

Beurteilung von Bildungsqualität: Akteure und Instrumente

Referat vor der Konferenz Schweizerischer Gymnasialrektoren
anlässlich der Jahresversammlung vom 18. Mai 2004

Hans-Martin Binder, Interface Institut für Politikstudien, Bereichsleiter Bildung

Sehr verehrte Damen und Herren

Allem voran möchte ich meinen aufrichtigen Dank aussprechen an die Organisatorinnen und Organisatoren Ihrer Jahresversammlung, dafür, dass Sie mich gebeten haben, vor diesem würdigen Gremium im schweizerischen Bildungswesens ein Referat zu halten. Ich habe diese Einladung von Herrn Hirschi sehr gerne angenommen.

Für eine Mehrheit unter Ihnen wird wohl nicht apriori einsichtig sein, vor welchem Hintergrund die Anfrage an mich ergangen ist, zu Ihnen zur Frage der Beurteilung von Bildungsqualität zu sprechen. Zwei Aspekte werden wohl dafür verantwortlich gewesen sein.

Erstens liegt ein Schwerpunkt meiner beruflichen Tätigkeit darin, Aspekte des Bildungssystems zu evaluieren – so zum Beispiel: die Frage von Aufgabenzuteilung und Belastungen von Schulleitungen, neue Formen der Schulaufsicht oder (und das hat vermutlich zu meiner Einladung an diesen Anlass wesentlich beigetragen) die Evaluation der Umsetzung des neuen Maturitätsanerkennungsreglementes MAR in verschiedenen Deutschschweizer Kantonen.

Zweitens führe ich im Rahmen der in vielen Kantonen etablierten Qualitätsmanagementsystemen im Bildungsbereich Aus- und Weiterbildungskurse für Qualitätsbeauftragte beziehungsweise Evaluationsbeauftragte von Schulen durch – mit dem Ziel und dem Auftrag, den betroffenen Lehrpersonen oder Schulleiterinnen und Schulleitern das Verständnis für Schulevaluation zu vermitteln und ihnen das Rüstzeug zu geben, schulintern zielführende und entwicklungsorientierte Selbstbeurteilungen vornehmen zu können.

„Ist Bildung messbar?“ – Das der Titel Ihres diesjährigen Jahreskongresses. Ich möchte mein Referat, bei dem es ja im Kern um die Frage gehen soll: *Wer beurteilt Bildungsqualität mit welchen Instrumenten?* unter drei Aspekten beleuchten. Nämlich – erstens – unter einer konzeptionellen Optik, zweitens unter einer wissenschaftlich-methodischen Optik sowie – drittens – unter einer emotionalen und schulkulturellen Optik. Diese drei Aspekte kommen auch im Zitat zum Ausdruck, das ich meinen Ausführungen voran stelle. Es handelt sich um den Auszug aus einem aktuellen schriftlichen Lehrerbericht einer Fachschaft an einer schweizerischen Kantonsschule.

„Seit einigen Jahren ist im Schulwesen, zuerst auf der Volksschulstufe, dann auch auf unserer Stufe, eine Entwicklung zu beobachten, die zu Besorgnis Anlass gibt. Unter dem Titel, die ‚Qualität‘ sichern beziehungsweise verbessern zu wollen, wurden und werden verschiedene Massnahmen verfügt, die sich auf die Arbeit der Lehrerinnen und Lehrer insgesamt negativ auswirken dürften: Einerseits, da sie eine zeitliche Mehrbeanspruchung bedeuten, andererseits, weil sie in der Tendenz einengend und uniformisierend wirken. (...)

Wir wenden uns gegen jegliche Absicht, ‚Qualität‘ messen und verordnen zu wollen, aus der Überzeugung, dass hier Dinge gemessen werden sollen, die nicht messbar sind. (...)

(aus: Lehrerbericht 2000/2004 einer Deutschschweizer Kantonsschule)

Wir erkennen in diesem Zitat – in einer markant kritischen Betrachtungsweise – die drei Aspekte des Themas „Beurteilung von Bildungsqualität“: Von einer neuen Entwicklung ist da die Rede, aus der heraus Massnahmen verfügt und Qualität verordnet werden – offenkundig eine Frage des Konzeptes. Dann der Aspekt, Qualität messen zu wollen, Dinge messen zu müssen, die nicht messbar sind – sicher eine Frage auf der empirischen, methodischen Ebene. Und dann noch die Befürchtung, dieses ganze mühselige Tun wirke einengend und uniformisierend und wirke sich negativ auf die Arbeit der Lehrerinnen und Lehrer aus – eine Befürchtung, die ich primär der emotionalen, aber auch der schulkulturellen Ebene zuordnen würde.

Meine Ausführungen zielen darauf ab,

- erstens: die Beurteilung von Bildungsqualität und die Frage, wer Bildungsqualität aus welcher Optik beurteilen soll, in einer Konzeption von Qualitätsmanagement zu verorten;
- zweitens: ein paar methodische Hinweise zu geben, wie Bildungsqualität auf empirischer Basis beurteilt werden kann, so dass die Ergebnisse Gültigkeit und Relevanz aufweisen und gleichzeitig wissenschaftlichen Standards und Kriterien genügen;
- drittens: das Bewusstsein dafür zu schärfen, dass die Beurteilung von Bildungsqualität – insbesondere wenn sie als formative, entwicklungsorientierte Beurteilung angelegt ist – eingebettet sein muss in eine Schulkultur, die den offenen Diskurs über die Qualität von Schule und Unterricht grundsätzlich zulässt und fördert.

Auf dieser Reflexionsbasis sollten dann – so meine Hoffnung – überkritische Positionen und Abwehrhaltungen gegenüber Sinn und Möglichkeiten der Beurteilung von Bildungsqualität, wie sie in der vorhin präsentierten Meinung zum Ausdruck kommen, keine Stichhaltigkeit und eigentlich auch keine Berechtigung mehr haben.

Als erstes nun die Verortung von Qualitätsbeurteilungen in einem Gesamtkonzept von Schulqualitätsmanagement.

In den vergangenen 10 bis 15 Jahren hat im Bereich des öffentlichen Bildungswesens eine Entwicklung stattgefunden von der Schule als Institution und Gebäude, in dem eine Anzahl von Lehrpersonen Schülerinnen und Schüler unterrichten, hin zu teilautonomen, geleiteten Schulen – verstanden als pädagogische und organisationale Einheit. Diese Entwicklung hat eine direkte Implikation auf die Frage der Beurteilung von Bildungsqualität. Ging es früher allein um die Frage nach dem „guten“ Lehrer beziehungsweise dem „guten“ Unterricht, so stellt sich die Qualitätsfrage unter der Prämisse der teilautonomen geleiteten Schule weiter gefasst, nämlich als die Frage: „Was ist gute Schule?“ beziehungsweise „Was ist eine gute Schule?“ Und: Welche institutionellen, betriebsorganisatorischen, führungsmässigen, sozialen und ressourcenbezogenen Bedingungen fördern oder determinieren guten Unterricht und die erfolgreiche Vermittlung von Bildung?

Früher liess sich also die Schul- oder Bildungsqualität allein aus der Perspektive der pädagogischen – und allenfalls didaktischen – Kompetenz der einzelnen Lehrperson betrachten. Diesem Konzept entsprach ja dann die Funktion der Qualitätsbeurteilung in kontrollierendem Sinn durch Laienschulpflegen beziehungsweise Schulinspektoren mit Unterrichtsbesuchen und Visitationsberichten. Die Qualitätsfrage liess sich demzufolge reduzieren auf die relativ simple Gleichung: „Guter Lehrer“ oder „gute Lehrerin“ gleich „guter Unterricht“ gleich „gute Schule“. Punktum.

Die jüngere Entwicklung hin zu teilautonomen geleiteten Schulen hat dagegen zur Konsequenz, dass die Sorge und die Verantwortung um die gute Schule – um Bildungsqualität – neu an die Schule als pädagogische Einheit und Organisationseinheit delegiert wird. Nicht mehr die einzelne Lehrperson, die von Abgesandten des hoheitlichen Bildungssystems kontrolliert wird, steht in der primären Verantwortung, sondern die innerhalb einer Schuleinheit an Schule Beteiligten – jede und jeder entsprechend seiner Rolle und seiner Funktion – tragen die gemeinsame Verantwortung für die Erhaltung, Förderung und Weiterentwicklung der Schul- und Bildungsqualität.

Dieses Paradigma der eigenverantwortlichen Schule impliziert also das Verständnis, wonach die einzelne Lehrperson, die Bildung schafft und vermittelt, nicht isoliert in ihrem Klassenzimmer wirkt, sondern in einem vielfältigen institutionellen, sozialen, kulturellen und atmosphärischen Umfeld steht, und die Tätigkeit der einzelnen Lehrperson durch dieses Umfeld geprägt und mit determiniert wird.

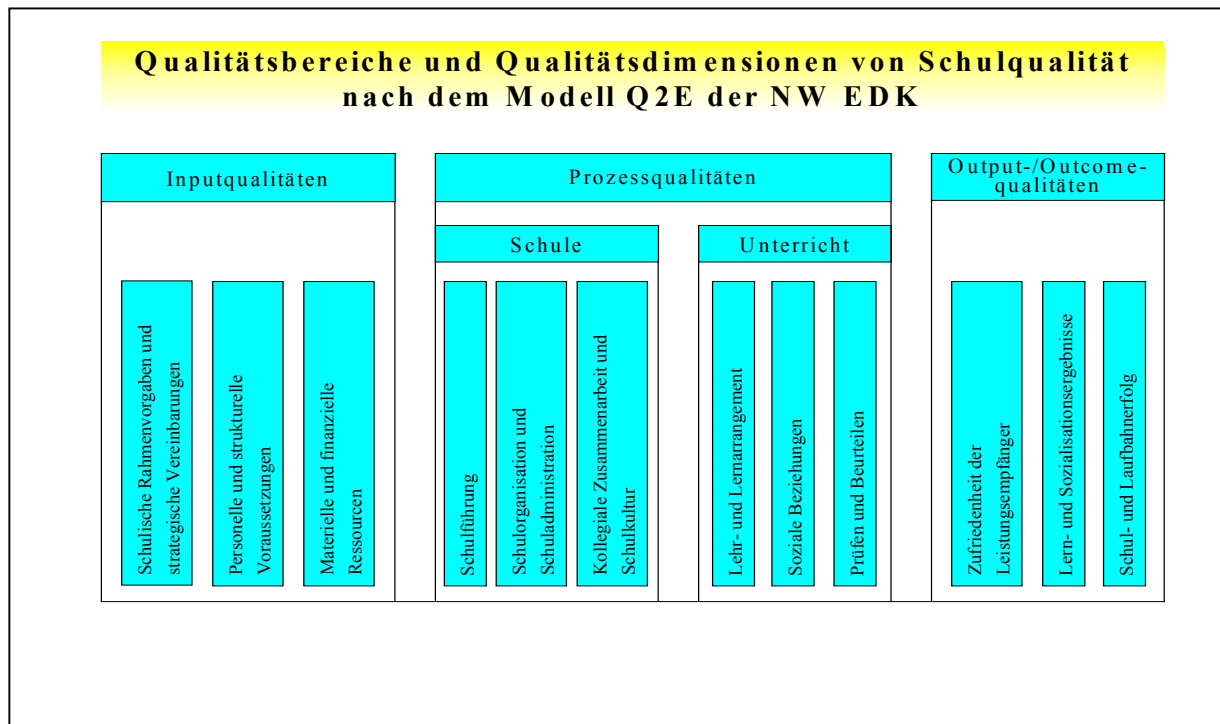
Aus dieser Sichtweise wiederum lässt sich ableiten, welche Elemente in einem Konzept, das Bildungsqualität erfassen und beurteilen will, mit einbezogen werden müssen. Nämlich:

1. verschiedene Akteure mit unterschiedlichen rollen- und funktionsspezifischen Haltungen, Erfahrungen und Sichtweisen;
2. verschiedene strukturelle Ebenen: Einzelne Lehrpersonen, Schulhausteams, die einzelne Schule als pädagogische und organisationale Einheit, Institutionen im Umfeld der Einzelschule sowie das Bildungssystem als Ganzes;
3. verschiedene thematische Zugänge: Neben dem Unterricht sind das auch Aspekte der Kommunikation, der Zusammenarbeit, der Schulorganisation, der Ressourcen usw.

Das knappe Fazit bis dahin: Der Blick auf Bildungsqualität ist in hohem Mass komplexer geworden.

Versucht man nun diese differenzierte Sicht auf Bildungsqualität konzeptionell zu fassen, dann gilt es im wesentlichen zwei Ebenen zu unterscheiden: Einerseits die Ebene von inhaltlichen Dimensionen von Qualität, andererseits die verschiedenen Ebenen des Bildungssystems.

Auf der *inhaltlichen Ebene* hat beispielsweise das speziell für den Bildungsbereich entwickelte Qualitätssystem Q2E – ein Projekt der Nordostschweizer Erziehungsdirektorenkonferenz NW EDK – schul- und bildungsrelevante Qualitätsbereiche und Qualitätsdimensionen in einem Modell zusammengefasst.



Insgesamt zwölf Dimensionen werden hier für die Bereiche Inputqualitäten, Prozessqualitäten sowie Output- beziehungsweise Outcomequalitäten (also: Qualitäten auf der Wirkungsebene) definiert. Mit diesem Gerüst lässt sich Bildungsqualität inhaltlich schon sehr differenziert fassen und gleichzeitig in klare und abgrenzbare Bereiche gliedern.

Eine wesentliche Voraussetzung für zielführende und inhaltlich fokussierte Qualitätsbeurteilungen ist, von Anfang an Transparenz und Klarheit darüber zu schaffen, was denn genau in den Fokus der Beurteilung genommen werden soll. Das Modell von Q2E ist dazu ein bewährtes Instrument.

Auf der *systemischen Ebene* kann einer Gesamtkonzeption des Schulqualitätsmanagements ein Modell zu Grunde gelegt werden, wie es etwa die Bildungsdirektion des Kantons Zürich – in sehr ähnlicher Form existiert das Modell auch im Kanton Luzern – etabliert hat.

Elemente des kantonalen Schulqualitätsmanagements

	Interne Beurteilung: Selbstbeurteilung	Externe Beurteilung: Fremdbeurteilung
Ebene „Personen“ Lehrpersonen, Schulleitung	Systematische Selbstbeurteilung z.B. Feedbackgruppen, gegenseitige Hospitation	Systematische Mitarbeiterbeurteilung (LQS)
Ebene „Einzelne Schule“	Systematische Selbstbeurteilung der Schule zu selbst definierten relevanten Themen	Periodische systematische Fremdbeurteilung der Schulen durch professionelle EvaluatorInnen
Ebene „Bildungssystem“ Systemmonitoring	Kontinuierliche Sammlung und Auswertung von bildungsstatistischen Daten	Regelmässige wissenschaftliche Evaluationen zu besonderen Fragen und Projekten

In dieser Matrixform lässt sich sehr klar definieren, aus welcher Sicht – Innensicht oder Fremdsicht – Schul- und Bildungsqualität auf den drei Ebenen „einzelne Lehrperson“, „Schule als pädagogische und organisationale Einheit“ sowie „Bildungssystem“ erfasst und beurteilt werden soll. Auch dieses Modell schafft Transparenz und Klarheit.

Nachdem wir eine Verortung der verschiedenen Konzepte von Beurteilungen von Bildungsqualität vorgenommen haben, will ich im nächsten Schritt auf die methodischen Aspekte der Beurteilung von Schulqualität und Bildungsqualität eingehen.

Ich habe bereits eingangs erwähnt, dass es uns bei der Frage, wie Bildungsqualität beurteilt werden kann, darum geht, dass diese Beurteilung auf empirischer Basis erfolgt, so dass die Ergebnisse Gültigkeit und Relevanz aufweisen und gleichzeitig wissenschaftlichen Standards und Kriterien genügen. Wenn der Begriff „Beurteilung“ in diesem Sinn verstanden wird – nämlich: die Bewertung von Konzepten, Programmen, Projekten oder Massnahmen auf der Basis von wissenschaftlich erfassten Daten und Informationen im Hinblick auf die Überprüfung ihrer Übereinstimmung mit den Zielen, der Effektivität und Effizienz im Vollzug sowie ihrer Wirksamkeit – dann sprechen wir auch von „Evaluation“. Ich verwende diesen Begriff in meinen Ausführungen fortan synonym mit dem Begriff der „Beurteilung“.

Doch bevor wir zu Fragen der Methodik von Erfassung und Beurteilung im engeren Sinn kommen – nämlich zur Frage des *Wie?* – stellen sich methodische Fragen nach dem *Wer?* und dem *Wozu?*

Nehmen wir die Frage des *Wozu?* an den Anfang, weil sich aus dieser Perspektive heraus oft die Antwort auf die Frage nach dem *Wer?*, also nach den evaluierenden, beurteilenden und den zu evaluierenden Akteuren folgerichtig ergibt.

Die Frage nach dem *Wozu?* – also die Frage: Zu welchem Zweck evaluieren wir? – muss zwingend am Anfang einer Evaluation gestellt werden. Deshalb, weil sie verschiedene Implikationen auf das *wer?* und das *wie?* hat. Es soll ja nicht evaluiert werden, allein deshalb, weil es „von oben verordnet“ wird. Verordnete Evaluation, deren Nutzen – im engeren wie im weiteren Sinn – von den Beteiligten nicht als plausibel erkannt werden kann, ist – das würde ich einmal so salopp sagen – zum Scheitern verurteilt. Dies allein schon deshalb, weil man in Evaluationen nur dann zu gültigen und verlässlichen Daten und Informationen – auf denen die anschliessende Beurteilung aufbaut – kommen kann, wenn die entsprechenden Datenlieferanten die notwendige Bereitschaft und Offenheit haben zur Preisgabe von relevanter Information. Unter dem dritten Fokus meiner Ausführungen – den individuellen und schulkulturellen Aspekten von Schul- und Unterrichtsbeurteilung – werde ich nochmals auf diese Frage der Nutzenorientierung und Nutzenplausibilität von Evaluationen zurückkommen.

Zurück zur Frage: Zu welchem Zweck soll Schule und Unterricht beurteilt, evaluiert werden? Wir können hier zwei grundsätzlich unterschiedliche *Funktionen* unterscheiden.

Evaluation kann einerseits die Funktion der *Kontrolle* haben. Wir sprechen in diesem Zusammenhang von *summativer* oder bilanzierender Evaluation. Mit einer summativen Evaluation will sich das Bildungssystem beispielsweise darüber ins Bild setzen, ob, in welchem Umfang und mit welcher Wirkung eine Massnahme oder ein Programm an den Schulen umgesetzt worden ist. Oder sie will Rechenschaft darüber einfordern, wie gut die im Lehrplan gesetzten Lernziele einer Stufe erreicht werden. Summative Evaluationen sind in diesem Sinn meistens auf der Ebene des Bildungssystems angesiedelt – wenn Sie sich an die vorhin präsentierte Matrix der Systemebenen erinnern – und können wichtige Daten und Erkenntnisse für das Bildungsmonitoring eines Bildungssystems darstellen. Im Kanton Zürich sind zum Beispiel aufwändige Studien über Schulleistungen auf verschiedenen Klassenstufen im gesamten Kanton erstellt worden. Oder: Das Projekt „Klassenscockpit“ stellt ein standardisiertes Instrument zur bilanzierenden Beurteilung des Leistungsstandes einer Klasse in einem bestimmten Fach dar.

Evaluation kann andererseits die Funktion haben, Erkenntnisse im Hinblick auf Qualitätsverbesserungen und Qualitätsentwicklungen von Schulprojekten, Institutionen oder ganzer Schulen zu gewinnen. In diesem Fall sprechen wir von *formativen* oder entwicklungsorientierten Evaluationen – in den meisten Fällen prozessbegleitend –, weil sie darauf angelegt sind, die Basis für Entwicklungen, für institutionelles, interpersonales oder individuelles Lernen abzugeben. In diese Kategorie gehören sehr viele Evaluationen. Und zwar – wenn wir wiederum auf die Systemmatrix Bezug nehmen – sowohl im Bereich der Fremdbeurteilung als auch im Bereich der Selbstevaluationen, auf der Ebene der Schule wie auf der Ebene der Einzelperson. Aber auch externe Evaluationen auf der Ebene des Bildungssystems können formativ, entwicklungsorientiert angelegt sein. So habe ich beispielsweise im Kanton Zürich nach einer gewissen Zeit der Versuchsphase die neue Schulaufsicht an der Volksschule mit einem formativen Ansatz evaluiert: Es ging darum, Erkenntnisse über die Zweckmässigkeit der Durchführung, die Wirksamkeit und die Zielorientierung des neuen Systems zu gewinnen, um allenfalls Verbesserungen am Konzept oder an den Abläufen anzubringen. Oder die über mehrere Jahre sich erstreckenden Evaluationen zur Umsetzung des MAR in den Kantonen Luzern und Aargau waren formative Evaluationen, bei denen jeweils nach einzelnen, themenspezifischen Evaluationsphasen die entsprechenden Ergebnisse mit Empfehlungen in die Projektleitung MAR und an die verantwortlichen Schulleitungen zurückgespiesen wurden.

Nach der Frage des *Wozu?* wenden wir uns der Frage zu, *wer* denn beurteilen soll. Es geht also primär darum, zu definieren, ob eine Thema, ein Programm, ein Projekt oder eine Institution von externen Fachpersonen – also im Rahmen einer Fremdevaluation – oder von den Beteiligten intern – also im Rahmen einer Selbstevaluation – beurteilt werden soll. An diesem Punkt nun möchte ich einen Rückbezug herstellen zu meinen Ausführungen im ersten Teil bezüglich der konzeptionellen Grundlage von Qualitätsbeurteilung im Bildungsbereich. Ich habe dargelegt, dass die Entwicklung zu teilautonomen geleiteten Schulen mit verbunden ist mit der weitgehenden Übertragung der Verantwortung für die Qualitätssicherung und die Qualitätsentwicklung der Schule und des Unterrichts an die einzelne Schuleinheit. Diese Entwicklung hat deshalb auch zur Folge, dass der Stellenwert von Selbstbeurteilungen – sei dies auf der Ebene der Schule als pädagogische und organisationale Einheit oder auf der Ebene der einzelnen Lehr- oder Schulleitungsperson – markant gestiegen ist. Sie alle wissen das, weil Sie in den vergangenen Jahren oder aktuell mit dem Aufbau eines wirkungsorientierten Qualitätsmanagements an Ihrer Schule beauftragt wurden.

Ich möchte bei meinen weiteren Ausführungen deshalb den Fokus auf die drei Matrixfelder der Selbst- und Fremdevaluation auf der Ebene der einzelnen Schule sowie auf die Selbstbeurteilung auf der Ebene der einzelnen Lehr- und Schulleitungsperson richten. Denn: Die Übertragung der Verantwortung für das Schulqualitätsmanagement an die einzelne Schule beantwortet im Grunde die wichtige Frage: „Für wen wird evaluiert?“ Und Sie spüren ja sogleich, welche Saite hier anklingt: Eine Schule mit hoher Autonomie und grosser Eigenverantwortung für die Bildungsqualität evaluiert in erster Linie *für sich selbst*. Das betrifft sowohl die Schule als pädagogische und organisationale Einheit als auch die einzelnen Lehrpersonen und verantwortlichen Schulleitungsmitglieder auf der Ebene des Unterrichts und der Führung.

Kommen wir schliesslich zur Frage, wie Qualität im Bildungsbereich – auf der Ebene der einzelnen Schule wie auch auf der Ebene der einzelnen Lehrperson – beurteilt werden kann.

Eine Grundproblematik der Beurteilung stellt sich in der Frage nach dem Massstab, nach der Referenzgrösse. Denn: Beurteilen – oder im härteren numerischen Sinn von Messen – lässt sich etwas nur auf der Basis einer Referenzgrösse, eines definierten Soll-Zustandes. Es stellt sich also die zentrale Frage nach der Operationalisierbarkeit von Qualität oder: die Frage nach der Definition von Zielgrössen und von Qualitätsindikatoren.

Nun, die Frage ist zwar schwierig und sie lässt sich – das muss hier auch klar eingestanden werden – auch nicht immer und in jedem Fall befriedigend klären. Aber immerhin: Wir verfügen im Schul- und Bildungswesen doch über eine ganze Anzahl von Zieldimensionen, von Zielkategorien und von Qualitätsindikatoren. Wollen wir einige Beispiele nennen:

- Jede Schulstufe hat Lehrpläne mit definierten Lern- und Bildungszielen, bisweilen sogar operationalisiert bis auf die Stufe von Unterrichtszielen.
- Schulen verfügen heute fast allesamt über Leitbilder – ein Katalog von Schulzielen auf den verschiedensten Ebenen: Ziele zu Pädagogik und Klassenführung, zu Unterrichtsgestaltung, zu verschiedenen Kompetenzbereichen, zum schulischen Umfeld, zu Teamarbeit und Führung und andere mehr.
- Auf der Basis dieser Leitbilder definieren Schulen im Rahmen ihrer Jahresplanung operationalisierte Teilziele und thematische Entwicklungsprogramme.
- Schulen verfügen auch über verschiedene Instrumente wie Schulordnungen, Pflichtenhefte usw., in denen der schulische Alltag, Formen des schulinternen Zusammenlebens usw. in Normen und Regeln gefasst werden.

- Und schliesslich verfügen Schulleiterinnen und Schulleiter von geleiteten Schulen über ein Management der Mitarbeiterförderung, das Fremd- und Selbstbeurteilung wie auch die Festlegung individueller Entwicklungsziele beinhaltet.

Sie sehen: Ganz ohne Zielgrössen und Qualitätsmassstäbe stehen wir nicht da.

Anzumerken ist im übrigen auch, dass in der letzten Zeit mancherorts – im Bereich der Bildungsforschung wie durch führende Institutionen der Bildungsplanung in verschiedenen Kantonen – starke Anstrengungen unternommen worden sind, Indikatoren für Bildungsqualität in verschiedenen Dimensionen und Teilbereichen zu entwickeln. Diese Indikatoren sind zum Teil definiert worden, um die externen Schulevaluationen – als Form der neuen Schulaufsicht – mit den notwendigen Standards zu untermauern und/oder um ein aussagekräftiges Bildungsmonitoring als Steuerungsinstrument des Schulsystems aufzubauen. So gesehen scheint mir die These (sie erinnern sich an die eingangs zitierte Meinung einer Fachschaft), Bildung und Bildungsqualität sei nicht messbar und beurteilbar, fälschlicherweise etwas allzu apodiktisch.

Wenn wir einen Schritt weiter gehen auf der methodischen Ebene der Frage nach dem wie beurteilen, dann stellt sich dabei die Frage, *aus welcher Sicht* die Beurteilung von Schule und Unterricht, von verschiedenen Aspekten der Bildung, erfolgt. Gerade weil wir es in sehr vielen Fällen von Beurteilungen mit Inhalten zu tun haben, die nicht oder nur partiell mit normengestützten, quantifizierbaren und skalenbildenden Indikatoren und Messgrössen erfasst werden können – also zum Beispiel Leistungstests, Maturaquoten und ähnliches – sind wir bei der Beurteilung und Bewertung von Bildungsqualität, Unterrichtsqualität oder Schulqualität auf *qualitative, interpretierende* Analysemethoden angewiesen. Und weil wir in diesem Sinn die Realität nicht in statistischen Messgrössen abgebildet erhalten, müssen wir empirische Methoden zur Anwendung kommen lassen, die uns der Realität durch eine *mehrperspektivische* Sicht annähern lässt. Unsere empirische Tätigkeit lässt gewissermassen ein kaleidoskopartiges Bild der Wirklichkeit entstehen, das den sozialwissenschaftlichen Gütekriterien zu entsprechen vermag, nämlich: der *Reliabilität* (also die Verlässlichkeit, das was gemessen werden soll auch exakt zu erfassen), der *Validität* (also die Genauigkeit, das zu erheben, zu erfragen und zu beobachten, was wir zu messen vorgeben) sowie durch die Umschreibung des *Geltungsbereichs* (also die Umschreibung der sozialen Gegebenheiten, unter denen die Ergebnisse der Untersuchung gelten).

Einen solchen mehrperspektivischen Ansatz – oder eine *Triangulation*, wie wir das in den empirischen Sozialwissenschaften nennen – erreichen wir auf drei Ebenen, nämlich durch *Daten-Triangulation*, durch *Methoden-Triangulation* und durch *Akteur-Triangulation*. Das heisst nichts anderes, als dass wir zur Erfassung der Realität unseres Beurteilungs- oder Evaluationsgegenstandes verschiedene Daten und Datentypen mittels verschiedener Erhebungsmethoden bei verschiedenen beteiligten oder betroffenen Akteuren erfassen und interpretieren. Lassen Sie mich das durch ein Praxisbeispiel illustrieren: Wir haben im Rahmen unserer MAR-Evaluationen unter anderem die Frage der Motive der Schülerinnen und Schüler bei der Wahl ihrer Schwerpunktfächer analysiert. Wir haben vier Zugänge gewählt, um diese Frage untersuchen und interpretieren zu können: Erstens haben wir die verschiedenen Gründe für die Wahl eines Schwerpunktfaches im Rahmen einer schriftlichen Umfrage bei allen Studierenden zweier Klassenstufen aller Kantonsschulen eines Kantons (also eine Vollerhebung bei zwei Jahrgängen) befragt und statistisch ausgewertet. Zweitens haben wir Vertreterinnen und Vertreter aller Fachschaften aller Kantonsschulen eines Kantons in leitfadengestützten Gruppengesprächen zu ihrer Einschätzung und ihren Erfahrungen bezüglich der Gründe und Motive der Schülerinnen und Schüler befragt und sie dabei auch mit den Ergebnissen aus der Schülerumfrage konfrontiert. Ausgewertet haben wir die Gesprächsprotokolle mit einem Textauswertungssystem, das uns zu gewichteten Typen und Kategorien von Aussagen ge-

führt hat. Drittens haben wir die Schulleitungen aller Kantonsschulen eines Kantons in Einzelinterviews zu ihren Einschätzungen und Erfahrungen befragt und sie dabei auch mit den Ergebnissen aus der Schülerumfrage und der Lehrerumfrage konfrontiert. Viertens haben wir die Ergebnisse der Untersuchung eines Kantons mit den Ergebnissen aus dem anderen Kanton verglichen. Unser Untersuchungsansatz hat akteur- beziehungsweise rollendifferenzierte und gewichtete Aussagen möglich gemacht, der überdies fachspezifische und schulspezifische Differenzierungen sowie eine Differenzierung auf der Zeitachse – das heisst: mögliche Veränderungen zwischen den Jahrgängen – ermöglicht hat.

Sie werden mir jetzt vielleicht sagen: Schön und gut, eine derart elaborierte Untersuchung kann sich Herr Binder als professioneller Evaluator mit einem hübschen Projektbudget und analytischen Ressourcen im Hintergrund vielleicht leisten. Auf der Ebene einer einzelnen Schule sind solche Untersuchungsdesigns im Rahmen von schulinternen Selbstbeurteilungen aber kaum realisierbar. Das mag zum Teil durchaus stimmen. Und es ging mir bei diesem Beispiel auch nicht darum, einen zwingend einzuhaltenden konzeptionellen Standard zu definieren. Es ging mir nur darum, ihnen den Ansatz der Triangulation zu veranschaulichen. Ein Ansatz, der aber im Grunde auch in einfacheren Evaluationen angewendet werden kann. Ich habe im Rahmen meiner MAR-Evaluation im Kanton Luzern Selbstevaluationen von Kantonschulen gesehen, die für einzelne Fragestellungen ebenfalls mehrperspektivische Ansätze zur Anwendung gebracht haben.

Bei diesem Teil meiner Ausführungen ist es darum gegangen, Ihnen ein paar methodische Hinweise zu geben, wie Bildungsqualität auf empirischer Basis beurteilt werden kann, so dass die Ergebnisse Gültigkeit und Relevanz aufweisen und gleichzeitig wissenschaftlichen Standards und Gütekriterien genügen.

Lassen Sie mich die gemachten Ausführungen kurz zusammenfassen mit der Auflistung der wichtigsten Gelingensbedingungen von Wirkungs- und Qualitätsbeurteilungen.

Erstens: Jede Evaluation muss im voraus die folgenden vier Fragen beantworten:

Was soll beurteilt werden? Es geht darum, das Thema, den Beurteilungsgegenstand zu definieren. Das Modell Q2E mit den verschiedenen Dimensionen von Schulqualität kann dabei eine Hilfe zur Verortung der definierten Thematik darstellen.

Für welchen Zweck und für wen soll beurteilt werden? Die Frage stellt sich: Geht es um eine summative, bilanzierende Evaluation, die – aus der Sicht der Schule – vielleicht primär eine Kontrollfunktion erfüllt, oder geht es um eine formative, entwicklungsorientierte Evaluation, bei der der Verwendungszweck – aus der Sicht der Schule – primär bei der Schule selber liegt und der Optimierung und Verbesserung von Qualitätsaspekten geht? In diesem Fall geht es ganz zentral darum, dass die Evaluation für die Schule nützlich ist und verwendbare, umsetzbare Ergebnisse erzeugt.

Wer soll die Evaluation durchführen? Hier stellt sich die Frage, ob ein Thema oder eine Fragestellung von den Betroffenen im Rahmen einer internen Evaluation selbst oder von einer externen Fachperson oder Fachinstitution (z.B. einer Fachstelle für Schulevaluation) durchgeführt werden soll. Selbstverständlich ist auch eine Kombination beider Ansätze möglich.

Wie soll evaluiert werden? Es geht darum zu klären, mit welchen methodischen Instrumenten welche Akteure zur Erfassung der für eine Beurteilung notwendigen Daten eingesetzt werden und wie die Beurteilung erfolgen kann, damit relevante Aussagen unter Einhaltung sozialwissenschaftlicher Minimalstandards resultieren. Und schliesslich geht es auch um die Klärung der Frage, mit welchem Aufwand und welchen notwendigen Ressourcen die Evaluation durchgeführt werden kann.

Zweitens: Jede Evaluation muss für alle Beteiligten transparent und plausibel verortet sein in einem Gesamtsystem des Qualitätsmanagements – Sie erinnern sich an die Systemsmatrix – und damit auch erkennbar machen, zu welchem Zweck und mit welcher Zielsetzung Evaluationen durchgeführt werden.

Drittens: Jede Evaluation muss die an einer Fragestellung Beteiligten in geeigneter Weise in die Untersuchung mit einbeziehen. Die Frage der Wahlmotive von Schülerinnen und Schülern bei der Schwerpunktfachwahl kann nicht zu glaubwürdigen und relevanten Erkenntnissen führen, ohne die Studierenden in die Informationsgewinnung mit einzubeziehen. Oder ein anderes Beispiel: Die Beurteilung der Qualität von Lehrerleistungen ist mit dem kollegialen Feedback im Rahmen einer Hospitation nicht umfassend möglich – auch wenn das fraglos ein sehr wichtiges Instrument einer entwicklungsorientierten Selbstevaluation auf der Ebene der einzelnen Lehrperson ist. Zu einer umfassenden Beurteilung würde – im Sinne einer 360-Grad-Beurteilung – auch die Sicht der betroffenen Schülerinnen und Schüler, der Schulleitung und der Eltern gehören.

Zudem muss angefügt werden, dass der Einbezug von mitbeteiligten Akteuren – insbesondere auf der Ebene der Selbstevaluationen – die Akzeptanz von Beurteilungen ganz wesentlich steigert.

Im letzten Teil meines Referates möchte ich noch kurz auf den emotionalen und schulkulturellen Aspekt von Evaluationen im Schul- und Bildungsbereich eingehen. Um den Faden für diesen Aspekt aufzunehmen, möchte ich nochmals zurück gehen zum eingangs erwähnten Zitat aus dem Lehrerbericht einer Fachschaft. Darin kommt zum Ausdruck, dass Lehrpersonen den Eindruck haben, Massnahmen zur Beurteilung von Schul- und Unterrichtsqualität würden sich primär negativ auf ihre Arbeit auswirken: sie stellten eine Mehrbelastung dar, würden einengend und uniformisierend wirken. Und sie würden Dinge messen wollen, die nicht messbar sind.

In dieser Haltung kommt zum Ausdruck, dass eine grosse Zahl von Lehrpersonen offenkundig nach wie vor dem schulkulturellen und pädagogischen Selbstverständnis verhaftet sind, wonach Schul- und Bildungsqualität primär – wenn auch nicht ausschliesslich – durch die pädagogische und didaktische Qualität der einzelnen Lehrperson determiniert ist. Zudem gewährleistet – in dieser Sichtweise – die intrinsische Motivation der einzelnen Lehrperson apriori die Qualität des Unterrichts und der vermittelten Bildung.

Es wäre nicht korrekt, diese Haltung einfach als falsch oder störend zu verurteilen oder sie auszublenden oder zu bagatellisieren. Vielmehr muss durch sie unser Bewusstsein dafür geschärft werden, dass die Beurteilung von Schul- und Bildungsqualität eingebettet sein muss in eine Schulkultur, die den offenen Diskurs und Dialog über die Qualität von Schule und Unterricht zulässt und aktiv fördert.

Einerseits muss diese Schulkultur allen Beteiligten klar vermitteln, dass für die Bildungsqualität nicht allein der Unterricht der einzelnen Lehrperson verantwortlich ist. Dass dieser Unterricht vielmehr in einem institutionellen und sozialen Umfeld stattfindet und von zahlreichen Akteuren mit unterschiedlichen Rollen, Funktionen und Erwartungen mit geprägt ist.

Andererseits muss die Schulkultur und die Umsetzung von Qualitätsmanagementkonzepten für die Beteiligten erkennbar und erfahrbar machen, dass die formative Beurteilung von Teilqualitäten im Schul- und Unterrichtsbereich zu Entwicklungsprozessen und Qualitätsverbesserungen führen kann, die schliesslich auf individueller Ebene zu erhöhter Berufszufriedenheit, Motivation und Leistungssteigerung führt.

Im Rahmen der Evaluation der Gymnasialreform im Kanton Luzern haben wir mit gezielter und nachhaltiger Aus- und Weiterbildung der Evaluations- und Qualitätsbeauftragten der verschiedenen Kantonsschulen dazu beigetragen, die notwendigen Kompetenzen für schulinterne Selbstevaluationen aufzubauen. Alle Schulen haben in den vergangenen vier Jahren je zwei interne Selbstbeurteilungen zu selbst ausgewählten relevanten Themen durchgeführt. Im Verlauf dieses Prozesses ist an allen Schulen die Einsicht in die Nützlichkeit solcher Qualitätsbeurteilungen im Hinblick auf Verbesserungen und Entwicklungen gewachsen. Selbstevaluationen haben sich in dieser Zeit als fixe Bestandteile eines Qualitätsentwicklungssystems etabliert. Gute Erfahrungen mit nützlichen und verwertbaren Evaluationen lassen dieses „Schreckgespenst“ oder „Unding“ in Schulhausteams und Kollegien zu einem akzeptierten und verbindlichen qualitätsfördernden Instrument werden.

Aus diesem Grund nenne ich zwei wichtige zusätzliche Gelingensbedingungen für Qualitätsbeurteilungen: Erstens kann Schul- und Unterrichtsevaluation nicht einfach auf simple Art verordnet oder angeordnet werden. Und zweitens müssen Beteiligte in möglichst hohem Mass einbezogen werden in die Erfassung und Interpretation von Ist-Qualitäten und in die Weiterbearbeitung der Evaluationsergebnisse mit Blick auf ziel- und wirkungsorientierte Entwicklungsschritte und Verbesserungsmaßnahmen.

Und schliesslich das noch: Wenn ich von der Notwendigkeit einer diskursorientierten Schulkultur spreche, so ist mir bewusst – und das Zitat des Lehrerberichtes zeigt das in aller Deutlichkeit – dass wir es hier mit einem notwendigen Kulturwandel zu tun haben. Und Kulturwandel braucht seine Zeit. Wir müssen der Etablierung eines Bewusstseins für Qualitätsmanagement und Qualitätsbeurteilungen im Schul- und Bildungsbereich diese Zeit gewähren. Vorgefasste Meinungen lassen sich nur über den Weg der positiven Erfahrung und des Erkennens von persönlichem und institutionellem Nutzen überführen in neue Einsichten und die Bereitschaft, bei Massnahmen zur Qualitätsbeurteilung und Qualitätsentwicklung aktiv und konstruktiv mitzutun.

Deshalb – und damit komme ich zum Schluss meiner Ausführungen:

Die Beurteilung von Schul- und Bildungsqualität muss

- im Rahmen von relevanten Teilbereichen,
- auf der Basis definierter Zielgrößen und Indikatoren und
- einer Triangulation von empirischen Methoden
- unter Einbezug der verschiedenen beteiligten und betroffenen Akteure mit ihren rollen- und funktionsspezifischen Sichtweisen
- zu transparenten, gültigen und nachvollziehbaren Ergebnissen und Erkenntnissen führen und
- in kooperativen Prozessen zu ziel- und wirkungsorientierten Entwicklungsschritten und qualitätsrelevanten Massnahmen leiten.