

Jürgen Oelkers

Schule, Bildungsstandards und realer Unterricht^{)}*

1. *Schule, Staat und Bildung*

Bildung wurde im 19. Jahrhundert gleichermassen verstaatlicht und verschult. Der Staat übernahm mit den Kosten auch die curriculare Steuerung, einschliesslich der Normierung der didaktischen Formate und der zur Verfügung gestellten Zeit. Unterricht wurde auf ein einheitliches Zeitmass eingestellt und mit Effektivitätsanforderungen konfrontiert, die umso mehr Organisation erforderten, je engmaschiger sie verstanden wurden. Notwendig für den Aufbau der Organisation waren die Professionalisierung der Lehrerschaft, die Standardisierung der Lehrmittel, die Normierung der Überzeugungen und vieles mehr. Die Ausbreitung der Bildung geschah flächendeckend und unter der Voraussetzung eines bestimmten institutionellen Arrangements, das bis heute massgebend ist.

Die Schulen, könnte ich auch sagen, werden in ihrer Grundform nicht ein zweites Mal erfunden, und wir müssen mit ihnen leben. Zu Schulen gehören Standards und das ist offenbar nicht einfach eine Last. Die Entwicklung des modernen Schulsystems war in vielen Hinsichten eine Erfolgsgeschichte und eine Fortschrittserfahrung. Das gilt materiell wie symbolisch:

- Die Investitionen stiegen im historischen Längsschnitt rapide an,
- die staatlichen Haushalte waren von einem bestimmten Zeitpunkt an nicht mehr umkehrbar,
- die Bildungsversorgung wurde zu einer festen gesellschaftlichen Erwartung
- und im Gegenzug stiegen die Leistungserwartungen und gewann die Schule öffentliches Ansehen.

Tatsächlich ist auf diesem Wege Bildung erweitert und verbreitet worden, und dies nicht einfach numerisch, sondern im Blick auf die durchschnittlich erreichte Qualität. Ohne öffentliches Schulwesen hätte die heutige Bildungsqualität nicht aufgebaut werden können, wäre es unmöglich, Basisfertigkeiten über Generationen konstant zu halten und könnte keine ungefähre Gleichverteilung des Angebots erreicht werden.

Das setzte und setzt ein Kalkül des Nutzens voraus, keine Gesellschaft könnte sich auf Bildung einlassen, wenn sie nichts davon hätte. Der Bildungsnutzen wird mit Abschlüssen und Berechtigungen kalkuliert, was schon im 19. Jahrhundert Anlass war für heftige Kritik, ohne die Utilitätsrichtung des Systems wirklich zu tangieren.

^{*)} Vortrag in der Gutenbergschule Realschule Eltville am 4. Februar 2013.

- Schulische Bildung musste sich lohnen und war so nie zweckfrei, wie sich an den am Ende erreichten Abschlüssen ablesen liess.
- Die Patentierung der Leistung hiess immer auch, Gewinner und Verlierer unterscheiden zu müssen.
- Selbst dort, wo - wie in Finnland - die obligatorische Schulzeit selektionsfrei ist, gibt es anschliessende Zuweisungen, die ungleich sind.

Alle Bildungssysteme sind selektiv, anders könnte das Grundprinzip der Beurteilung nach Leistung nicht durchgehalten werden. Leistungen sind unterschiedlich, also können damit nicht gleiche Berechtigungen verbunden sein. Aber genau diese Ungleichheit erscheint als anstössig, was umso mehr gilt, wenn der Verdacht aufkommt, persönlich lohne sich das Ganze nicht, weil die Schulerfahrung eine künstliche ist, die nichts wirklich mit dem Leben zu tun hat.

Auch dieser Verdacht hat seine Geschichte. Wenn es auf dem umkämpften Markt der originellsten Schulkritik periodisch immer wieder heisst, die Schule sei eine einzige „Bildungslüge“ (Fuld 2004), weil sie überflüssiges Wissen vermittele, das nicht zu dem führe, was in den Zielen der Bildung behauptet wird, so ist das seit der Antike geläufig. Geändert hat diese Kritik nichts, sie rechnet nicht mit den institutionellen Bedingungen der Schule und verfügt nicht über wirkliche Daten, die den zunehmenden Verfall der Bildung erhärten könnten. Letztlich läuft das Argument der „Bildungslüge“ darauf hinaus, sich entweder einen Zustand ohne Schule vorzustellen oder aber anzunehmen, früher sei alles besser gewesen, wofür im Falle der Entwicklung gesellschaftlicher Bildung nichts spricht. Auffällig ist demgegenüber, dass immer Gebildete Kritiker der Bildung sind, für die es leicht ist, das eigene Privileg in Frage zu stellen, weil sie selbst es nicht mehr verlieren können.

Man sollte aber deutlich trennen zwischen dem, was organisierte Bildung leistet und dem, was die Kritik der Intellektuellen fordert. Oft verwendet die Kritik Massstäbe, denen die Schule mit Sicherheit *nicht* folgt, und oft ist gerade die emphatische Kernidee der Bildung an der Schule vorbei gedacht. Wenn Schüler von der Fron schulischen Lernens befreit werden sollen, dann immer in Namen einer besseren Idee der Bildung, die gleichsam hinter der Realität gesucht wird und dann aber daran nichts wirklich ändern kann.

- Schulen sind Grosssysteme, die aus eigenen Prämissen heraus lernen und keineswegs auf den Anstoss der Kritik warten und dies umso weniger, wenn die Kritik zu konkreten Problemen gar nichts zu sagen hat.
- Sie ist nur radikal, und das ist im Blick auf träge Systeme wie das der Schule zu wenig.
- Andererseits hat diese Kritik wiederum ihre Geschichte.

Seit Beginn des 18. Jahrhunderts, besteht die Möglichkeit, zwischen „Bildung“ und „Schulbildung“ zu unterscheiden, also die Schule und den Kanon der Unterrichtsfächer von der persönlichen Kultivierung abzugrenzen. Sie führt zu dem, was in der englischen Bildungstheorie die „inward form“ genannt wurde, die unverfügbar ist und sich selbst aufbauen muss.¹

¹ Der englische Platonismus hat die Entwicklung der deutschen Konzeption von „Bildung“ oder der „inneren Form“ massiv beeinflusst. Ich verweise in diesem Zusammenhang auf die Zürcher Dissertation von Rebekka Horlacher (2004).

Daraus entstand das humanistische Konzept der Selbstformung des Menschen, das Bildung auf *Welt* und nicht lediglich auf *Schule* bezieht. Die Differenz erlaubt, eine Gleichsetzung von Schule mit Bildung zu verhindern. Bildung geht nicht auf in einem bestimmten institutionellen Arrangement, Denken und Empfinden, Geschmack und Urteilskraft sind nicht einfach die Folgen organisierter Bildung, sondern setzen Welterfahrung voraus, die sich nicht patentieren lässt.

Die Bildungsromane der deutschen Klassik nutzten diese Chance ebenso wie die literarische Schulkritik des 20. Jahrhunderts, die zwischen Upton Sinclair, Thomas Mann und George Bernard Shaw illustre Namen zu verzeichnen hat. Auf der Linie dieser Kritik sind „Individualität“ und „Bildung“ kaum noch zu unterscheiden; was der Mensch von der Welt sich aneignet, ich zitiere den jungen Humboldt,² *ist* dann seine Bildung, ohne damit didaktische Formate oder Standards des Lernens zu verbinden. Diese Art Bildung ist frei, sie geschieht spontan und ihr Anlass ist nichts weniger als die ganze Welt, die nicht schulisch sortiert sein muss, um sich *in* ihr und *mit* ihr bilden zu können. Die Bildung ist sozusagen der Roman des Lebens, der sich wohl aufschreiben, aber nicht institutionalisieren lässt. In diesem Sinne ist Bildung Selbstformung, nicht Schulabschluss.

Parallel zu dieser Kritik ist das Schulsystem ausgebaut worden und entstanden die heutigen „Bildungsanstalten“, die mit dem Paradox umgehen müssen, auf das Leben vorbereiten zu wollen, ohne die freie Erfahrung von „Welt“ zuzulassen. Die Schulzeit für das, was dann euphemistisch „allgemeine Bildung“ genannt wurde, ist kontinuierlich gewachsen. In Schweizer Kantonen waren lange Zeit nur sechs Schuljahre obligatorisch, heute sind es neun Jahre, de facto aber zwölf und mehr, weil kaum jemand umhin kommt, weiter, als das Obligatorium reicht, zur Schule gehen zu müssen. Berufsschulen oder Gymnasien sind für nahezu alle Jugendlichen unumgängliche Erfahrungen, die mit dem zugebracht werden, was im schulpädagogischen Jargon „Lernen auf der Sekundarstufe II“ genannt wird.

- Dass wir von *Stufen* ausgehen, hat seinen Grund darin,
- dass die schulische Bildung als stetiger Zuwachs konzipiert wird,
- während tatsächlich nie die gleiche Qualität erreicht wird und grosse Unterschiede in Kauf genommen werden.

In diesem Sinne muss die Erfolgsgeschichte relativiert werden, oder anders, bei der Verschulung grosser Zahlen und angesichts heterogener Verhältnisse ist ausgeschlossen, dass die Schule oder auch nur Unterricht überall gleich erfolgreich sein kann. Dafür sprechen die vorliegenden Daten etwa der PISA-Studien:

- Nach neun Jahren obligatorischer Schulzeit verfügt ein knappes Viertel aller Schüler über eine Lesekompetenz, die dem Stand der fünften Klasse entspricht.
- Für sie hört die Entwicklung ihrer Kompetenz irgendwann einfach auf, während sie weiterhin zur Schule gehen müssen.

Wir wissen übrigens nicht, ob das je anders war. Die PISA-Instrumente erheben keinen historischen Längsschnitt, sondern sind periodisch wiederholte Momentaufnahmen, die

² Bruchstück *Theorie der Bildung des Menschen* (Humboldt 1980, S. 234-240). Der Text stammt vermutlich aus dem Jahre 1793, den Titel hat der Herausgeber Albert Leitzmann besorgt. Was „Theorie der Bildung“ bei Humboldt heissen soll oder nur heissen kann, ist umstritten. Ich verweise in diesem Zusammenhang auf die Lizentiatarbeit von Werner Hürlimann (Zürich).

allerdings zuverlässig die durchschnittliche Verteilung von Qualität in bestimmten Fächern beschreiben. Sie kann grösser oder kleiner sein, aber ist nie gleich.

Andere Schüler lernen schneller als vorgesehen, aber auch sie müssen das Obligatorium erfüllen, allenfalls werden sie früher promoviert. Es ist nicht klar, ob man wirklich neun oder zwölf und mehr Jahre braucht, um die Ziele zu erreichen, gleichwohl ist die Schulzeit auf Zuwachs angelegt; aber nur weil Schulen „Bildungsanstalten“ genannt werden, ist nicht gesagt, dass sie auch wirklich besorgen können, was der Ausdruck „Bildung“ verspricht.

- Das PISA-Paradox - Leseunkundigkeit trotz jahrelangem Schulbesuch - ist durchaus erklärbar.
- Unterricht allein garantiert nicht bereits die Wirksamkeit,
- weil Randfaktoren prägender sein können als die Gestaltung der schulischen Lernsituation und nicht alle Schüler ihre Schulzeit von Anfang bis Ende als stetiges Vorankommen erleben.

Die Institutionen der Bildung kontrollieren sogar zunehmend *weniger* das gesamte Lernfeld, was paradoxerweise Bildung in einer bestimmten Hinsicht befördert hat, denn niemand kann sich mehr mit dem begnügen, was in der Schule gelernt wurde, wie gut es immer gewesen sein mag. Die schulpädagogische Grundidee der „Ausrüstung für das Leben“ ist vom Leben überholt worden, falls sie überhaupt je Realitätsgehalt hatte. In der Schule lernt man nicht alles wirklich perfekt und vieles nur vorläufig, schon aus diesem Grunde verlagern sich Bildungsaufgaben ins Leben, während Schulen mit Qualifizierung befasst sind, die zu realisieren zusehends schwieriger wird.

Ihr Ziel ist nicht Bildung im Sinne der Selbstformung, sondern der Erwerb von Kompetenzen, die in der Breite der Schülerschaft mühsam aufgebaut werden müssen. Schulbildung ist profan, ihre Leistungen zu erfassen, ist kein magischer Begriff von Bildung nötig. Die Differenz zwischen „Bildung“ und „Schulbildung“ hat Vorteile, sie begrenzt die Macht der Schule und sichert auf der anderen Seite ihre Funktion. Der Schule wird ja zugleich alles und nichts zugetraut, je nachdem, welches Hoffnungs- oder Enttäuschungsmanagement gerade betrieben wird. Aber das ist die Folge von Erwartungen, die zu hoch und zu unspezifisch sind, um dem gerecht zu werden, was in den Schulen geschieht und auch nur geschehen kann.

Die öffentliche Schule erfüllt viele der an sie gerichteten Erwartungen *nicht*, weil sie ihre institutionelle Form oder, wie manche Bildungshistoriker sagen, ihre „Grammatik“, nicht beliebig verändern kann. Für viele Ansprüche, also, ist die Schule schlicht der falsche Adressat:

- Sie kann nicht zugleich die Persönlichkeit bilden,
- den Lehrplan erfüllen,
- auf den Kanon der Literatur zurück kommen,
- das naturwissenschaftliche Weltbild befestigen,
- die Erziehungsausfälle der Gesellschaft beseitigen und für die fortlaufende Motivation von Schülern sorgen,
- deren grundlegende Freizeiterfahrungen schulferne und gelegentlich auch schulzynische Kinder- und Jugendkulturen sind.

Hier müssen klare Grenzen definiert werden, die deutlich machen, was Schulen leisten und vor allem, was sie *nicht* leisten. Das hiesse auch, „Schulbildung“ oder die Qualifizierungsleistungen der staatlichen Institution Schule von allen unpassenden Ansprüchen der philosophischen Bildungstheorie zu entlasten.

Aber auch die Schulreform muss sich überprüfen lassen. Jede neue Methode kann mit der Gewöhnung an sie ihren Reiz verlieren und die Nachhaltigkeit ist nicht allein dadurch gegeben, dass jeder formal nach eigenem Tempo lernt. Zudem muss man verschiedene Methoden vergleichen, wenn man wissen will, wie wirksam sie sind. Das gleiche gilt für die bildungspolitische Steuerung und deren Wirksamkeitserwartungen. Es reicht nicht aus, mit einem Plan oder ein Kerncurriculum zu verabschieden und die Umsetzung dann einem Leitfaden zu überlassen. Das gilt umso mehr, wenn sich „Wirksamkeit“ tatsächlich auf die Leistungen der Schülerinnen und Schüler beziehen soll. Voraussetzung dafür ist das Geschehen im Klassenzimmer.

Reformideen kann man danach sortieren, wie wirksam sie sind. Nicht alles, was gut klingt, ist auch wirksam; die pädagogische Sprache erlaubt es, jedes Konzept als „modern“ und „innovativ“ hinzustellen, obwohl es weder neu noch erprobt ist. „Selbstorganisiertes Lernen“ und „altersdurchmisches Lernen“ etwa sind heutige Diskursfavoriten, auf die man schwören kann, ohne dadurch schon zu wissen, ob und wie sie wirken. Das Gleiche gilt natürlich auch für den Ruf nach „mehr Disziplin“. Wer ruft, wird nicht schon gehört. Also, was weiss man über Wirksamkeit?

2. Fragen der Wirksamkeit

John Hattie von der University of Auckland hat im Jahre 2009 die bislang aufwendigste Metaanalyse von internationalen Wirksamkeitsstudien vorgelegt, die sich auf den Bereich Schule und Unterricht beziehen. Das inzwischen auch in Deutschland kontrovers diskutierte Buch trägt den Titel *Visible Learning* (Hattie 2009). Der Titel ist Programm, denn nur die Wirkungen machen sichtbar, was es mit einem pädagogischen Konzept oder einem Curriculum auf sich hat. In der Bildungspolitik ist das eher umgekehrt. Hattie untersuchte über 800 Metanalysen der angelsächsischen Wirksamkeitsforschung, vereinheitlichte die Terminologie und konzentrierte die Daten.

Der Begriff „Wirksamkeit“ bezieht sich einzig auf *achievement*, also die messbaren Leistungen von Schülerinnen und Schülern in bestimmten Fächern. Natürlich geht „Leistung“ nicht in Messbarkeit auf, aber mit Leistungsstudien werden belastbare Daten geliefert, die pädagogischen Streit objektivieren können. Man muss nicht jeweils Partei ergreifen und der einen oder der anderen Seite Recht geben. Die Stärke der Effekte („effect size“) wird von Hattie nicht umgangssprachlich beschrieben, sondern von 138 einzelnen Faktoren aus errechnet, die zu sechs Gruppen zusammengefasst werden.

Hattie unterscheidet: Die Schule als Organisation, den Unterricht, die Curricula der Schule, die Lehrerinnen und Lehrer, die Schülerinnen und Schüler, die Familie und die soziale Herkunft (ebd., S. 31). Vergleicht man diese sechs Gruppen, dann ergibt sich klares und unstrittiges Ergebnis:

- Die grösste Effektstärke kommt den *Lehrpersonen* zu.
- Von ihnen hängt es primär ab, welche Leistungen die Schülerinnen und Schüler zeigen,
- allerdings von ihnen nicht einfach *als* Personen, sondern unter der Voraussetzung eines elaborierten beruflichen Könnens und so eines Feldes von Faktoren, die dafür ausschlagend.
- Man kann daher sagen, nicht auf die Lehrperson kommt es an, sondern auf die *Qualität* der Lehrperson.

Dieses Feld der Qualität von Lehrpersonen umfasst bei Hattie acht Punkte. Die ersten drei werden wie folgt bestimmt: Worauf es ankommt, ist

- Die Qualität des Unterrichts, so wie die *Schülerinnen und Schüler* sie wahrnehmen.
- Die Erwartungen der Lehrpersonen an sich selbst und die Lernenden.
- Die Konzeptionen der Lehrpersonen über Unterricht, Leistungsbeurteilung sowie über die Schülerinnen und Schüler.

Der letzte Punkt bezieht sich auf die Sichtweisen (views) der Lehrerinnen und Lehrer, etwa ob sie glauben, dass alle Schülerinnen und Schüler im Lernen Fortschritte machen und ob sich die Leistungen sich ändern oder stabil bleiben. Ein Problem ist auch, wie der Lernfortschritt von den Lehrpersonen verstanden und artikuliert wird, also wem oder was sie den Fortschritt zuschreiben und wie die Lernenden davon in Kenntnis gesetzt werden. Wenn der entscheidende Faktor etwa in der sozialen Herkunft gesehen wird, dann ziehen Lehrkräfte andere Schlüsse, als wenn sie die Begabung im Vordergrund sehen.

Die fünf weiteren Punkte für den Einfluss der Lehrkräfte auf die Lernleistungen der Schülerinnen und Schüler sehen so aus:

- Die Offenheit der Lehrkräfte oder wie sie darauf eingestellt sind, sich überraschen zu lassen.
- Das sozio-emotionale Klima im Klassenzimmer, wo Fehler und Irrtümer nicht nur toleriert werden, sondern willkommen sind.
- Die Klarheit, mit der die Lehrpersonen Erfolgskriterien und Leistungsanforderungen artikulieren.
- Die Unterstützung der Lernanstrengung.
- Das Engagement aller Schülerinnen und Schüler (ebd., S. 34).

Wenn heute auch in der angelsächsischen Diskussion gesagt wird, „teachers make the difference“, dann ist das zunächst nur ein Mantra. Der *Unterricht* macht den Unterschied, aber unterrichtet wird verschieden und nicht alle Lehrpersonen sind gleich erfolgreich in der Beförderung des Lernens.

Bestimmte Lehrkräfte erfüllen die Aufgaben besser als andere und die kritische Frage ist, bis zu welchem Grad das der Fall ist. Deutlich wird gesagt: Nicht alle Lehrerinnen und Lehrer unterrichten „effektiv“, nicht alle sind „Experten für Lernen“ und nicht alle haben grossen Einfluss auf die Lernenden. Die wichtige Frage ist, in welchem Ausmass sie Einfluss auf die Leistungen haben und was den grössten Unterschied macht“ (ebd.).³

Geht man nicht von einer Faktorengruppe aus, sondern bezieht sich auf einzelne Faktoren, dann sind die drei Faktoren mit der grössten Effektstärke:

- **Self-reported grades:** Die Einschätzung ihres aktuellen Leistungsstandes durch die Schülerinnen und Schüler (ebd., S. 43).
- **Piagetian programs:** Die altersgerechte Gestaltung des Unterrichts nach den Stufen von Piaget (ebd.).
- **Providing formative evaluation of programs:** Die ständige Erhebung des Lernfortschritts und die direkte Rückmeldung an die Schülerinnen und Schüler (ebd., S. 181).

Am anderen Ende der Skala sind Faktoren, deren Effektstärke bezogen auf die Lernleistung so schwach ist, dass sie mit grosser Wahrscheinlichkeit als wirkungslos gelten müssen. Diese Faktoren beziehen sich auf Konzepte, die in der Alternativpädagogik, aber auch in der heutigen Schulreformediskussion, hoch gehandelt werden, weil sie als besonders „kindgemäss“ gelten. Es handelt sich etwa um:

- **Multi-grade/multi-age classes:** Jahrgangsübergreifender Unterricht und altersdurchmischtes Lernen (ebd., S. 91ff.).
- **Student control over learning:** Selbstbestimmtes Lernen (ebd., S. 193/194).
- **Open vs. traditional:** „Offener“ versus „traditioneller“ Unterricht (ebd., S. 88/89).

Vergleicht man die Effekte von offenem und traditionellem Unterricht, dann sind die Unterschiede generell nicht sehr gross. Bezogen auf die Leistungen hat der traditionelle Unterricht leichte Vorteile, offener Unterricht wirkt eher in Bereichen wie Kreativität oder Selbstvertrauen, aber mit beiden lässt sich die Fragen nicht beantworten, was im Blick auf die Wirksamkeit den Unterschied macht (ebd., S. 89). Die Antwort ist nicht das Konzept, sondern die Qualität des Unterrichts. Dogmen der Didaktik, auf die man zu schwören gelernt hat, wirken nicht.

Neben der Wirksamkeit von Methoden und Programmen lassen sich auch Faktoren bestimmen, die sich auf den letztendlichen Schulerfolg beziehen. Die negativsten Faktoren sind dabei Repetitionen von Jahrgängen (**retention**) (ebd., S. 97ff.), hoher Fernsehkonsum (**television**) (ebd., S. 67) und häufige Umzüge der Familie (**mobility**) (ebd., S. 81/82). Im

³ „Not all teachers are effective, not all teachers are experts, and not all teachers have powerful effects on students. The important consideration is the extent to which they do have an influence on students achievements, and what it is that makes the most difference“ (Hattie 2009, S. 34).

ersten Fall verliert man ein Jahr und gewinnt kaum etwas, im zweiten Fall distanziert man sich von schulischen Lernformen, die Anstrengungsbereitschaft verlangen, und im dritten Fall muss Anschluss an ein fremdes System gefunden werden, das den Wechsel nicht belohnt, sondern eher bestraft. Im Übrigen haben auch Sommerferien negative Effekte, weil sie das Vergessen befördern (ebd., S. 81).

Schulaufsicht und Schulleitungen haben mittlere Effekte im Blick auf die Lehrkräfte (ebd., S. 83ff.) und das gilt auch für eine so umstrittene Grösse wie die Klassengrösse (ebd., S. 85ff.).⁴ Die Programme der Lehrerinnen- und Lehrerbildung haben niedrige bis schwache Effekte, weil die Studierenden das Angebot nach dem Vorwissen sortieren, das sie mitbringen und sich so im Studium selbst bestätigen (ebd., S. 109ff.). Ziele und Klarheit der Anforderungen sind Erfolgsfaktoren (ebd., S. 125ff., 163ff.), die direkte Instruktion hat nicht ausgedient (ebd., S. 204ff.) und das „problem-solving teaching“ ist effektiver als das „problem-based learning“ (ebd., S. 210ff.), weil das Problem nicht langwierig herausgefunden werden muss. Schliesslich: Kooperative Lernformen sind wirksamer als kompetitive (ebd., S. 212ff.).

Mit diesen Daten und Rangfolgen der faktoriellen Wirksamkeit lässt sich jedenfalls nicht mehr behaupten, „offene“ Lernformen seien „traditionellen“ grundsätzlich überlegen und es käme darauf an, überall jahrgangsübergreifenden Unterricht zu realisieren, der ja nur die Gruppierung der Schülerinnen und Schüler betrifft und nicht dadurch schon deren Leistung. Auch selbstbestimmtes Lernen, also Formen, in denen die Schülerinnen und Schüler über ihr Lernen selbst verfügen können, hat im Blick auf das Leistungsverhalten gänzlich andere Effekte als die Alternativpädagogik annimmt.

Deutsche Studien zeigen ähnlich, dass offene Lerngelegenheiten im Blick auf individuelle Förderung wenn überhaupt, dann eher im affektiven Lernbereich wirksam sind und kognitive Kompetenzen dadurch kaum beeinflusst werden (Klieme/Warwas 2011). Fördern kann man auf verschiedene Weisen, die aber nie einfach *als Konzept* überlegen sind; auch Formen des Trainings oder adaptiver Unterricht sind im Blick auf das Lernen der Schülerinnen und Schüler nur dann wirksam, wenn sie didaktisch durchdacht und gut strukturiert sind.

Das gilt etwa auch für den Projektunterricht, der vielfach in der Realität, also in der Wahrnehmung der Lehrenden und Lernenden, gar nicht dem entspricht, was in der einschlägigen Literatur angenommen wird, nämlich „selbstgesteuertem Lernen“ (Traub 2011). Es heisst so, aber man macht nicht, was das Konzept versprochen hatte. Oft wird in der Theorie nicht antizipiert, dass sich die Lernleistungen auch im Projektunterricht unterschiedlich verteilen und entsprechend die Profite für die Lernenden keineswegs identisch sind.

- Hattie plädiert für die Normalform eines von der Lehrperson vorbereiteten, strukturierten und realisierten Unterrichts.
- Der Lehrer und die Lehrerin wird nicht in die Rolle eines konstruktivistischen Beobachters versetzt; die Lehrperson darf und soll agieren.

⁴ „It appears that the effects of reducing class size *may* be higher on teacher and student work-related conditions, which *may* or *may not* translate into effects on student learning“ (Hattie 2009, S. 86).

- Für ihn oder für sie als Qualitätsannahme gilt, dass er oder sie „with the eyes of the students“ wahrnehmen kann (Hattie 2009, S. 238) und mit seinem Unterricht ihrem Lernen dienlich ist.
- Das ist nicht einfach dann der Fall, wenn man, wie es heute heisst, „selbstorganisiert“ lernt.

Schon Franz Weinert (1996) hat darauf verwiesen, dass besonders bei anspruchsvollen Aufgaben die kompetente Unterstützung durch die Lehrkraft notwendig ist, damit es zu einem Aufbau systematischen und fehlerfreien Wissens kommen kann. Dahinter steht ein Konflikt, der deutlich benannt werden muss. Selbstreguliertes Lernen kann nicht einfach mit offenem Unterricht gleichgesetzt werden, wie das in der didaktischen Literatur fast regelmässig geschieht. „Instruktion“ steht dann gegen „Konstruktion,“ während der Unterricht in staatlichen Schulen von Aufgaben gesteuert wird, die auf die jeweilige Lernsituation hin angepasst werden müssen.

Resultate der amerikanischen Leseforschung zu Beginn der neunziger Jahre haben zuerst deutlich gemacht haben, dass offener Unterricht im Wesentlichen nur den Kindern aus bildungsnahen Schichten zugutekommt, wohingegen Kinder, die in prekären Verhältnissen aufwachsen und wenig Bildungsanregung erfahren, mit klar strukturiertem Unterricht viel besser zurechtkommen und mehr profitieren. Man tut daher den, wie sie in der amerikanischen Literatur genannt werden, „low-income children“ keinen Gefallen, wenn sie mit offenen Lernsituationen konfrontiert werden, die sie nicht bewältigen können.

Die amerikanischen Untersuchungen verweisen darauf, dass Kinder aus armen Familien in den Anfangsklassen der Regelschule leistungsmässig kaum benachteiligt werden, dann aber zurückfallen, wenn sich die fachlichen Anforderungen des Unterrichts signifikant erhöhen (Chall/Jacobs/Baldwin 1990, S. 167). „However, when classrooms did provide the instruction and the literate environments needed for making the transition to the more mature reading of the intermediate and upper elementary grades, these children learned well. Such classrooms supplied the children with sufficient and appropriate textbooks, reference works, library books, and writing materials“ (ebd., S. 167/168).

Auf der anderen Seite müssen sich die Schülerinnen und Schüler sich als die „Lehrer ihrer selbst“ verstehen und dürfen die Verantwortung für den Lernerfolg nicht abwälzen. Sie dürfen also nicht davon ausgehen, dass sie erst lernen, wenn sie ausreichend motiviert worden sind, was letztlich nur heisst, dass sie die Dosis des Motiviertwordenseins bestimmen. Die Psychologie des Konstruktivismus ist für Hattie eine Form des Wissens und nicht des Unterrichtens (Hattie 2009, S. 243). Wer sich darauf beruft, kommt nie weiter als bis zur Alternativpädagogik (so schon Diesbergen 2000).

Real müssen die Lehrerinnen und Lehrer jeweils herausfinden, was für ihre Klasse oder Lerngruppe je der richtige Mix im Angebot ist. Und die Schülerinnen und Schüler lernen, wie sie gegebenen die Ressourcen nutzen oder eben auch nicht nutzen. Wie sie das tun, ist unterschiedlich und muss vor dem Hintergrund ihrer Schulkarriere verstanden werden. Es gibt, anders gesagt, einen Zusammenhang zwischen Lernleistung und Schulerfolg.

3. *Lernleistung und Schulerfolg*

Der Alltag heutiger Schulen ist gekennzeichnet von Heterogenität, und dies von Anfang an und in vielen Fällen unausweichlich. Die bildungspolitisch vorrangige Frage ist, wie sich neue Methoden oder am Schreibtisch entworfene Curricula auf diesen Tatbestand auswirken, also ob die Risiken des Schulbesuchs dadurch minimiert werden können oder nicht. Schülerzentrierte Methoden tragen nie einen Risikostempel, sie werden in bester Absicht verwendet und stossen aber auf eine harte Wirklichkeit, die den guten Absichten oft im Wege steht. Das Problem lässt sich an einem zentralen Bereich zeigen, nämlich die Entwicklung des Lernstandes über die Schulzeit.

Die Bildungsdirektion des Kantons Zürich hat im Jahre 2003 ein Forschungsprojekt begonnen, das eine Kohorte von anfänglich rund 2.000 Schülerinnen und Schülern während ihrer gesamten Schulzeit beobachtet. Die Schüler wurden bei Schuleintritt getestet und dann nachfolgend am Ende der dritten und sechsten Klasse, ein Test am Ende der neunten Klasse folgt im nächsten Jahr. Auf diese Weise soll die Entwicklung des Lernstandes erfasst werden. Dieser Längsschnitt ist einmalig in der Schweiz, Ansätze dazu gibt es inzwischen auch im deutschsprachigen Ausland. Die PISA-Studien erfassen demgegenüber nur Momente, aber nicht Prozesse, also nicht Veränderungen an verschiedenen Zeitpunkten.

Die Ergebnisse der ersten drei Studien liegen inzwischen vor (Moser/Stamm/Hollenweger 2005; Moser/Keller/Tresch 2008; Moser/Angelone/Keller/Hollenweger/Buf 2010; Nach sechs Jahren Primarschule 2011) und lassen sich so zusammenfassen:

- Bei Schuleintritt haben etwa vier Fünftel der Kinder den Lernstoff der ersten Klasse in Mathematik teilweise bewältigt und ein knappes Fünftel befindet sich leistungsmässig bereits in der zweiten Klasse.
- Die Lesekompetenzen liegen weiter auseinander. 5 Prozent der Kinder kennen keine Buchstaben, etwa 35 Prozent erfassen die Verbindung zwischen Buchstaben und Lauten nur im Ansatz, 15 Prozent lesen bereits schwierige Wörter und längere Sätze.

Noch disparater ist der Wortschatz, also die Fähigkeit, Bilder mit eigenen Wörtern richtig zu bezeichnen.

- Nur knapp ein Drittel der Kinder beherrscht die richtigen Bezeichnungen,
- ein weiteres Drittel überbrückt geschickt die Lücken,
- das letzte Drittel beherrscht nur ganz elementare Bezeichnungen, also kann sagen, dass ein Hammer auf einem Bild ein „Hammer“ ist, mehr jedoch nicht.
- Korrekte Bezeichnungen für andere Bilder können nicht abgerufen werden.

Die zweite Studie nach drei Jahren Schulzeit zeigt, dass der Unterricht starke, aber auch ungleiche Effekte hat. Die Schülerinnen und Schüler erreichen wie gesagt die erste Klasse der Volksschule mit grossen Unterschieden im Lernstand.

- Unterschiede konnten in den ersten drei Jahren ausgeglichen werden, sofern qualitativ guter Unterricht stattgefunden hat.
- Er macht den Unterschied, was mit der Kompetenz der Lehrkräfte, der Zusammensetzung der Klasse, den je erreichten Lernfortschritten und der Unterstützung der Eltern zu tun hat.

Die dritte Studie nach sechs Jahren Schulzeit, die vor zwei Jahren veröffentlicht worden ist, zeigt, dass die Schülerinnen und Schüler mit Deutsch als Zweitsprache, die ihre Schulzeit mit einem geringen Wortschatz in der Unterrichtssprache begonnen haben, diesen Rückstand inzwischen aufholen konnten. Die Erstsprache ist kein Hindernis mehr, das Verhalten wird gesteuert durch die Akzeptanz der Schulkultur einschliesslich des Unterrichts in Deutsch als Zweitsprache von der ersten Klasse an. Auch ein altersbedingter Vorsprung schwindet im Laufe der Primarschulzeit.

Anders steht es mit Wissensvorteilen.

- Wer schon bei Beginn der Schulzeit über ein grosses Vorwissen verfügt hat, profitiert davon über die gesamten sechs Schuljahre.
- Noch besser lassen sich die Leistungen am Ende der Primarschule mit den Leistungen am Ende der dritten Klasse vergleichen.
- Wer hier gut war, bleibt gut - und umgekehrt.

Es gibt deutliche Risikogruppen, solche Kinder nämlich, die die Lernziele nicht erreicht haben und mit hoher Wahrscheinlichkeit im weiteren Verlauf ihrer Schulkarriere weiter zurückbleiben. Der bildungspolitische Slogan „No Child Left Behind“ erweist sich so einfach als Zweckoptimismus. Die Zürcher Studie wirft die Frage auf, wie mit der Risikogruppe umgegangen werden soll, und das verweist auf die Notwendigkeit einer konsequenten Leistungsentwicklung, für die der andere Slogan „Fördern und Fordern“ zur Verfügung steht. Der Slogan selbst hat keinen Inhalt und besagt eigentlich nur, dass irgendwie häufiger verfahren werden soll. Die wirklichen Probleme sind deswegen schwer zu bearbeiten, weil sie in den Klassen entstehen und trotz erheblichem Einsatz der Lehrkräfte auftreten.

Der Abstand zwischen den leistungsstarken und den leistungsschwachen Schülern muss im Verlauf der Schulzeit keinesfalls geringer werden, eher ist anzunehmen, dass er mit den fachlichen Anforderungen grösser wird. Bestimmte Gruppen von Schülerinnen und Schülern verbessern ihren Lernstand nicht, sondern bleiben auf dem einmal erreichten Niveau stehen.

- Die dritte Studie zeigt, dass die Zunahme der Leistungsunterschiede erst auf der Mittelstufe erfolgt, also bei sechs Jahren Primarschule nach der dritten Klasse.
- Bis dahin liegt der Anteil der Schülerinnen und Schüler, die die Ziele des Lehrplans in Deutsch und Mathematik nur teilweise erfüllen, bei 10 Prozent.
- Am Ende der sechsten Klasse sind das 17 Prozent für Deutsch und 18 Prozent für Mathematik, was deutlich eine Folge ist der unausweichlich steigenden Anforderungen.

Für die letzten Schuljahre ist eine weitere Öffnung der Leistungsschere zu erwarten. Für die leistungsschwächeren Schülerinnen und Schüler fehlen dann lohnende Aufgaben und Lernanlässe, die mehr sein müssen als die immer neue Bestätigung ihrer Schwächen. Dieses Problem der negativen Differenzierung über die Schulzeit lässt sich nicht einfach durch punktuelle Fördermassnahmen bearbeiten, sondern stellt sich im Blick auf den gesamten Verlauf einer Karriere als Schülerin und Schüler.

Wie schwer das ist, zeigen weitere Ergebnisse der dritten Studie: Die soziale Zusammensetzung der Klasse wirkt sich dann auf die Leistungen positiv aus, wenn die soziale Herkunft besonders hoch und so besonders homogen ist. Das sind 14 Prozent der Klassen im Kanton Zürich. Die Leistungsbeurteilung durch die Lehrkräfte kann diesen Vorteil durch strengere Bewertungsmaßstäbe in eine Benachteiligung verwandeln, was bei sozial belasteten Klassen umgekehrt gilt. Aber das ist eine heikle Strategie, die einer Quotenregelung gleichkommt.

Alle drei Studien verweisen darauf, dass die Entwicklung des Lernstandes zwischen verschiedenen Klassen auch in derselben Schule höchst unterschiedlich sein kann. Das gilt ebenso für die Effekte der Förderung. In den einen Klassen gelingt es, die Leistungsschere zu verringern, in anderen jedoch nicht, was primär damit zu tun hat, dass sehr verschieden unterrichtet und gefördert wird. Verbunden damit ist allerdings auch ein strukturelles Problem, das deutlich benannt werden muss.

- Fortlaufender Unterricht, der in einem bestimmten Zeitraum und ohne zusätzliche Ressourcen Ziele erreichen soll, die für alle gleich gelten, kann nur sehr begrenzt Nachteile einzelner Schüler ausgleichen.
- Kinder bleiben auch deswegen zurück, weil erreichbare Ziele für sie gar nicht bestehen.

Herkunftsbedingte Ungleichheit können Schulen nicht einfach kurzfristig ausgleichen. Die Beherrschung der Unterrichtssprache dagegen kann durch gezielte und frühe Förderung gelernt werden. Daher ist eine konsequente Sprachförderung ein zentrales Mittel zur Gewährleistung von Erfolgen in der Schule. Das gilt ganz generell und besonders im Blick auf die zentralen Leistungsfächer Deutsch und Mathematik. Notwendig ist dafür auch eine gezielte Elternarbeit. Neuere deutsche Studien zeigen, dass generell mit der Dauer der Schulzeit die Motivation der Schülerinnen und Schüler abnimmt. Diese Tendenz wird vor allem dort gemindert, wo es einen vollumfänglichen Ganztagsbetrieb gibt. Das spricht dafür, mit der Ganztagsbeschulung weiter fortzufahren und sie auszubauen (Nachweise in Oelkers 2011).

Allgemein gesagt: Schulerfolg hängt wesentlich ab von der Beherrschung der Unterrichtssprache. Wer dem Unterricht sprachlich nicht folgen kann, hat in fast allen Unterrichtsfächern Nachteile. Auch die soziale Herkunft oder die Nähe bzw. Ferne zu den Bildungsanforderungen der Schule spielen eine wichtige Rolle. Wenn beides negativ zusammenkommt, ist trotz aller Anstrengungen die Wahrscheinlichkeit hoch, dass der Erfolg im Leistungsbereich ausbleibt. Alle Methoden müssen sich dieser Frage stellen und dürfen nicht lediglich als persönliche Vorlieben der Lehrkräfte Verwendung finden. Das gilt ebenso für Lehrpläne und Lehrmittel. Wie soll etwa selbstreguliertes Lernen wirksam sein, wenn es für die leistungsschwächeren Schülerinnen und Schüler keine geeigneten Lehrmittel gibt? Aber helfen da nicht „Standards“?

4. Ein Blick auf „Standards“

Jede bildungspolitische Reform beginnt mit der Rhetorik der Verheissungen. Dabei weiss man, dass es gute Reformideen gibt, die nie verwirklicht werden, und schlechte, die die Praxis sehr schnell erreichen. Schlechte Reformideen treten meistens gebündelt auf und haben

Belastungsfolgen, nicht für die, die die Idee in die Welt gesetzt haben, sondern für die, die sie umsetzen müssen. Genauer sollte ich sagen, dass es eigentlich gar keine schlechten Bildungsreformen gibt, weil ja jede eine grosse Verheissung ist und im besten Licht dargestellt wird. Belastungen scheint es keine zu geben. Aber nach zehn Jahren Reform Erfahrung sind in der Schweiz deutliche Belastungsfolgen erkennbar, die klar benannt werden müssen.

Die Bruchstellen zeigen sich in Schweizer Belastungsstudien (Albisser et.al. 2006; Nido et.al. 2008; Windling et. al. 2011) deutlich. Fragt man die Lehrkräfte und nimmt ernst, was sie sagen, dann lassen sich folgende Belastungsfaktoren bestimmen:

- Die kaum durchschaubaren Reformwellen mit ihren Erlassfolgen,
- die unerreichbaren Zielsetzungen oder die Diskrepanz zwischen Rhetorik und Praxis,
- die geringe Unterstützung bei der täglichen Arbeit,
- die verschiedenen Formen der Rechenschaftslegung, also interne und externe Evaluationen,
- und nicht zuletzt die Veränderung der Schülerschaft.

Der Aufwand, insbesondere an schriftlichen Stellungnahmen und Äusserungen, ist rasant gestiegen, ohne dass damit alleine eine Verbesserung des Ertrages erreicht worden wäre. In der Schweiz setzen sich Schulen gegen unnötige Belastungen zur Wehr und bilden im Gegenzug eigene Prioritäten und verfolgen Ziele, die tatsächlich erreichbar sind und zu erkennbaren Verbesserungen führen.

Das Problem entsteht bei der Umsetzung von Reformen oder dort, wo neuerdings immer das Wort „Implementation“ auftaucht und schon durch die Wortwahl technische Erwartungen erweckt werden, als könnte man, wie die Schweizer sagen, an den Schulen „herumschrauben“ und damit Erfolg haben. Sie entnehmen meiner Wortwahl, dass ich dieser Meinung nicht bin. Der Begriff „Implementation“ hat Karriere gemacht, seitdem man einen Algorithmus in ein Computerprogramm umsetzen kann. Auch der Ausdruck „Bildungsstandards,“ der in den achtziger Jahren in den Vereinigten Staaten geprägt wurde, unterstellt eine technische Normierung, so wie sie in der Industrie üblich ist.

Schulischer Unterricht ist aber ersichtlich nicht mit einer Industrienorm erfassbar, wenn die so definiert ist, dass ein bestimmtes Format an jedem Ort seiner Anwendung gleich sein muss.

- Unterricht ist Interaktion mit ungleichem Verlauf und Ausgang; eine technische Norm wie zum Beispiel das DIN-A-4 Seitenformat kann damit nicht in Verbindung gebracht werden.
- Trotzdem erweckt der Ausdruck „Standard“ den Eindruck einer Normierung, die die Widrigkeiten des Unterrichts überspringen und gleichsam direkt für den gewünschten Effekt sorgen könnte.

Ein weiterer Modebegriff ist „Kompetenz“, der inzwischen für alles herhalten muss, was irgendwie innovativ klingt und doch selten das traditionelle Begriffspaar „Wissen und Können“ übersteigt. Niemand musste hier lange nach „Strategien der Implementation“ suchen, es genügte einfach die Ausbreitung der Expertensprache, die heute bis in die Buchtitel

der Fachliteratur und die Lehrpläne der Schule vorgedrungen ist. Aber wenn man etwas anders bezeichnet, ist es nicht schon anders. Praktische Innovation ist weit schwieriger zu bewerkstelligen als semantischer Austausch und der Wandel der Rhetorik von Experten, mit dem oft nur Eindrucksmanagement verbunden ist. Aber was ist damit gewonnen, wenn alles, was als Zielformel gut klingen soll, „Kompetenz“ genannt wird?

Keine didaktische Theorie seit dem 19. Jahrhundert hat Unterrichtserfolg mit dem blossen Nachvollzug von Schulwissen gleichgesetzt; es ging immer um das Verstehen, Durcharbeiten und Anwenden des Stoffes durch die Schüler, also um zunehmendes Können oder um stetig verbesserte Kompetenz.

- Der Gründer des Pädagogischen Seminars der Universität Leipzig, der Herbartianer Tuiskon Ziller (1884, S. 240),⁵ hatte dafür auch eine plausible Regel:
- Der Zögling, wie man die Schüler im 19. Jahrhundert nannte, „darf durch den Unterricht nicht geistig schwächer werden“.⁶

Wie oft das der Fall war, ist nie untersucht worden, aber das Problem der Nachhaltigkeit des Unterrichts ist nicht neu, und auch Zillers Lösung, dass die Schulübel verschwinden, „wenn nur die Lehrer bessere Methoden zu gebrauchen wüssten“ (ebd., S. 243),⁷ ist bis heute angesagt. Sie alle kennen der „Klippert-Effekt“, also das Nachbessern mit Methoden, die eigentlich die Ausbildung hätte vermitteln sollen.

Unterricht und fortlaufende Leistungsüberprüfung machen das Kerngeschäft der Schule aus. In Schweizer Gymnasien gibt es im Blick auf die Prüfungspraxis wohl formal sehr genau festgelegte Verfahren, etwa bei der Verrechnung der einzelnen Noten, aber was die Noten inhaltlich bedeuten, ist weitgehend offen und wird von jeder Lehrkraft selbst festgelegt. Es gehört zur Berufsidentität von Lehrerinnen und Lehrern, über die Leistungen der Schüler in ihrem Fach autonom und nach eigenem Urteil bestimmen zu können. Absprachen über die Noten oder gar eine genaue Beschreibung der Notenstufen sind bislang weitgehend unbekannt, was sich nicht allein für die Gymnasien sagen lässt. Prüfen nach *gemeinsamen* Standards für die Noten gibt es in den wenigsten Schulen. Aber das wäre der Kernstandard jeder Schule, der heute weitgehend unthematisiert bleibt.

In der Prüfungspraxis sind bis heute Noten zentral, also die Einschätzung der Leistungen von Schülerinnen und Schüler auf einer für alle Lehrkräfte verbindlichen Skala, die die Unterschiede von Fähigkeiten in Sachgebieten erfassen soll. Trotz einer Vielzahl von scheinbaren oder tatsächlichen Alternativen ist die Notenskala das bei weitem gebräuchlichste Instrument der Leistungsbeurteilung. Für dieses Instrument spricht, dass Leistungen in Schulklassen tatsächlich immer mit Niveauunterschieden zustande kommen. Wer sie abbilden will, muss daher ein gestuftes Schema verwenden, wobei das Problem nur ist, welche Stufen

⁵ Tuiskon Ziller (1817-1882) war Gymnasiallehrer und habilitiert sich als Jurist. 1854 begann er mit Vorlesungen im Pädagogik, 1861 erfolgte die Gründung des Pädagogischen Seminars, das ein Jahr später mit einer Übungsschule verbunden wurde. Zillers *Einleitung in die Allgemeine Pädagogik* von 1856 ist die Begründungsschrift des Pädagogischen Herbartianismus.

⁶ Sperrung im Zitat entfällt.

⁷ Sperrung im Zitat entfällt.

zur Anwendung kommen und wie die tatsächlichen Leistungsunterschiede in der Beurteilung abgebildet werden.

Das Prinzip der Stufung wird auch in den Bildungsstandards für die Schweizer Volksschule deutlich. Für sie wurden Kompetenzmodelle entwickelt, die davon ausgehen, dass sich die Leistungen auf verschiedenen Stufen erfassen lassen. Die Erreichbarkeit der Leistungsziele ist je nach Aufwand und Talent unterschiedlich. Das unterscheidet die neuen Aufgabekulturen von den klassischen Lehrplänen, die so etwas wie „Regelstandards“ beschrieben haben.

- Solche Vorgaben sollten *alle* Schülerinnen und Schüler erreichen können, während tatsächlich immer Leistungsabstände das Verhalten bestimmen.
- Aus diesem Grunde gibt es in der Schweiz „Mindeststandards“ und nicht wie in Deutschland Regelstandards.
- Alle Schüler müssen eine *bestimmte* Stufe erreichen, nicht den gesamten Lehrplan, was weder möglich noch sinnvoll wäre.

Wenn es Zukunft auch in der Schweiz vergleichende Leistungstests geben wird und verbindliche Standards für die Volksschule eingeführt werden, dann stärkt das in gewisser Hinsicht den Primat des Fachunterrichts. Auch der Kompetenzaufbau wird verbessert, weil und soweit die Leistungsniveaus getestet werden. Das läuft der Alternativpädagogik zuwider, die auf jeglichen Leistungsvergleich am liebsten verzichten würde. Aber jedes Kind vergleicht sich in seinem Leistungsverhalten mit anderen Kindern.

Viele Bildungssysteme verfügen heute über Standards, Leistungstests und komplexere Formen der Evaluation. Das elaborierteste System in Europa ist seit 1988 in England aufgebaut worden, und selbst in den liberalen Gesamtschulsystemen Skandinaviens werden inzwischen Standards und Tests verwendet. Es gibt also überall mehr als nur die gewohnte Prüfungspraxis in den einzelnen Schulen. Die Frage ist, wie diese Praxis so verändert werden kann, dass zur Gesamtentwicklung passt und zugleich den einzelnen Schulen sichtbare Vorteile bringt. Neue Formen der Beurteilung, anders gesagt, müssen einen Beitrag zur Unterrichtsentwicklung leisten, sonst werden sie nicht akzeptiert.

Für die Schweizer Schulen sind Leistungsstandards nichts Neues, neu wären externe Leistungstests, die derzeit aber nicht abgesehen werden können. Sie würden die internen Prüfungen auch nicht ersetzen, sondern allenfalls ergänzen. An flächendeckende Tests ist derzeit nicht gedacht, und auch ein „High-Stakes-Testing,“ also Tests mit Folgen für die Schule oder gar für die einzelnen Lehrkräfte,⁸ steht nicht zur Diskussion. Oft werden Leistungstest so verstanden, als würden sie wesentlich nur in dieser Funktion verwendet, was jedoch nicht der Fall ist.

Wenn die Prüfungspraxis optimiert werden soll, das muss die Schule das zu einem vorrangigen Entwicklungsziel erklären und die Implementation zur gemeinsamen Sache der Akteure machen. Das verlangt Führungsarbeit seitens der Schulleitung, das System selbst wird von den Lehrkräften ausgearbeitet, die Notwendigkeit der Entwicklungsarbeit und der

⁸ Einzelne Schuldistrikte in den Vereinigten Staaten, etwa in Cincinnati, koppeln die Löhne an die Ergebnisse der Leistungstest.

Sinn des neuen Systems muss mit allen Akteuren kommuniziert werden. Die Eltern sowie die Schülerinnen und Schüler gewinnen dann einen Einblick, nach welchen Kriterien benotet und geprüft wird. Sie können sich dann darauf einstellen, was vor allem für die Akzeptanz der Noten von grosser Bedeutung wäre.

In einem transparenten System der Leistungsbeurteilung wären Noten gestufte Formate mit einem gemeinsamen Massstab. Die Notengebung würde sich nach Kriterien vollziehen, die für alle Beteiligten klar sind und die fortlaufend kommuniziert werden. Dann gibt nicht jede Lehrkraft „ihre“ Noten, vielmehr werden die Notenstufen pro Fach oder Fachbereich festgelegt und auf das eingestellt, was als Leistungsstandard gilt. Standards unterscheiden sich nach Stufen und darauf beziehen sich die Noten. Die einzelnen Stufen erhalten eine schriftliche Fassung, die den Unterschied beschreibt. Sprachlich möglich ist das durch Graduierungen, die zu den Stufen passen.

Ein Beispiel für die Benotung einer längeren Klausur aus dem Universitätsbereich sähe so aus:

Noten:

- *Exzellent* (6.0): eigenständige, umfangreiche, zutreffende und mit herausragenden weiterführenden eigenen Reflexionen angereicherte Darstellung.
- *Sehr gut* (5.5.): eigenständige, zutreffende und weiterführende Darstellung.
- *Gut* (5.0): Zutreffende und weiterführende Darstellung.
- *Befriedigend* (4.5.): Zutreffende Darstellung mit leichten Mängeln.
- *Ausreichend* (4.0): Zutreffende Darstellung mit Mängeln.
- *Nicht ausreichend* (unter 4.0): Darstellung mit gravierenden Mängeln.

Das auszuhandeln und festzulegen wäre in den Schulen Aufgabe der einzelnen Fachschaften, wobei die Schulleitung ein allgemeines Format für die Beschreibung auch vorgeben kann. Die Anpassung muss auf die Besonderheiten der Fächer Rücksicht nehmen. Entscheidend ist, dass transparent gemacht wird, wann welche Note vergeben wird und was sie für ihren Referenzraum bedeutet. Ein zunehmend wichtige Frage geht dahin, welche Verantwortung die Schule dabei übernimmt und welche Rolle bei der Bewertung von Leistungen Tests spielen werden.

5. Die Leistung von Leistungstests

Ein zentrales Stichwort lautet „accountability,“ die Verantwortung der Schule für das Zustandekommen und die Bewertung der Leistungen ihrer Schüler. Natürlich hängt der Erfolg des Unterrichts auch davon ab, wie die Schülerinnen und Schüler ihre Lernressourcen einsetzen und wie sie das schulische Angebot nutzen. In diesem Sinne ist die Verantwortung immer geteilt; wenn Schüler eine minimalistische Strategie der Leistungserfüllung wählen, ist das ihre Entscheidung und nicht die der Schule. Sie muss nur klar machen, was die Konsequenzen einer solchen Lernstrategie sind, nämlich schlechtere Bewertungen und gegebenenfalls das Nichterreichen des Leistungsziels.

In diesem Sinne sind „eigenverantwortete Schulen“ immer solche, die die Verantwortung teilen und zwischen den verschiedenen Akteuren festlegen. Nicht teilbar ist die Verantwortung der Schule für die Leistungsziele und die Art und Weise der Leistungsüberprüfung. Beides zusammen stellt das Angebot der Schule dar, das möglichst transparent dargestellt und kommuniziert werden muss. Grundlegend sind klare Ziele und durchschaubare Kriterien, die eindeutig beschrieben werden und so auch Geltung finden. Die Schulen müssen sich auf Ziele beziehen, die sie auch erreichen können, und sie müssen ihre Entwicklungsprozesse regelmässig bilanzieren, damit sie wissen, wo sie stehen.

Ein Projekt über den Nutzen von Leistungstests heisst „Check Five“ und ist im Schweizer Kanton Aargau durchgeführt worden. Der Kanton Aargau ist einer der letzten Kantone der Schweiz, in dem der Übergang zu den Sekundarschulen nach der fünften und nicht nach der sechsten Klasse erfolgt. „Check Five“ ist ein vergleichender Leistungstest, der im Verlaufe des fünften Schuljahres durchgeführt wurde. Das Projekt ist vom kantonalen Parlament in Auftrag gegeben worden und hat den Zweck, ein Instrument zu entwickeln, mit dem sich die Ergebnisse von Evaluationsstudien an die Lehrkräfte rückvermitteln lassen. Wenn derartig aufwändige Studien wie die PISA-Tests auch einen praktischen und nicht bloss einen politischen oder wissenschaftlichen Wert haben sollen, dann ist die Frage unausweichlich, was die Lehrkräfte mit solchen Daten anfangen können. „Check Five“ ist darauf eine Antwort.

Der Test wurde im vergangenen Jahr durchgeführt und betraf vier Dimensionen, nämlich Mathematik, Deutsch, kooperatives Problemlösen und selbst reguliertes Lernen. Weil hinter dem Test eine politische Absicht steht, rechnete das Projektteam mit erheblichen Widerständen in der Lehrerschaft und wäre zufrieden gewesen, wenn sich 40 Klassen gefunden hätten. Von den 370 Lehrpersonen, die im Kanton Aargau in der fünften Klasse unterrichten, meldeten sich 262, der Test wurde schliesslich mit 140 Klassen und 2.531 Schülern durchgeführt.

Die Akzeptanz solcher Verfahren ist grundsätzlich offenbar vorhanden. Die Lehrkräfte wurden vor Durchführung des Tests befragt, mit welchen Einstellungen und Erwartungen sie an das Projekt herangehen, für das sie sich freiwillig gemeldet haben. Jede einzelne Lehrkraft erhielt zwei Monate nach dem Test eine Rückmeldung, die zeigt, wie ihre Klasse im Vergleich zum Gesamtergebnis abgeschnitten hat. Die Ergebnisse mussten mit den Schülern in einer frei gewählten Form kommuniziert werden. Dabei konnten die Lehrkräfte auch die Eltern einbeziehen.

Nach der Rückmeldung der Ergebnisse erfolgten noch zwei weitere Befragungen, die die Umsetzung erhoben. Sie hatten vier hauptsächliche Ziele:

- Wie gehen die Lehrkräfte mit den Daten ihrer Klasse um, zumal dann, wenn sie nicht gut sind?
- Wie analysieren sie die Ursachen?
- Wie reflektieren sie die Qualität ihres Unterrichts im Lichte der Daten?
- Und welche Massnahmen ergreifen sie zur Verbesserung der Qualität?

Das Projekt ist abgeschlossen und die Ergebnisse liegen vor. Zentral sind wiederum die grossen Leistungsunterschiede zwischen den Schülerinnen und Schülern, aber auch zwischen den einzelnen Klassen.

- Ein Teil der Schüler konnte nahezu alle Mathematikaufgaben richtig lösen, ein anderer Teil löste nach fünf Jahren Unterricht nur zwischen 10 und 20 Prozent der Aufgaben richtig.
- Diese Schüler sind in ihrer fachlichen Entwicklung stehen geblieben.
- Zwischen den Klassen waren die Unterschiede am grössten in den Bereichen Grammatik und Arithmetik, also klar strukturierten Lernfeldern, die offenbar höchst verschiedenen unterrichtet werden.

Weitaus geringer waren die Unterschiede im Textverständnis, auch weil das nicht nur im Deutschunterricht abverlangt wird. Die Unterschiede lassen sich nicht mit der sozialen Herkunft erklären,⁹ sondern sind hausgemacht.

Die Lehrkräfte sind umso mehr aufgefordert zu reagieren. Die Studie zeigt, dass die Testergebnisse mit den Eltern sowie den Schülerinnen und Schülern zum Teil aufwändig kommuniziert wurden. Die Lehrkräfte geben an, mit den Ergebnissen die fachlichen Leistungen ihrer Schüler besser einschätzen zu können. Sie ergänzen damit das klasseninterne Bezugssystem und kommen zu objektiveren Bewertungen, die auch den Übergang verbessern können. Hunderte von Massnahmen zur Sicherung der Unterrichtsqualität wurden vorgelegt, von denen mehr als die Hälfte auch umgesetzt wurde. 131 der 140 Lehrpersonen würden ein zweites Mal an einem solchen Test teilnehmen. Die Akzeptanz von Leistungstests steigt mit der positiven Erfahrung der Verwendbarkeit der Resultate. Akademische Vorbehalte teilen die Lehrkräfte nicht (Tresch/Moser 2005; Tresch 2007)). Sie nutzen, wenn man ihnen eine Handreichung zum Umgang mit Ergebnissen von Leistungstests anbietet (Departement Bildung, Kultur und Sport 2005).

Das Verfahren stösst auf hohe Zustimmung auch bei den Schülern, was durch andere Studien bestätigt wird. Objektivierete, transparente Verfahren der Leistungsmessung sind keine Horrorerfahrungen, sondern wirksame Instrumente, sofern sich die Daten ins Feld rückübersetzen lassen und die Lehrkräfte erreichen. Die Forschung dient der Praxis und besteht nicht aus Belehrungen. Die Akteure werden ernst genommen und die Theorie behauptet nicht, mehr zu wissen als die Praxis. Die Probleme werden aufbereitet und in Form von verständlichen Beschreibungen zurückgespielt. Die Lösungen vor Ort werden durch eine bessere Sicht auf die Probleme angereizt und nicht etwa abgestossen. Diese Praxisnähe der Forschung ist im Übrigen nicht nur in der Schweiz möglich.

Ein Projekt, das auf die Situation am Ende der obligatorischen Schulzeit reagieren soll, heisst „Stellwerk“. In einem Stellwerk werden Weichen gestellt. Analog dieser Metapher werden alle Schülerinnen und Schüler Mitte der 8. Klasse getestet, über welche Kompetenzen in zentralen Fächern sie tatsächlich verfügen. Der Test ist die Grundlage für gezielte Nachbesserungen während der verbleibenden Schulzeit.

- Die Schüler erhalten eine objektivierete Rückmeldung, wie ihr Lernstand beschaffen ist und wo sie stehen.
- Sie können fehlende Kompetenzen aufholen, ohne dass der Unterricht auf den Lehrplan Rücksicht nehmen muss.

⁹ Die Segregation der Schülerpopulation nach bildungsrelevanten Merkmalen im Kanton Aarau ist weit geringer als etwa im Kanton Zürich (Tresch/Moser 2005, S. 68).

- Das Ziel ist, die Anforderungen zu erfüllen, die heute an qualifizierte Lehrberufe gestellt werden.
- Die öffentliche Schule ist dafür das Stellwerk: Sie ist verantwortlich, dass Lücken geschlossen werden und am Ende grundlegende Kompetenzen vorhanden sind.

Je nach Resultat entscheiden die Lehrkräfte zusammen mit den Schülern und den Eltern, wie und in welchen Bereichen die Förderung ansetzen soll. Dieses individuelle, auf den tatsächlichen Bedarf hin zugeschnittene Lernprogramm ersetzt in Teilen den Lehrplan, der zugunsten von gezielten Nachbesserungen in bestimmten Kompetenzbereichen gelockert oder suspendiert wird. Ein solches Verfahren, bei dem die Schüler sehen, wo sie stehen und wo ihre Lücken sind, die nicht verhandelt werden können, ist weit wirksamer als Ermahnungen der Lehrkräfte. Oft glauben die Jugendlichen nicht, was nach der Schule auf sie zukommt, oft überschätzen sie sich auch oder sehen nicht, wie sie sich verbessern können. Dazu brauchen sie eine glaubwürdige, unzweideutige Rückmeldung sowie ein auf sie zugeschnittenes Lernprogramm, das in der gegebenen Zeit bewältigt werden kann.

Es ist illusorisch zu erwarten, dass die Schülerinnen und Schüler bis Mitte der 8. Klasse einen angeglichenen Leistungsstand erreichen werden, dagegen spricht die Leistungshierarchie; aber die Schule kann mit gezielten Förderprogrammen die Chancen auf dem Arbeitsmarkt verbessern und so wirksam auf die Erwartungen der Abnehmer eingehen. Es ist gleichermassen illusorisch, dass die Firmen nachbessern, was bei Beginn der Berufsausbildung fehlt. Die tatsächlich erreichte Kompetenz in den allgemeinbildenden Fächern am Ende der Schulzeit ist der Schlüssel zum Lehrvertrag oder für den weiteren Schulbesuch und so zur Chancennutzung. Mehr können Schulen nicht tun, aber das sollten sie tun.

Literatur

Albisser, St./Kirchhoff, E./Meier, A./Grob, A.: Anforderungsverarbeitung und Gesundheit im Berufszyklus von Lehrpersonen. Paper vorgestellt auf der Tagung „Balancieren im Lehrberuf“. Unveröffentlichtes Manuskript. Bern 2006.

Chall, J.S./Jacobs, V.A./Baldwin, L.E.: The Reading Crisis. Why Poor Children Fall Behind, Cambridge/Mass., London: Harvard University Press 1991.

Department Bildung, Kultur und Sport des Kantons Aargau (BKS) (Hrsg.): Leistungen messen und beurteilen. Handreichungen zum Umgang mit Ergebnissen von Leistungstests. Aarau: BKS 2005.

Diesbergen, C.: Radikal-konstruktivistische Pädagogik als problematische Konstruktion. Eine Studie zum Radikalen Konstruktivismus und seiner Anwendung in der Pädagogik. 2. Aufl. Bern et. al.: Peter Lang Verlag 2000. (= Explorationen. Studien zur Erziehungswissenschaft, hrsg. v. J. Oelkers, Band 22)

Hattie, J.: Visible Learning. A Synthesis of Over 800 Meta-Analyses Relating to Achievement. London/New York: Routledge 2009.

Horlacher, R.: Bildungstheorie vor der Bildungstheorie. Die Shaftesbury-Rezeption in Deutschland und der Schweiz im 18. Jahrhundert. Würzburg: Königshausen+Neumann 2004.

Humboldt, W. v.: Werke in fünf Bänden, hrsg. v. A. Flitner/K. Giel, Bd. I: Schriften zur Anthropologie und Geschichte. 3. Aufl. Darmstadt 1980.

- Hürlimann, W.: „Es wäre ein grosses und treffliches Werk...“ Sprangers Leistung zur vermeintlichen Bildungstheorie Humboldts. Lizentiatsarbeit Universität Zürich, Pädagogisches Institut (Fachbereich Allgemeine Pädagogik). Ms. Zürich 2003.
- Klieme, E./Warwas, J.: Konzepte der individuellen Förderung. In: Zeitschrift für Pädagogik Beiheft (2011).
- Moser, U./ Stamm, M./Hollenweger, J.: Für die Schule bereit? Lesen, Wortschatz, und soziale Kompetenzen bei Schuleintritt. Oberentfelden: Sauerländer Verlage AG 2005.
- Moser, U./Keller, F./ Tresch. S.: Schullaufbahn und Leistung. Bildungserfolg und Lernverlauf von Zürcher Schülerinnen und Schülern am Ende der dritten Volksschulklasse. Bern: h.e.p. Verlag 2008.
- Moser, U./Angelone, D./Keller, F./Hollenweger, J./Buff, A.: Lernstandserhebung am Ende der 6. Klasse. Schlussbericht zuhanden der Bildungsdirektion. Unveröff. Ms. Zürich: Institut für Bildungsevaluation/Pädagogische Hochschule Zürich 2010.
- Nach sechs Jahren Primarschule. Lernstand der Schulanfängerinnen und Schulanfänger von 2003 vor ihrem Übertritt in die Sekundarstufe I. Zürich: Bildungsdirektion Kanton Zürich 2011.
- Nido, M./Ackermann, K./Ulich, E./Trachlser, E./Brüggen, S.: Arbeitsbedingungen, Belastungen und Ressourcen von Lehrpersonen und Schulleitungen im Kanton Aargau 2008. Ergebnisse der Untersuchung im Auftrag des Departements Bildung, Kultur und Sport (BKS) des Kantons Aargau. Zürich: Institut für Arbeitsforschung und Organisationsberatung 2008.
- Oelkers, J.: Expertise zum Thema: „Ganztagsschule“ in der Ausbildung der Professionen. Zuhanden der deutschen Kinder- und Jugendstiftung. Zürich: Institut für Erziehungswissenschaft 2011.
- Schmid, Chr.: Lernen und Transfer: Kritik der didaktischen Steuerung. Bern: h-e-p Verlag 2006.
- Spörer, N./Brunstein, J.C.: Erfassung selbstregulierten Lernens mit Selbstberichtsverfahren. Ein Überblick zum Stand der Forschung. In: Zeitschrift für Pädagogische Psychologie 20, 3 (2006), S. 147-160.
- Traub, S.: Selbstgesteuert lernen im Projekt? Anspruch an Projektunterricht und dessen Bewertung aus Sicht von Lehrenden und Lernenden. In: Zeitschrift für Pädagogik Band 57, Heft 1 (2011), S. 93-113.
- Tresch, S.: Potenzial Leistungstest. Wie Lehrerinnen und Lehrer Ergebnisrückmeldungen zur Sicherung und Steigerung ihrer Unterrichtsqualität nutzen. Bern: h.e.p. Verlag 2007.
- Tresch, S./Moser, U.: Check 5. Schlussbericht 2005 zuhanden des Departements Bildung, Kultur und Sport des Kantons Aarau. Verf. Ms. Zürich: KBL 2005.
- Weinert, F.: Für und Wider die „neuen Lerntheorien“ als Grundlage pädagogisch-psychologischer Forschung. In: Zeitschrift für Pädagogische Psychologie 10 (1) (1996), S. 1-12.
- Windlin, B./Kuntsche, E./Delgrande, Jordan, M.: Arbeitsüberforderung und -unzufriedenheit von Lehrpersonen in der Schweiz. National repräsentative Ergebnisse demografischer, klassen- und schulbezogener Faktoren. In: Schweizerische Zeitschrift für Bildungswissenschaften Band 33, Nr. 1 (2011), S. 125-144.
- Ziller, T.: Grundlegung zur Lehre vom erziehenden Unterricht. 2. verb. Aufl. Hrsg. v. Th. Vogt. Leipzig: Verlag von Veit&Comp. 1884.