn zugleich Auswege offen stehen; aber es war natürlich eine Defizitbestimmung, die zum Grundmuster der Qualitätsbeschreibung wurde und sich bis zu dem hinzieht, was die deutschen Medien leicht hysterisch „PISA-Schock“ genannt haben. Defizite sind die Grundneigung der Pädagogik, die leicht sagt, was fehlt und schlecht ist, aber sich schwer damit tut anzugeben, was erreicht werden kann. Das nämlich würde einen Wechsel des Blickes verlangen, weg vom Defizit und hin zum Potential.

Eine Potentialanalyse kennt erst die neueste Diskussion, die davon ausgeht, dass ein Schock die Wahrnehmung mehr beeinträchtigt als anregt. Soll man ängstlich alle drei Jahre auf PISA warten? „Bildungskatastrophen“ sind nur dann gut, wenn sie Medienereignisse bleiben und nicht dazu führen, bei sich selbst nur noch Schwächen zu sehen. Das Potential der Bildung geht aus von der Bestimmung der eigenen Stärken, die danach in ein Verhältnis zu den erkennbaren Schwächen gesetzt werden. Sonst überlässt man das Feld den professionellen Schwarzsehern, die es reichlich gibt und die das sind, ohne dass man sie gerufen hätte.

Das moderne Bildungssystem hat immer versucht, ohne Schwarzseher auszukommen, auch wenn es oft die Lehrkräften waren, die den Sinn ihres Systems in Frage stellten, sei es aus Gründen der Entlöhnung oder der Entfremdung. Aber die Bestimmung des eigenen Potentials, also der Stärken und Schwächen, war immer eine Notwendigkeit und richtete sich nicht aus an der radialen Schulkritik. Die Frage war schon früh, auf welcher Datenbasis die Schulaufsicht erfolgen sollte. Lange waren die probaten Mittel Inspektionsreisen und Schulvisitationen, verbunden oft mit öffentlichen Examen und der Fortbildung der Lehrkräfte im Anschluss daran. Das war praktisch, weil die sichtbaren Schwächen gleich korrigiert werden konnten - Gemeint waren die Schwächen der Lehrkräfte, die sich in den Leistungen der Schüler spiegelten.

Der Beruf des „Schulinspektors“ entstand im 19. Jahrhundert. Er war solange bei den Lehrkräften wohl gelitten, als Loyalität vorausgesetzt werden konnte. Manche Inspektoren verbreiteten Angst und Schrecken, vor allem dann, wenn ihre Berichte dem Bild widersprachen, das die Lehrkräfte von sich selbst hatten. Ihre Reaktionen sind historisch kaum dokumentiert, aber sie hängen mit dem zusammen, was Siegfried Bernfeld das „beleidigte Pädagogengemüt“ nannte, also die Enttäuschung der Effekterwartung trotz bester Absichten. Bernfeld war Psychoanalytiker, aber er hat nie gesagt, ob sich seine Beschreibung auf die Neurosenlehre beziehen lässt.

Inspektoren sind in der Schweiz irgendwann in den siebziger Jahren abgeschafft oder im Auftrag „redimensioniert“ worden, wie man dort sagt. Irgendwie schienen sie nicht mehr zeitgemäss zu sein und wenn man Behörden abbauen kann, dann tut man das in der Schweiz. Die Konsequenz war, dass die Schulen keinen Blick von Aussen mehr nutzen konnten. Da die Inspektoren verschiedene Schulen kannten, konnten sie Vergleiche ziehen und die Unterschiede thematisieren. Auch das entfiel mit der Auflösung ihrer Funktion. Aber Schulaufsicht kann nicht einfach nur in Verwaltungsarbeit bestehen, so dass die Frage der Datengewinnung zurückkehrte, allerdings in einer anderen Gestalt.

Die Idee der Rückkoppelung von Daten, der Rechenschaftslegung von Schulen und damit der Übernahme von Verantwortung (*accountability*) hat als einer der ersten der amerikanischen Bildungsökonom Henry M. Levin 1974 in einem Aufsatz für die Zeitschrift *School Review* dargelegt. Die Praxis der Evaluation ist älter, die Theorie wurde mit dem

Aufkommen der Kybernetik in der Ökonomie entwickelt. Die Verantwortung der Schule wird von Levin modellhaft mit der gezielten Evaluation ihrer Anstrengungen und der dabei je erzielten Resultate in Verbindung gebracht.

Seit diesem Begründungsaufsatz ist das Thema politisch stark aufgewertet worden, ohne an dem grundlegenden Modell von Levin viel zu verändern. Es beschreibt, was „Qualitätssicherung“ im Kern ausmacht und wozu sie in der Praxis dienlich sind. Die Verantwortung der Schule kann nicht abstrakt definiert werden, sondern verlangt die Beobachtung des Prozesses und so eine Datenbasis. Anders kann jeder Akteur Verantwortung reklamieren, ohne zur Rechenschaft verpflichtet zu sein. Oder es kann jeder Verantwortung bestreiten, obwohl die Zuständigkeit gegeben ist. Die Rechenschaft über das Erreichte und der Nutzen des bisherigen Prozesses ist die zentrale Begründung hinter dem Modell.

Man sollte das auch deutlich als Kernaufgabe einer demokratischen Schule verstehen und nicht als „neo-liberalen“ Irrweg, wie von manchen Kritikern behauptet wird. Sie verkennen, das für Schulen in einer Demokratie Transparenz grundlegend ist, die aber ohne eine verlässliche Datenbasis nicht erreicht werden kann. Wenn Schulen fragen sollen, wo stehen wir und was wollen wir in Zukunft erreichen, dann brauchen sie Evaluationen ihrer Qualität. Bilanzen bestehen nicht einfach aus Bekenntnissen, sondern aus Zahlen und Analysen, wie sie Schulen bislang noch nicht haben liefern können.

Was heute „Evaluation“ genannt wird, ist in der Praxis von Schulen allerdings keine neue Erscheinung. Ein Blick in die Geschichte lohnt sich, auch weil damit abgeschätzt werden kann, was erreicht werden kann und was vermieden werden muss, damit die Implementation Erfolg haben kann. Wenn der Schock über eine „Bildungskatastrophe“ aufgrund schlechter Daten wohl die Erhöhung der Ressourcen rechtfertigen kann, insofern politisch nützlich ist, nicht jedoch die Entwicklung vorantreibt, wie können dann Schulevaluationen nutzbar gemacht werden, also Akzeptanz gewinnen und nicht abschrecken?

Die erste grosse Vergleichsuntersuchung amerikanischer Schuldistrikte wurde Ende des 19. Jahrhunderts durchgeführt, veranlasst durch Medien und nicht durch den Staat. Verantwortlich für die Studie war der Kinderarzt Joseph Mayer Rice, der einige Jahre in New York praktiziert hatte.¹ Er gab seine Praxis 1888 auf und ging nach Europa. Rice studierte Psychologie und Pädagogik an den Universitäten Jena und Leipzig. Der Grund war die theoretische Fundierung einer Erziehung, die auf dem Prinzip der Induktion beruhen und das Lernen der Kinder in den Mittelpunkt stellen sollte. Was sie lernen oder nicht lernen, sollte den Massstab für die Schulqualität darstellen.

Nach seiner Rückkehr untersuchte Rice in grossem Stil den Zustand der öffentlichen Schulen in den Vereinigten Staaten. Er wird heute der „Vater der Evaluation“ genannt, weil erstmalig mit einem Blick von Aussen und vergleichend der Zustand der Schulen beschrieben wurde. Es gab wohl in den meisten Schuldistrikten jährliche Berichte der Superintendenten, die die Distrikte leiten; aber diese Berichte stellten kaum Vergleiche an und hatten keine unabhängig erzeugten Daten zur Verfügung. Superintendenten sind Teil des Systems, und sie berichten über Erfolg oder Misserfolg ihrer eigenen Arbeit, was die Kritikfähigkeit begrenzt, um es freundlich zu sagen.

¹ Joseph Mayer Rice (1857-1934) stammte aus Philadelphia war zweiter Sohn deutscher Einwanderer. Er promovierte 1881 am College of Physicians and Surgeons in New York. Rice war zunächst Arzt am Mount Sinai Hospital in New York und eröffnete 1886 eine eigene Praxis.

Rice war kein Schulinspektor und auch kein Superintendent, sondern ein externer Beobachter, der vermutlich erste in dieser Rolle, der auch die Differenz stark machte. Wer nicht unabhängig ist, sollte diese Rolle gar nicht ausüben. Rice verbrachte für seine erste Studie sechs Monate in Schulen, hospitierte im Unterricht und nahm am Schulleben teil. Die Methode war die einer teilnehmenden Beobachtung mit Interviews. Rice besuchte insgesamt 36 Schuldistrikte und veröffentlichte seine Resultate 1893 in der Studie *The Public School System in the United States* (Rice 1893), das zu einem der folgenreichsten Bücher der amerikanischen Pädagogik werden sollte.

Rice wollte wissen, nicht welche guten Absichten die Lehrkräfte vertreten, sondern wie die Schülerinnen und Schüler tatsächlich lernen, unter welchen Umständen und mit welchen Wirkungen Effekten. Auch die besten pädagogischen Absichten sind nichts als Hypothesen, die über den tatsächlichen Effekt nichts aussagen. Oft wird die Realität hinter der pädagogischen Rhetorik verschleiert, so dass ein unabhängiger Blick nötig ist. Entsprechend waren die Beurteilungskriterien. Der Bericht von Rice basierte auf folgenden Parametern:

- Das Erscheinungsbild der Klassenzimmer,
- die Haltung der Lehrkräfte gegen den Kindern,
- die Praxis des Rezitierens,
- die Art der Beschäftigung der Schüler (busy work),
- Antworten der Lehrkräfte auf zwölf allgemeine Fragen,
- Teilnahme an Versammlungen der Lehrerinnen und Lehrer,
- Anstrengungen der Lehrkräfte zur eigenen Fortbildung (ebd., S. 5).

Man erkennt leicht die Kriterien wieder, die auch heute bei Evaluationen angelegt werden. Lehren und Lernen war genau so ein Thema wie die Professionalität der Lehrkräfte, die Schulkultur oder das Management der Schulen, um die Qualitätsbereiche zu nennen, die hinter SEIS stehen.

Das Buch von Rice war ein Erfolg, weil es nicht beschönigte, sondern schonungslos die Missstände beschrieb. Einige der Resultate lauteten so: Die schlechtesten Lehrkräfte wurden durchgehend in den unteren Klassen eingesetzt, weil die Schulen davon ausgingen, dass die jüngeren Schüler weniger Aufwand und so geringere Kompetenz erfordern. Durchgehend beschrieb Rice die Praxis des *Rote-Learnings*, also des ständigen Wiederholens und Nachsprechens von Mustersätzen, die auswendig gelernt wurden. Üblich war auch die Methode der *Group Recitations*, bei der die ganze Klasse beschäftigt wurde. Auf ein Signal des Lehrers hin musste die Klasse aufstehen und Merksätze so lange wiederholen, bis sie „sassen,“ wie man zur gleichen Zeit in der deutschen Schulpädagogik sagte. Diese Praxis war verbreitet und vor den Studien von Rice aber kaum öffentlich.

Von den 36 Distrikten, die Rice besuchte, wurden nur vier als positiv und fortschrittlich beurteilt.² Vor allem dieser Befund löste den Schock aus, die amerikanischen Schulen insgesamt schienen im Abseits zu stehen und wurden als dringend reformbedürftig hingestellt. Dagegen liefen die Lehrkräfte Sturm, die grössere Öffentlichkeit allerdings akzeptierte den Befund, der sich mit den Erfahrungen der Eltern deckte und der so alles andere als unglaublich war. Rice zog sich aber den Zorn der Lehrkräfte zu, die ihm Auskunft gegeben und ihre Klassenzimmer für seine Untersuchung geöffnet hatten. Andere

² Alle diese Schulsysteme waren im Mittleren Westen, in Indiana, Indianapolis, in La Porte, Indianapolis, die Lincoln Schule in Minneapolis sowie die Francis Parker Schule im Cook County.

Lehrkräfte befürchteten, von ihm besucht zu werden. Aber das Prinzip der externen Visite oder der Evaluation unabhängig von der Schulaufsicht war etabliert.

Es gab auch positive Beispiele in seiner Evaluation, die typischerweise auf einen zentralen Faktor zurückgeführt wurden, die Bereitschaft der Lehrkräfte, selbst zu lernen, und dies nicht nur in professioneller Hinsicht, sondern auch bezogen auf ihre allgemeinen intellektuellen Fähigkeiten. Die grössten Fortschritte in der Schul- und Unterrichtsentwicklung werden dort erzielt, „where the teachers themselves are the most earnest students.“ Es war Rice, der den Satz prägte: „It is, after all, the teacher that makes the school“ (ebd., S. 18).

Rice war der Auslöser der Evaluations-Bewegung, die eine Sicht von Aussen anlegen will, um die Qualität von Schulen besser einschätzen zu können. Heute ist dieses Prinzip weitgehend anerkannt. Rice dagegen war umstritten, nicht nur bei den Lehrkräften, sondern vor allem auch bei den Vertretern der Schulaufsicht, die seine Daten und Befunde nicht hören wollten. Vor allem wollten sie nicht einsehen, dass *Rote-Learning* die Ursache der Misere war. Was sie bei ihnen Visitationen die „Übung der Kräfte des Kindes“ nannten, war nach Rice genau das, was das Lernen behinderte und die schlechten Resultate hervorbrachte. Aber das wurde von der Schulaufsicht massiv bestritten. Es dauerte bis 1915, bevor sich das Prinzip der wissenschaftlichen Evaluation in der amerikanischen Schulaufsicht durchsetzte.

Ein Grund für den Widerstand war die fehlende Selbstevaluation. Rice untersuchte Schulen, die nicht vorbereitet waren, und zwar weder auf seine Fragen noch auf seine Schlussfolgerungen. Dann ist die Folge Abwehr und nicht Öffnung für die Daten. Schulen sind erprobt darin, sich Zumutungen zu entziehen und entwickeln viel Geschick, Vorlagen zu unterlaufen und in ihrem Sinne anzupassen. Was sich nicht als nützlich erweist, verliert schnell jede Priorität. Manchmal vergisst das die Kultusbürokratie. Steuerung setzt Einsicht der Akteure und Nutzen für das Handlungsfeld voraus, also mehr als nur die Erlasslage (Oelkers/Reusser 2008).

Die Idee der datengestützten Steuerung geht zurück auf einen programmatischen Aufsatz von Leonard Porter Ayres,³ der begründete, warum Bildungsprozesse *durch Resultate* gesteuert werden sollten. Der Aufsatz erschien 1912 in der amerikanischen Zeitschrift *School Review* (Ayres 1912). Ayres war Statistiker und setzte auf Steuerung allein durch Daten, also durch Kennziffern, Leistungstests und Prüfungsergebnisse. Aber einzig das hat Schulen noch nie zum Handeln veranlasst, es sei denn der Staat droht mit Sanktionen wie die Schliessung der Schule bei zu schlechten Leistungsdaten. Das machen inzwischen amerikanische Bundesstaaten. Aber wer eine schlechte Schule schliesst, hat dadurch nicht schon eine bessere.

Im gleichen Jahr 1912 wurden auf der jährlichen Versammlung der amerikanischen Superintendenten fast 50 Anträge gestellt, die die Einführung von Leistungstests und anderen Massnahmen zur Sicherung von Schulqualität betrafen. Die Vorschläge waren hochgradig umstritten und lösten kontroverse Diskussionen aus. Erst drei Jahre später beschloss die gleiche Versammlung, den Weg frei zu geben für die Einführung von Evaluationsinstrumenten, die die amerikanische Schule seitdem mehr kennzeichnen als alle Methoden der progressiven Erziehung. Das Stichwort dafür lautete bereits vor dem ersten

³ Leonard Porter Ayres (1879-1946) war nach seinem Studium sechs Jahre lang als Lehrer und Superintendent auf Puerto Rico tätig. Ayres war Statistiker und leitete von 1908 an das Bureau of Education and Statistics der Russell Sage Foundation.

Weltkrieg „school efficiency.“⁴ Die Administration sollte sich darauf einstellen, dass Schulen nicht nur verwaltet werden können, sondern auch entwickelt werden müssen und dies gleichermassen effizient wie verständnisvoll.

Das schweizerische Projekt „HarmoS“ vollzieht diesen Schritt heute nach, in gebührendem historischem Abstand, wie das in der Schweiz üblich ist, der auch darüber belehrt, welche Fehler vermieden werden müssen. Sanktionsdrohungen sind sicher kein Mittel zur weiteren Entwicklung des Schweizer Schulwesens, allerdings werden jetzt auch hier die schlechten Schulen sichtbar und das wirft die Frage auf, wie mit ihnen umgegangen werden soll, zumal dann, wenn der Befund öffentlich wird. Die Antwort ist nicht Bestrafung, sondern Ursachenforschung, Akzeptanz der Gründe und klare Entwicklungsziele, die wenn nötig auch extern moderiert werden können.

Gesteuert wird das Bildungssystem nicht allein durch regelmässige Datenerhebungen, sondern auch durch Weiterbildung, Organisationsentwicklung und nicht zuletzt durch Lehrmittel, die immer vergessen werden, wenn es um Qualitätssicherung gehen soll. Und das amerikanische Beispiel wirft auch die Frage auf, warum die durchschnittliche Schulqualität nicht längst besser eingeschätzt wird, geworden ist, wenn doch seit mehr als hundert Jahren Leistungsmessungen und Schulevaluationen durchgeführt werden. Auffällig ist, dass gerade die amerikanische Bildung immer wieder - und dies regelmässig - unter das Feuer einer Kritik gerät, die die Leistungsfähigkeit des Systems grundsätzlich in Frage stellt. Jede neue Generation scheint ihr Krisenerlebnis zu wünschen, während die tatsächliche Datenlage eher auf langfristige Verbesserung hinweist.

Das Programm „Selbstevaluation in Schulen“ oder SEIS ist entwickelt worden, um den einzelnen Schulen Gelegenheit zu geben, nach Innen ihre Stärken und Schwächen zu bestimmen, ohne lediglich eine Rhetorik der guten Absichten zu benutzen. Diese Rhetorik hat ihren Platz in den Leitbildern, die zur Selbstverständigung notwendig sind, aber nicht mit der Wirklichkeit gleichgesetzt werden dürfen. Wenn man im Blick auf das Wünschbare übereinstimmt, hat man nicht schon etwas bewegt, insofern sind die neuen Verfahren der Steuerung auch ein Mittel gegen eine allzu idealistische Pädagogik, die sich mit Worten beruhigen kann, ohne den Worten Taten folgen lassen zu müssen.

SEIS erzeugt Daten. Die relevanten Gruppen der Schule werden befragt und die Antworten werden ausgewertet, ohne dass die Schulen eigene Fragebögen entwickeln und eigene Software herstellen müssten. Insofern ist SEIS ein Service, den jede Schule ohne allzu grossen Aufwand nutzen kann. Das System ist bewährt und kann jederzeit angereichert werden. Erhebung, Auswertung und Kommunikation der Daten bleiben intern. Es handelt sich um eine *Selbstevaluation*, mit der die Schulen versuchen, über sich ein realistisches Bild zu gewinnen, ohne zwanghaft nur die angenehmen Seiten sehen zu müssen.

Die Frage ist nicht ob SEIS funktioniert oder nicht, das System ist nach vier Jahren Erprobung ausgereift und kann in jeder Schule eingesetzt werden. Die Frage ist, was die die Schulen mit den von ihr selbst erhobenen Daten anfangen. Die Daten beschreiben ziemlich genau, wo die Eltern, die Lehrkräfte oder auch die Schülerinnen und Schüler die Stärken und Schwächen ihrer Schule sehen. Damit werden Potenziale sichtbar, aber auch kritische Punkte

⁴ *School Efficiency Series* hiess auch eine Buchreihe, die der Mathematiker und Berufspädagoge Paul Henry Hanus (1855-1941) im New Yorker Verlag World Book Company herausgab. Hanus war Gründer der Harvard School of Education und leitete zahlreiche Reformkommissionen, darunter von 1906 bis 1909 die Massachusetts Commission on Industrial Education.

und Schwachstellen, die bislang leicht übergangen werden konnten. Nunmehr antworten viele oder alle Eltern und dadurch entstehen ganz andere Prioritäten.

In der Schweiz bezeichnet man die Verweigerung der Weiterbehandlung von Berichten mit dem schönen Wort „schubladisieren.“ Wenn etwas in der Schublade verschwindet, ist zwar der Deckel drauf, aber auch eine Chance vertan. Rückmeldungen sind Chancen, auch wenn - oder weil - nicht immer nur positive Daten kommuniziert werden. Kritik ist nützlich und schadet nichts, vor allem aber begreift man Eltern, Lehrer und Schüler - oder wen immer man sonst befragt - als Ressource für die Schulentwicklung, die ernst genommen werden muss.

Allein die Tatsache, dass überhaupt gefragt wird, verändert die Situation nachhaltig. Wer nämlich einen Fragebogen ausfüllt, kann nachfragen, was aus den Antworten geworden ist und trägt so zur Transparenz bei. Die Schulen stellen sich mit SEIS nicht nur ihren Abnehmern, sondern nutzen zugleich deren Know How, auch und gerade dann, wenn die Rückmeldung nicht positiv ist. Wer die Elternarbeit verbessern will, darf nicht nur mit den Aktivisten reden und kann auch nicht von der Erwartung ausgehen, alle Eltern seien gleich und gar noch gleich positiv auf die Schule eingestellt. Die Grundregel sollte vielmehr sein, dass die Schule etwas von den Eltern lernen kann.

Eltern und Schüler beobachten die Schule mindestens ebenso intensiv wie die Lehrkräfte oder die Schulleitung, nur aus einer ganz anderen Perspektive. Wenn man bestimmen will, was SEIS für die Schulentwicklung bedeutet, dann muss man sich nur vor Augen halten, dass diese Perspektive bislang nie benutzt wurde. Eltern und Schüler haben immer Urteile über den Unterricht, die Schule und die Lehrkräfte, die sie erleben; sie haben oft auch Meinungen darüber, was verbessert oder anders gemacht werden könnte; aber dieses Know How ist von den Schulen nicht als Ressource und so auch nicht als Chance verstanden worden.

Eltern werden von den Lehrkräften gelegentlich als Störfaktor betrachtet, der besser nicht aktiviert werden sollte, während Unterricht umso mehr Erfolg hat, je mehr ihn die Eltern unterstützen. Das ist bislang eine stillschweigende Dienstleistung, die die Schule gerne in Anspruch nimmt, ohne sie honorieren oder auch nur beachten zu müssen. Dabei gehört die tägliche Einstellung der Kinder auf die Schule zu den Leistungen, die alles andere als selbstverständlich sind. Schon das spricht dafür, mit den Eltern ernsthaft und folgenreich zu kommunizieren.

Wirksamer Unterricht hat wesentlich mit den Eltern zu tun, auch wenn die Lehrkräfte das anders sehen. Die Förderung des Lernens ist nicht nur eine Aufgabe der Schule, aber oft ist unklar, wie die Eltern über moralische Appelle hinaus sinnvoll beteiligt werden können. SEIS bietet dazu eine Plattform, die konkrete Rückmeldungen erlaubt und von allen Eltern genutzt werden kann. Auch diese Form von Partizipation hat Folgen. Es könnte ja sein, dass das unüberwindliche Ghetto der „bildungsfernen Schicht“ einfach nur eine Frage der fehlenden Anreize und der mangelhaften Kommunikation ist.

Aber auch die Lehrkräfte wissen wenig, was die jeweils andere Gruppe über die Schule denkt. Wenn man überlegt, dass im Unterricht fast nie die Situation thematisiert wird, in der er stattfindet, und wenn das doch der Fall ist, dann aus Vorfällen und Krisen heraus, dann erkennt man den Wert einer gezielten Abfrage von Know How. Lehrkräfte gehen oft von der irrtümlichen Annahme aus, dass ihr Unterricht alle Schüler mehr oder weniger gleich erreicht, so dass dann auch deren Urteile über den Unterricht mehr oder weniger gleich sind.

Aber das ist nicht einmal ein Grenzwert. Die Schüler sehen den Unterricht unterschiedlich, oft sehr differenziert und aber fast nie als Rückmeldung an die Lehrkräfte.

Auch aus diesem Grund ist der Unterricht für Schüler und Eltern oft intransparent. Er findet statt, aber hat keine Rückkopplung. Oft nehmen sich die Lehrkräfte in einem Wirkungsdilemma wahr, sie wollen Ziele erreichen und sehen Streueffekte. Der Unterricht setzt eine Black Box voraus, die sich durch Daten nicht verunsichern lässt, aber nur deswegen, weil Daten gar nicht erhoben werden. SEIS ist der Einstieg nicht in eine schulische Audit-Gesellschaft, sondern in eine Rückmelde-Kultur, die auf Respekt beruht. Der soziale Bindeeffekt in einer Schule ist umso grösser, je mehr die verschiedenen Gruppen eingebunden sind und die Schule zu „ihrer“ machen. Und wie wäre man stärker eingebunden, als mit der Erfahrung, dass die eigene Meinung zählt?

Literatur

- Ayres, L-P.: Measuring Educational Processes through Educational Results. In: School Review Vol. 20 (1912), S. 300-309.
- Levin, H.M.: A Conceptual Framework for Accountability in Education. In: School Review, Vol. 82, No. 3 (May 1974), S. 363-391.
- Oelkers, J./Reusser, K.: Expertise: Qualität entwickeln - Standards sichern - mit Differenz umgehen. Unter Mitarbeit von E. Berner/U. Halbheer/St. Stolz. Berlin: Bundesministerium für Bildung und Forschung 2008.
- Rice, J.M.: The Public-School System of the United States. New York: The Century Co. 1893.