

Guy Kempfert
Hans-Günter Rolff

PÄDAGOGIK

Handbuch Qualität und Evaluation

5. Auflage

vollständig
überarbeitet



E-Book inside

BELTZ

Kempfert/Rolff

Handbuch Qualität und Evaluation

Guy Kempfert/Hans-Günter Rolff

Handbuch Qualität und Evaluation

Ein Leitfaden für Pädagogisches Qualitätsmanagement

5., vollständig überarbeitete und erweiterte Auflage

BELTZ

Guy Kempfert war Schulleiter an einem Schweizer Gymnasium, Schulinspektor in Zürich und anschließend in zwei internationalen Konzernen für die Personal- und Organisationsentwicklung verantwortlich. Er ist heute selbstständiger Trainer, Berater und Coach (guy.kempfert@gmail.com).

Hans-Günter Rolff ist emeritierter Professor am Institut für Schulentwicklungsforschung der Technischen Universität Dortmund (IFS), Vorsitzender des Akademierats der Deutschen Akademie für Pädagogische Führungskräfte (DAPF) und wissenschaftlicher Leiter des Fernstudiengangs Schulmanagement der Universität Kaiserslautern (hans-guenter.rolff@tu-dortmund.de).

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt
Jede Verwertung ist ohne Zustimmung des Verlags unzulässig.
Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen
und die Einspeicherung und Verarbeitung in elektronische Systeme.



Dieses Buch ist erhältlich als:
ISBN 978-3-407-83194-1 Print
ISBN 978-3-407-63096-4 E-Book (PDF)

5., vollständig überarbeitete und erweiterte Auflage 2018

© 2005 Beltz Verlag · Weinheim und Basel
Werderstraße 10, 69469 Weinheim
Alle Rechte vorbehalten

Lektorat: Dr. Erik Zyber
Herstellung und Satz: Michael Matl
Druck: Beltz Grafische Betriebe, Bad Langensalza
Umschlaggestaltung: Michael Matl
Zeichnungen: Dieter Surm
Printed in Germany

Weitere Informationen zu unseren Autoren und Titeln finden Sie unter: www.beltz.de

Inhaltsverzeichnis

Vorwort	8
I. Zwei Linien der Qualitätsentwicklung	9
1. Von der Schule ausgehende Entwicklung	10
2. Zentral administrierte Entwicklung	20
3. Auf dem Wege zum integrierten Qualitätsmanagement?	29
II. PQM – ein Modell für die Praxis	35
1. Qualitätsarbeit als Entwicklung von Einzelschulen	36
2. Das Modell im Überblick	45
3. Gütesiegel/Zertifikat als Vision	50
III. Praxisfelder: Alle sind für Qualität verantwortlich	55
1. Zielführendes Handeln als Ausgangspunkt	56
2. Feedbackkultur einführen und pflegen	61
3. Kooperative Unterrichtsentwicklung	68
3.1 Fachkonferenzen	69
3.2 Professionelle Lerngemeinschaften	85
IV. Interne Evaluation	93
1. Qualitätsindikatoren	96
2. Erhebungsinstrumente	102
3. Technische Hinweise	106
4. Kultur der Evaluation	111

V. Evaluation und Entwicklung des Unterrichts	119
1. Was ist guter Unterricht?	120
2. Unterrichtsrelevante Kompetenzen	126
3. Selbstevaluation von Lehrerinnen und Lehrern	130
4. Schülerelbstevaluation	140
5. Reziprokes Feedback mit Schülern	150
6. Digital gestütztes Feedback	165
7. Feedbackkonferenzen	168
8. Kollegiale Unterrichtshospitation	180
9. Leitungsfeedback	193
VI. Schulevaluation	199
1. Leitbild und Schulprogramm	201
2. Mentale Modelle	203
3. Schüler- und Elternbefragungen	207
4. Ratingkonferenzen	215
VII. Lernerfolgfeststellungen	227
1. Parallelarbeiten/Beispielaufgaben	229
2. Einzelschulische zentrale Lernstandserhebungen	234
3. Innerschulische Verarbeitung	238
VIII. Externe Evaluation muss sein: Konzepte und Beispiele	245
1. Formen externer Evaluation	247
2. Authentische Evaluation statt Fassadenevaluation	250
3. Peer Review: Ein Fallbeispiel aus der Schweiz	253
4. Knackpunkte und Hinweise	268
5. Einige Verallgemeinerungen	274
IX. Einzelne Schritte gehen, aber das Ganze sehen	277
1. Die Schulleitung führt den Prozess	278
2. Management durch eine Steuergruppe	284
3. Personalauswahl	286

4. Jahresgespräche	292
5. Zielvereinbarungen	303
6. Man kann jede Stelle dieses Buches zum Anfang nehmen	310
X. Zehn-Komponenten-Check von Qualitätsmanagement-Systemen	313
XI. Anhang	319
1. Standards für Evaluation – Kurzfassung	320
2. IFES: Externe Evaluation in der Sekundarstufe II in der Schweiz	325
3. Satire: Evaluation der Wiener Philharmoniker	336
Literaturverzeichnis	338
Sachregister	345

Vorwort

Qualitätsmanagement und Evaluation sind für die nächsten Jahre ein bedeutsames Thema, vielleicht sogar das bedeutsamste überhaupt. Qualitätsmanagement ist anspruchsvoll. Denn es muss ganzheitlich sein, auf einer umfassenden internen Evaluation beruhen, extern überprüft werden und sich womöglich auch noch für eine Zertifizierung eignen. Deshalb ist es unlösbar mit Evaluation verknüpft. Zudem sollte es nicht nur einen Managementfokus haben, sondern dezidiert pädagogisch orientiert sein.

Qualitätsmanagement von Schulen kann im Prinzip auf zweierlei Weise betrieben werden: von oben und von unten. Vermutlich ist eine Kombination von beidem das Effektivste. Über Qualität von oben herrscht eine intensive bildungspolitische Diskussion; bundes- bzw. landesweite Standards und nationale wie internationale Tests sind die Stichwörter. Über Qualität von unten gibt es keinen öffentlichen Diskurs und es ist auch wenig klar, was damit gemeint ist. Deshalb wird in diesem Buch dargelegt, wie Qualitätsentwicklung von unten, d. h. in der Einzelschule, aussieht und wie sie gelingen kann.

Der Vorläufer zu diesem Buch kam 1999 heraus und ist seitdem in drei Auflagen erschienen. 2005 wurde das Buch vollständig überarbeitet. Die neuen Inhalte betreffen vor allem die »Treiber« des Qualitätsmanagements – zielführendes Handeln, Feedbackkultur und Teamentwicklung – sowie Tests und Standards, innerschulische Verarbeitung der Testergebnisse, Parallel- und Vergleichsarbeiten, schulinterne Unterrichtsevaluation, kollegiale Hospitation und professionelle Lerngemeinschaften.

So sind die Umrisse eines umfassenden pädagogischen Modells des Qualitätsmanagements für Schulen entstanden. Wir nennen es »Pädagogisches Qualitäts-Management« (PQM). Es ist als System angelegt, aber nicht als bürokratisches. Es eignet sich für alle Schulen aller Schulformen.

Die vorliegende 5. Auflage haben wir zum Handbuch erweitert, das alle Komponenten von Qualität und Evaluation umfasst und durch ein Register effektiv erschließbar macht.

Das Autorengespann hat sich bei gemeinsamer Schulentwicklungsarbeit bewährt. Es handelt sich um einen konzeptorientierten ehemaligen Schweizer Gymnasialschulleiter und einen praxisorientierten deutschen Erziehungswissenschaftler. Wir danken den Kolleginnen und Kollegen des Gymnasiums Liestal (Schweiz) und der Deutschen Akademie für Pädagogische Führungskräfte (DAPF) für zahlreiche Anregungen und Dieter Surm für die Karikaturen.

I. Zwei Linien der Qualitätsentwicklung

Schon vor TIMSS und PISA war Qualitätsmanagement eine wichtige Aufgabe, nach den ernüchternden Ergebnissen dieser internationalen Fachleistungs-Vergleichsstudien gilt das erst recht. Jetzt erwartet die öffentliche Diskussion, dass eine Leistungssteigerung in den getesteten Hauptfächern erfolgt.

Die Bildungspolitiker aller Couleur setzen auf Qualitätsmanagement. Es werden die unterschiedlichsten Modelle diskutiert und teilweise auch ausprobiert. Die entscheidende Frage an diese Modelle lautet nach PISA:

»Wie sollen die Leistungsdefizite der Schülerinnen und Schüler behoben werden?«

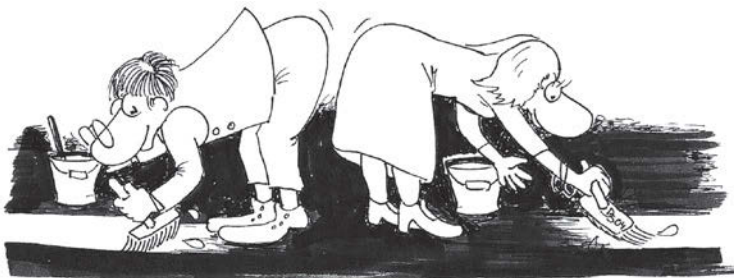
Vor allem aus der angelsächsischen Forschung wissen wir, dass es offenbar besonders schwierig ist, »schwache«, d. h. ineffektive Schulen zu verbessern (Stoll/Myres 1998). Deshalb müssen sich die Modelle des Qualitätsmanagements auch an der Zusatzfrage messen, ob sie in der Lage sind, die Qualität dieser Schulen mit geringem Förderpotenzial zu heben.

In der Theorie wie in der Praxis sogenannter Neuer Steuerung des Schulsystems sind zwei Linien zu erkennen, die sich deutlich unterscheiden und miteinander konkurrieren: das von der Zentrale her konzipierte Qualitätsmanagement und das von der Entwicklung von Einzelschulen her gedachte Modell.

Im Folgenden werden beide Linien zunächst als theoretische Idealtypen (im Sinne Max Webers) ausgeschärft und dann empirisch beschrieben sowie auf ihre Wirksamkeit für Qualitätsverbesserung hin untersucht.

Die zwei Einleitungskapitel sind ziemlich komplex und deshalb nicht leicht zu lesen. Die Komplexität ist darin begründet, dass Wirkung wesentlich aus Zusammenwirken entsteht und deshalb von ganzheitlichen Qualitätsentwicklungssystemen die größte Wirkung zu erwarten ist (vgl. Rolff 2016, S. 150 ff.).

Die darauffolgenden Kapitel sind deutlich einfacher aufgebaut und konkreter im Inhalt.



1. Von der Schule ausgehende Entwicklung

Der Grundtyp A ist der ältere von den beiden hier unterschiedenen Idealtypen. Er entstand in den 80er-Jahren, als das Scheitern der bis dato vorherrschenden zentralen Bildungsplanung offensichtlich wurde (vgl. Rolff 1991).

Initiierung: Einzelschule als Gestaltungseinheit

Der Grundtyp A baut auf die Einzelschule als Gestaltungseinheit oder – wie es auch formuliert wurde – als »Motor der Entwicklung«. Die dahinter stehende, durch zahlreiche empirische Studien erhärtete Idee ist, dass Qualität schulischer Arbeit und Ergebnisse nicht zentral erwirkt bzw. angeordnet werden kann, sondern sich nur *entwickelt*, wenn die Akteure innerhalb einer Schule dieses wollen und sie dabei unterstützt werden.

Grundtyp A steht für das Schulentwicklungsmodell. Schulen setzen sich selber Qualitätsziele und überprüfen regelmäßig, wie weit sie diese erreichen. Sie geben sich ein Qualitätsleitbild und stellen systematische Überlegungen an, wie sie es umsetzen (»implementieren«) können. Die Umsetzungsprogramme werden zumeist als Schulentwicklungsprogramme bezeichnet.

Kernprozess der Qualitätsentwicklung ist Unterrichtsentwicklung. Personal- und Organisationsentwicklung kommen hinzu, um das professionelle Selbst der in der Schule Handelnden zu fördern und für innerorganisatorische Abstützung zu sorgen.

Der Grundtyp A ist dezidiert als Entwicklungsmodell und gleichzeitig als Lernmodell zu verstehen. Da nicht davon auszugehen ist, dass die innerschulischen Akteure von vornherein in der Lage sind, ihre eigene Entwicklung selbst zu steuern, müssen Steuergruppen eingesetzt werden. Dies ist in einem bemerkenswerten Ausmaß auch seit 1989 an deutschen Schulen geschehen: Es gibt Bundesländer wie Hamburg oder NRW, in denen über zwei Drittel aller Schulen über Steuergruppen verfügen (oder verfügten), die für ihre Schulen in Form von Leitbildern bzw. Schulprogrammen weithin akzeptierte Qualitätsziele erarbeiteten und den Prozess der Umsetzung gestalten.

Das ist vermutlich als die größte Schulinnovation der letzten 100 Jahre anzusehen. Denn damit setzt sich ein neues Schulverständnis durch: An Stelle der Schule als bürokratische Lehranstalt tritt jetzt die Idee der »lernenden Schule«.

Qualitätsverständnis

Ausgangspunkt jedes Qualitätsmanagements muss sein zu klären, was unter Qualität überhaupt zu verstehen ist.

Der Begriff Qualität geht zurück auf das lateinische Wort »qualitas«, welches Beschaffenheit, Güte, Werthaltigkeit bedeutet. Für das zeitgemäße Qualitätsmanagement ist Qualität, »was den Anforderungen entspricht« (ISO 9000 ff.). In Wirtschaft und Verwaltung bestimmen die Kunden die Anforderungen. Wer bestimmt sie im Schulbereich? *Schulqualität* ist in erster Linie Lernqualität der Schüler. Aber die Schülerinnen und Schüler sind nicht Kunden, »sondern Koproduzenten des Lernens«, und sie bestimmen auch nicht die Anforderungen. Die Anforderungen entstehen aus:

- staatlichen Lehrplänen,
- gesellschaftlichen und wirtschaftlichen Erwartungen,
- der Wissenschaft,
- der Bildungs- und Erziehungstheorie und
- dem Leitbild/Schulprogramm.

Im Leitbild bestimmen die Schulen – im vorgenannten Rahmen – letztlich selbst, was sie unter Qualität verstehen.

Doch wie können die Schulen wissen, welche Qualität sie haben? Evaluation gibt die Antwort. Evaluation wird verstanden als Prozess des systematischen Sammelns und Analysierens von Daten/Informationen mit dem Ziel, an Kriterien orientierte Bewertungsurteile zu ermöglichen, die begründet und nachvollziehbar sind. Evaluation meint datengestützte Bewertung. Kriterien sind vor allem:

- Vorgaben der Behörde bzw. des Parlaments (auslegungsbedürftig),
- Leitbild/Schulprogramm,
- Projektziele/Projektauftrag/Leistungsauftrag,
- Vergleiche/Benchmarking,
- vereinbarte Qualitätsindikatoren und -standards (»Referenzrahmen«).

Insgesamt liegt hier ein breites Qualitätsverständnis zugrunde. Es bezieht sich auf Input-, Prozess- sowie Output-Faktoren und umfasst Lernleistungen wie Erziehung, Schulmanagement wie Ressourcennutzung, Schulkultur wie Lehrerprofessionalität (vgl. dazu Abb. I.1).

Evaluation: Primär intern, dazu extern

Evaluation spielt also in jedem QM-Modell eine zentrale Rolle. Beim Schulentwicklungsmodell hat interne Evaluation eine Priorität, und externe Evaluation gehört dazu. Mit Priorität ist sowohl zeitlicher Vorrang als auch Wichtigkeit gemeint.

1 Ergebnisse und Erfolge der Schule	2 Lernkultur – Qualität der Lehr- und Lernprozesse	3 Schulkultur	4 Schulmanagement	5 Lehrerprofessionalität und Personalentwicklung	6 Ziele und Strategien der Qualitätsentwicklung
1.1 Persönlichkeitsbildung	2.1 Schuleigenes Curriculum: Ziele und Inhalte	3.1 Soziales Klima in der Schule und in den Klassen	4.1 Schulleitungshandeln und Schulgemeinschaft	5.1 Zielgerichtete Personalentwicklung	6.1 Programm- und Leitbildentwicklung
1.2 Fachkompetenzen	2.2 Lernangebote zur Stärkung der Persönlichkeit	3.2 Lebensraum Schule – Lebensraum Klasse	4.2 Kooperative Wahrnehmung der Gesamtverantwortung	5.2 Weiterentwicklung beruflicher Kompetenzen	6.2 Konzepte und Evaluation der Unterrichts- und Erziehungsarbeit
1.3 Schlüsselqualifikationen	2.3 Unterrichtsgestaltung (Lehrerhandeln im Unterricht)	3.3 Beteiligung der Schüler- und Elternschaft	4.3 Schulleitungshandeln und Qualitätsmanagement	5.3 Lehrerkooperation (Arbeits- und Kommunikationskultur in der Schule)	6.3 Konzepte und Evaluation der Schulentwicklung (Schule als Ganzes)
1.4 Schulabschlüsse und Bildungsweg	2.4 Leistungsanforderungen und Leistungskontrolle	3.4 Öffnung von Schule und Kooperation mit gesellschaftlichen Partnern	4.4 Verwaltungs- und Ressourcenmanagement	5.4 Personaleinsatz der Beschäftigten	6.4 Maßnahmen zum schulübergreifenden Vergleich
1.5 Schulzufriedenheit und Schulimage	2.5 Schülerunterstützung im Lernprozess		4.5 Unterrichtsorganisation	5.5 Gesundheitsförderliche Arbeitsbedingungen	6.5 Dokumentation und Umsetzungsplanung
1.6 Gesamteindruck und Selbstpräsentation der Schule	2.6 Außerunterrichtliche Schülerbeurteilung				6.6 Verbesserung der Schulumfeldsituation und der Rahmenbedingungen

Abb. 1.1: Qualitätstableau; Sechs Qualitätsbereiche und 32 Qualitätsfelder (Quelle: KM-Niedersachsen)

Die Priorität interner Evaluation wird von Schulentwicklungsbegleitern empfohlen und durch Forschungen gestützt. Coleman (1992) und Newman u. a. (1997) haben überzeugende Belege dafür gefunden, dass die interne Evaluation erst den Boden bereitet, auf dem Qualitätsmanagement gedeihen und externe Evaluation wirken kann. Findet externe Evaluation statt, ohne dass in der Schule eine Evaluationskultur entstanden ist, »verpufft« sie: Sie wird von den Lehrpersonen nicht verstanden und schon allein deshalb nicht akzeptiert. Es ist kein Bewusstsein für die Nützlichkeit von Evaluation vorhanden, ebenso wenig wie eine Expertise im Umgang mit Daten.

Erst wenn eine »Kultur der Evaluation« entsteht, vermag Evaluation die erhofften Impulse für Unterrichts-, Personal- und Organisationsentwicklung, also Schulentwicklung zu geben. Lehrpersonen, die sich mit interner Evaluation beschäftigen, werden zu »reflektierenden Praktikern«, Schüler, die den Unterricht evaluieren und zudem ihre eigenen Lernfortschritte einzuschätzen wissen, beteiligen sich aktiv am Qualitätsmanagement, und Schulleiter, die Evaluationsergebnisse ernst nehmen, praktizieren datengestützte Schulführung.

Interne Evaluation ist beim Schulentwicklungsmodell vorrangig, externe notwendig. Die externe Evaluation, die der internen folgt, wird in der Regel als Inspektion (Sparka 2001, Maritzen 2016) durchgeführt.

Steuerung und Koordination: Kriterien- und indikatorenorientiert

Schulentwicklung supponiert ein dezentrales Qualitätsmanagement: Erst einmal müssen die Schulen selbstverantwortlich handeln (und dafür einen geeigneten Rechtsrahmen vorfinden), dann erst stellt sich die Frage der Koordination von Selbststeuerung und zentraler Steuerung. Der Prozess der »Verselbstständigung« von Schulen ist ein Lernprozess: Er benötigt Experimentierräume, Zeit, viel Unterstützung und für Selbstkontrolle auch Evaluation. Er ist ein aufwendiger Lernprozess. Er verlangt auch, dass zentrale Behörden Entscheidungsbefugnisse an die Schulen abgeben und Behörden nicht mehr in die Schulen »hineinregieren« (vgl. Rolff/Schmidt 2002).

Aber »Verselbstständigung« von Schulen ist ambivalent. Sie könnte auch auf das hinauslaufen, was die amerikanische Literatur »Balkanisierung« nennt: Jede Schule macht, was sie will, und sie verhält sich zudem aggressiv gegenüber jenen Schulen, die sie als Konkurrenten empfindet (Hargreaves 2000).

Deshalb ist es wichtig, die innerschulischen wie die außerschulischen Steuerungsmedien konklusiv aufeinander abzustimmen. Das ist in etlichen Ländern wie Schottland und den Niederlanden von vornherein der Fall gewesen, geschieht aber in anderen – wie in Deutschland – erst relativ spät, sozusagen im Nachhinein. Voraussetzung für eine GesamtAbstimmung ist ein Referenzsystem, auf das sich die innerschulische Qualitätsevaluation genauso beziehen kann wie die außerschulische. Sonst laufen die Aktivitäten des Qualitätsmanagements aneinander vorbei, und es entsteht keine Synergie. Ein Referenzsystem sollte von bedeutsamen Qualitätsbereichen ausgehen, diese

nach Qualitätsfeldern ausdifferenzieren und für jedes Feld Kriterien bzw. Standards formulieren, die anhand von Indikatoren überprüft werden können. Vor allem sollten sie sowohl von Schulen als auch von Behörden akzeptiert werden.

13. Steuerung der Q-Prozesse durch die Schulleitung	V. Qualitätsmanagement	1. Schulische Rahmenvorgaben und strategische Vereinbarungen		I. Input-Qualitäten
		2. Personelle und strukturelle Voraussetzungen		
		3. Materielle und finanzielle Ressourcen		
14. Praxis des Individualfeedbacks und der individuellen Q-Entwicklung		4. Schulführung	II. Schule	II./III. Prozessqualitäten
		5. Schulorganisation und Schuladministration		
		6. Kollegiale Zusammenarbeit und Schulkultur		
		7. Lehr- und Lernarrangement	III. Unterricht	
		8. Soziale Beziehungen		
		9. Prüfen und Beurteilen		
15. Praxis der Schulevaluation und der Schulentwicklung		10. Zufriedenheit der Leistungsempfängenden		IV. Output-/ Outcome-Qualitäten
		11. Lern- und Sozialisationsergebnisse		
		12. Schul- und Laufbahnerfolg		

Abb. I.2: Basisinstrument: Qualitätsbereiche und Dimensionen nach Q2E

Unter den deutschen Bundesländern verfügte als erstes Niedersachsen über ein solches Referenzsystem, »Orientierungsrahmen Schulqualität in Niedersachsen« genannt. Der niedersächsische Orientierungsrahmen für Qualitätsmanagement besteht aus folgenden Bausteinen:

- 6 Qualitätsbereiche,
- 32 Qualitätsfelder (vgl. Abb. I.1),
- 86 Kriterien und
- zahlreichen Hinweisen zu Indikatoren und anderen Instrumenten und Methoden der Qualitätsevaluation.

Abbildung I.1 zeigt die Qualitätsbereiche und -felder des niedersächsischen Referenzrahmens, der die Gestalt eines Qualitätstableaus hat. Zu jedem Qualitätsfeld gehören zwei bis vier Kriterien, die so etwas Ähnliches wie Qualitätsziele darstellen.

Die Kriterien des niedersächsischen Qualitätstableaus müssen anhand von Indikatoren evaluierbar, d. h. überprüfbar gemacht werden. Diese Indikatoren können die Schulen selber konstruieren oder aus einer Indikatorensammlung übernehmen, die den meisten Qualitätsrahmen beigelegt ist. Das Erreichen der Kriterien wird auf einer vierstufigen Skala gemessen, die ähnlich definiert ist wie die Stufen in Schottland, den Niederlanden oder auch bei Q2E (siehe nächste Seite).

Für die drei unmittelbar lernerfolgsbezogenen Q-Felder 1.2 Fachkompetenzen, 1.3 Schlüsselqualifikationen und 2.1 Schuleigenes Curriculum (Ziele und Inhalte) hat die deutsche Kultusministerkonferenz (KMK) der Länder Standards formuliert, an denen sich die Schulen der Sekundarstufe I orientieren müssen (siehe unten).

In der deutschsprachigen Schweiz wurde ein Q-Modell entwickelt, das ebenfalls von der Einzelschule ausgeht und der internen Evaluation Priorität gibt. Es wird Q2E genannt als Kürzel von »Qualität durch Evaluation und Entwicklung« (Landwehr/Steiner 2002). Es bietet ebenfalls einen elaborierten Referenzrahmen für Schulqualität in Form eines sogenannten Basisinstruments (vgl. Abb. I.2).

Das Basisinstrument verfügt über zwölf Qualitätsdimensionen von »1. Schulischen Rahmenvorgaben und strategischen Vereinbarungen« bis »12. Schul- und Laufbahnerfolg«. Sie sind nach dem QM-üblichen Ablaufschema nach Input-Prozess- und Output-/Outcome-Qualitäten gegliedert, wobei die Prozessqualitäten wiederum nach Schule und Unterricht ausdifferenziert sind. Zu jeder Dimension sind Kriterien bzw. Indikatoren angegeben, an denen man Qualität erkennen und bewerten kann. Sie werden bei Q2E Qualitätsmerkmale genannt.

Abbildung I.3 zeigt eine Auswahl solcher Kriterien/Indikatoren für die Unterrichtsentwicklung zentraler Qualitätsdimensionen »7. Lehr- und Lernarrangement«. Die Aussagen sind z. T. als Kriterien für Qualität in diesem Bereich zu verstehen, z. T. als Indikatoren, die anzeigen, woran man Qualität für diesen Bereich erkennen kann. Q2E versteht Indikatoren als »Hinweise darauf, dass bestimmte qualitätsbestimmende Werte vorhanden sind, gelebt und umgesetzt werden« (Bd. 4, S. 43).

Das Basisinstrument umfasst schließlich noch drei Bereiche des Qualitätsmanagements: I. Steuerung der Q-Prozesse durch die Schulleitung, II. Praxis des Individualfeedbacks und der individuellen Q-Entwicklung und III. Praxis der Schulevaluation und der Schulentwicklung. Für diese Bereiche des Qualitätsmanagements sind wiederum Kriterien/Indikatoren angegeben und – wie gehabt – vielfach gestuft, sodass sogenannte Bewertungstabellen entstehen.

Abbildung I.4 zeigt eine solche Bewertungstabelle für den Q-Bereich »Steuerung der Q-Prozesse durch die Schulleitung«.

Der Begriff des Standards bezieht sich auf die Stufen solcher Bewertungstabellen. In der Q2E-Bewertungssystematik werden – wie international üblich – vier Qualitätsstufen unterschieden, die jedem Kriterium/Indikator zugeordnet am rechten Rand von Abbildung I.3 zu ersehen sind. Sie werden beim Q2E-Ansatz präzise wie folgt definiert:

- **Stufe 1:** Wenig entwickelte Praxis: Die Praxis befindet sich an einem defizitären, unbefriedigenden Zustand. Es gibt einen dringenden Handlungsbedarf.
- **Stufe 2:** Funktionsfähige Praxis: Die grundlegenden Anforderungen sind erfüllt; es gibt aber noch viele Verbesserungsmöglichkeiten.
- **Stufe 3:** Differenzierte, gut entwickelte Praxis: Die Praxis ist auf gutem Niveau etabliert; die »üblichen« Erwartungen werden erfüllt.
- **Stufe 4:** Exzellente Praxis: Die Praxis weist eine hohe situative Passung auf, d. h., sie wird flexibel den unterschiedlichen Situationen angepasst, um eine optimale Wirkung zu erreichen. Gleichzeitig ist eine nachhaltige Qualitätsentwicklung im Sinne einer lernenden Organisation gewährleistet (Bd. 5, S. 34).

c) Gestaltung der Lehr- und Lernprozesse (methodisch-didaktisches Arrangement)		1	2	3	4
1	Die Lehrperson legt Wert darauf, dass die Schülerinnen und Schüler Einsicht erhalten in die Ziele und Absichten des Unterrichts sowie in die Bedeutung der Lernziele und Inhalte.				
2	Der Lehrperson gelingt es, komplexe Lernprozesse unter Berücksichtigung der Erfahrungen und des Vorwissens aufzubauen und schwierige Sachverhalte in einer schülergerechten Sprache zu erläutern.				
3	Es gelingt den Lehrpersonen, das Interesse der Schülerinnen und Schüler für die Unterrichtsthemen zu wecken und eine konzentrierte/engagierte Mitarbeit herbeizuführen.				

Abb. I.3: Kriterien/Indikatoren zur 7. Q-Dimension

»Innerhalb der drei Bewertungstabellen zum Qualitätsmanagement kommt der Bewertungsstufe 3 eine besondere Bedeutung zu: Sie wird im Rahmen des Projekts Q2E als »Standard-Stufe« für die Zertifizierung angenommen. Dies bedeutet, dass die in Stufe 3 formulierten Qualitätsaussagen als Soll-Vorgaben betrachtet werden, die von der Schule zum überwiegenden Teil (70 Prozent) erreicht werden sollten« (Landwehr/Steiner 2002, Bd. 5, S. 25).

Die Hinweise auf 70 Prozent bringen allerdings eine Unklarheit in den Referenzrahmen, da nicht klar ist, worauf er sich genau bezieht.

Die genannten zwölf Q-Dimensionen sollen einen Bezugsrahmen für die schulinterne Evaluation und die Formulierung schulspezifischer Qualitätsziele bieten. Die

	Stufe 1 Wenig entwickelte (defizitäre) Praxis	Stufe 2 Grundlegende An- forderungen an eine funktionsfähige Pra- xis sind erfüllt	Stufe 3 Praxis ist auf gutem Niveau entwickelt	Stufe 4 Exzellente Praxis mit hoher situativer Pas- sung und Optimie- rungsschleufe
Praxisgestaltung	<p>Es gibt kein eigent-liches Qualitätsmanage-ment an der Schule. Einzelne Maßnahmen der Q-Entwicklung und Q-Sicherung (z. B. Mitarbeitergespräch, Schulbesuche des Inspektorats, Fortbil-dungsmaßnahmen) stehen unverbunden und relativ zufällig nebeneinander.</p> <p>Die Schulleitung ver-fügt über wenig Wissen und Know-how bzgl. der Umsetzung eines wirksamen, systema-tischen Qualitätsma-nagements sowie bzgl. ihrer eigenen Rolle in diesem Prozess.</p> <p>Der Aufbau eines sys-tematischen Qualitäts-managements ist aus Sicht der Schulleitung keine notwendige bzw. erstrebenswerte Ziel-setzung »Keine Zeit«; »Es gibt Wichtigeres zu tun«. Q-Maßnahmen werden nicht als zum Schulleitungsauftrag gehörend wahrgenom-men.</p> <p>Die Instrumente, die als Bestandteile der Q-Sicherung zum Einsatz kommen, werden schematisch angewandt und blei-ben weitgehend unre-flektiert.</p>	<p>Ansätze eines systema-tischen Qualitätsma-nagements zur Förde-rung der Schul- und Unterrichtsqualität so-wie zur Förderung der einzelnen Mitarbeiten-den sind vorhanden.</p> <p>Die Schulleitung hat Einblick in die Anfor-derungen an ein schuli-sches Q-Management. Sie baut in der Schule gezielt Q-Manage-ment-Know-how (In-strumente, Verfahren bzw. entsprechende Erfahrungen) auf.</p> <p>Die Schulleitung ini-tiiert und unterstützt schulische und indi-viduelle Q-Projekte. Sie sorgt dafür, dass Fragen der Schul- und Unterrichtsqualität thematisiert werden und stellt Ressourcen für den Q-Manage-ment-Aufbau bereit.</p> <p><i>Abgrenzung zu 3: Qua-litätsprozesse besitzen für die Schulleitung »Projektstatus«, d. h., es sind Sondermaßnah-men, die noch nicht in den Schulführungspro-zess integriert sind.</i></p> <p>Der Aufbau des Q-Managements wird als eine von außen ini-tiierte Notwendigkeit gesehen. (Man muss etwas im Bereich Q-Management tun, um den Erwartungen der vorgesetzten Behör-den, der öffentlichen Meinung, anderer Schulen usw. zu ent-sprechen.)</p>	<p>Die Schulleitung be-trachtet den Aufbau und die Umsetzung eines funktionsfä-higen und wirksamen Q-Managements als einen wichtigen Aspekt der Schulführungsauf-gabe.</p> <p>Die Schulleitung verfügt über ein dif-ferenziertes Wissen (konzeptionelles Wis-sen und praktisches Know-how), das eine adäquate Umsetzung eines ganzheitlichen Q-Managements er-möglicht.</p> <p>Die Schulleitung ver-fügt über geeignete In-strumente und Verfah-ren, um gravierende Qualitätsdefizite recht-zeitig zu erkennen und wirksam anzugehen.</p> <p>Die Schulleitung vertritt die Anliegen des Qualitätsmanagements mit dem not-wendigen Gewicht gegenüber der Lehrerschaft und dem übrigen Schulper-sonal. Sie sorgt dafür, dass die Instrumente und Verfahren der Qualitätssicherung und -entwicklung für alle Beteiligten transpa-rent sind und mit der notwendigen Verbind-lichkeit und Seriosität umgesetzt werden.</p> <p style="text-align: center;">↑ Standardstufe</p>	<p>Die Schule verfügt über ein professionel-les Q-Management mit kreativen Lösungen für die schulspezifischen Anliegen.</p> <p>Die Schulleitung hat das QM-Konzept ver-innerlicht. Die Umset-zung erfolgt nicht re-zepthaft-stur, sondern unter Einbezug der Gegebenheiten und als Ableitung von all-gemeinen Prinzipien. Die Schulleitung führt das Q-Management situationsadäquat; sie kann im Rahmen des Q-Konzepts varian-tenreich und flexibel reagieren.</p> <p>Es liegt ein vielfäl-tiges Repertoire an Instrumenten und Erfahrungen des Q-Managements vor. Die Q-Instrumente, die von der Schule eingesetzt werden, berücksichti-gen die Möglichkeiten und die Eigenheiten der Schule in einer differenzierten Weise. (Sie sind »maßge-schneidert« mit Bezug auf das spezifische Schulprofil.)</p> <p>Die laufenden Über-prüfungen und Opti-mierungen des Quali-tätsmanagements sind ein spürbares Anliegen der Schulleitung und werden von ihr auch umgesetzt.</p>

Abb. I.4: Bewertungstabelle zur Steuerung des QM durch die Schulleitung

zuletzt aufgeführten drei Q-Bereiche dienen als Referenzrahmen für die externe Evaluation bzw. die Zertifizierung.

Die Referenzrahmen dienen bei beiden Modellen als zentrales Steuerungsmedium, an das sich Schulen wie Behörden halten müssen. Sie definieren beide das Qualitätsverständnis relativ breit. Dazu gehören Fachkompetenzen und Schlüsselqualifikationen, aber auch die Lehr- und Lernformen, die Schulkultur, die Lehrerprofessionalität oder das Schulmanagement.

Der wesentliche Unterschied liegt darin, das beim niedersächsischen Modell die externen Evaluatoren die Schule, vor allem auch den Unterricht direkt evaluieren, während beim Schweizer Modell die innerschulischen Q-Dimensionen in direkter Weise nur schulintern evaluiert werden: Die externe Evaluation bezieht sich nicht auf die Q-Dimensionen selbst, sondern darauf, wie professionell die Schulleitung und die Lehrerschaft diese Evaluation gestalten und ob sie eine Feedbackkultur aufgebaut haben. Beim niedersächsischen Modell handelt es sich also um externe Direktevaluation und beim Schweizer Modell – genau wie beim Qualitätsmanagement in der Wirtschaft (ISO, EFQM) – um Meta-Evaluation (um Evaluation der Evaluation).

Zu erwähnen ist allerdings, dass das Schweizer Modell ein »Fenster« für direkte Evaluation durch Externe vorgesehen hat, die dort primäre Evaluation genannt wird. Sie ist allerdings für Schulen nicht verpflichtend, sondern optional (Landwehr/Steiner 2002, Bd. 5, S. 31).

Implementationsprobleme

Über die Wirkungen beider Modelle auf die tatsächliche Qualität von Schule (und hier auf die Qualität des Unterrichts) ist bisher nicht geforscht worden. Dieser Frage vorgelagert, stellt sich das Implementationsproblem: Wie und wie weit sind die Modelle des Qualitätsmanagements in Schulen überhaupt umgesetzt worden, in »Reinform« oder abgewandelt, insgesamt oder nur teilweise, schulweit oder nicht in einzelnen Jahrgruppen, Fachschaften oder Abteilungen? Hierüber ist wenig bekannt, außer dass die Zahl der infrage kommenden Schulen wächst. In der Schweiz haben seit Anfang 2004 zahlreiche Berufsschulen und einige Gymnasien das Q2E-Modell umgesetzt. Das niedersächsische Modell ist nur annähernd komplett umgesetzt worden.

Die allgemeinen Bedingungen von Implementation haben Huberman und Miles in ihren richtungsweisenden Untersuchungen bereits 1984 herausgearbeitet. Demnach gelingt eine Implementation am ehesten in einem Kontext, der durch eine Kombination von »pressure and support« gekennzeichnet ist. In der deutschen Diskussion hat sich die Formel Druck und Zug eingebürgert, die nicht ganz so präzise ist, weil das Moment der Unterstützung nicht benannt wird.

Aus ersten Evaluationsberichten (Bastian/Rolff 2002 und Rolff u. a. 2004) wissen wir, dass

- sich nur wenige Schulen freiwillig einer externen Evaluation stellen,
- sich interne Evaluation in Schulen verbreitet,
- zumeist die Evaluation von Unterricht vermieden wird, schon gar von Fachunterricht,
- die Daten der Öffentlichkeit am liebsten vorenthalten werden, sogar der Schulöffentlichkeit,
- dort, wo externe Evaluation verpflichtend ist, die interne intensiviert wird,
- Evaluation insgesamt angstbesetzt ist und auf ernsthafte Vorbehalte stößt,
- konkrete Erfahrungen mit Evaluation, vor allem klare Regeln und Normen zum Umgang mit Daten geeignet sind, diese Ängste und Vorbehalte in beachtlicher Weise zu mindern (vgl. Müller 2002).

2. Zentral administrierte Entwicklung

Der Grundtyp B der Steuerung der Qualität von Schulen geht zurück auf die großflächigen Leistungsuntersuchungen, deren Karriere in Deutschland 1995 mit der Veröffentlichung der TIMSS-Ergebnisse begann und sich mit PISA und IGLU fortsetzt. Diese international Large-Scale-Assessments genannten Untersuchungen dienen dem zentralen System-Monitoring. Vor ihrem Hintergrund hat die KMK erstmals am 04.12.2003 bundesweit gültige und insofern »nationale« Bildungsstandards beschlossen, die die Qualität des Unterrichts sichern und der Weiterentwicklung dienen sollen. Diese Standards wurden zentral entwickelt, und sie werden von den Länderregierungen zentral administriert.

Initiierung durch Standards

Bildungsstandards nach dem Verständnis der KMK beschreiben erwartete Lernergebnisse. Sie sind so etwas wie »nationale Kompetenzerwartungen« (Helmke).

Bildungsstandards greifen allgemeine Bildungsziele auf und benennen Kompetenzen, die Schülerinnen und Schüler bis zu einer bestimmten Jahrgangsstufe an zentralen Inhalten erworben haben sollen. Sie konzentrieren sich auf Kernbereiche eines Faches. Damit wird die bisherige Orientierung an Inhalten und Qualifikationen aufgegeben. Dies ist ein weltweiter Trend, der durch die international vergleichenden Schülerleistungstests TIMSS und PISA ausgelöst wurde.

Bildungsstandards nach dem vorherrschenden Verständnis formulieren fachliche und fachübergreifende Kompetenzen, die für die weitere schulische und berufliche Ausbildung von Bedeutung sind und die anschlussfähiges Lernen ermöglichen. Die Standards der KMK lagen zunächst für den Schulabschluss am Ende der Pflichtschulzeit und für die Fächer Deutsch, Mathematik und Englisch vor. Weitere Fächer folgten ebenso wie Standards für das Ende der (vierjährigen) Grundschulzeit.

Die nationalen Bildungsstandards sollten zunächst als Mindeststandards definiert werden, wurden aber letztlich als sogenannte Regelstandards beschlossen. Sie greifen die Grundprinzipien des jeweiligen Unterrichtsfaches auf. Die nationalen Bildungsstandards:

- beschreiben die fachbezogenen Kompetenzen einschließlich zugrunde liegender Wissensbestände, die Schülerinnen und Schüler bis zu einem bestimmten Zeitpunkt ihres Bildungsganges erreicht haben sollten;

- zielen auf systematisches und vernetztes Lernen und folgen so dem Prinzip des kumulativen Kompetenzerwerbs;
- beschreiben erwartete Leistungen im Rahmen von Anforderungsbereichen;
- beziehen sich auf den Kernbereich des jeweiligen Faches und geben den Schulen Gestaltungsräume für ihre pädagogische Arbeit;
- weisen ein mittleres Anforderungsniveau aus;
- werden durch Aufgabenbeispiele veranschaulicht.

Diese Standards basieren auf vorläufigen fachspezifisch definierten Kompetenzmodellen, die aus der Erfahrung der Schulpraxis heraus entwickelt wurden. Sie beziehen international anerkannte Standardmodelle ein – unter anderem theoretische Grundlagen der PISA-Studie und den gemeinsamen europäischen Referenzrahmen für Sprachen.

Aufgabenbeispiele veranschaulichen die fachlichen Bildungsstandards. Die deutschen Bundesländer sind übereingekommen, weitere Aufgabenbeispiele zu entwickeln und in landesweiten bzw. länderübergreifenden Orientierungs- und Vergleichsarbeiten oder in zentralen oder dezentralen Prüfungen festzustellen, in welchem Umfang die Standards erreicht werden. Das Prinzip kompetenzorientierter Standards soll am Fach Mathematik illustriert werden:

Beispiel Mathematik

Mit dem Erwerb des Schulabschlusses sollen Schülerinnen und Schüler über die nachfolgend genannten allgemeinen mathematischen Kompetenzen verfügen, die für alle Ebenen des mathematischen Arbeitens relevant sind. Diese Kompetenzen werden immer im Verbund erworben bzw. angewendet. Sie sind aus Abbildung I.5 zu ersehen.

Einige der oben benannten mathematischen Kompetenzen werden im Folgenden beispielhaft konkretisiert:

(K 1) Mathematisch argumentieren

Dazu gehört:

- Fragen stellen, die für die Mathematik charakteristisch sind (»Gibt es ...?«, »Wie verändert sich ...?«, »Ist das immer so ...?«), und Vermutungen begründet äußern, mathematische Argumentationen entwickeln (wie Erläuterungen, Begründungen, Beweise), Lösungswege beschreiben und begründen.

(K 2) Probleme mathematisch lösen

Dazu gehört:

- Vorgegebene und selbst formulierte Probleme bearbeiten, geeignete heuristische Hilfsmittel, Strategien und Prinzipien zum Problemlösen auswählen und anwenden, die Plausibilität der Ergebnisse überprüfen sowie das Finden von Lösungsideen und die Lösungswege reflektieren.

Mathematische Leitideen

Die KMK betont, dass die oben beschriebenen allgemeinen mathematischen Kompetenzen von Schülerinnen und Schülern in der Auseinandersetzung mit mathematischen Inhalten erworben werden. Insofern bleiben Inhalte wichtig. Dementsprechend werden die allgemeinen mathematischen Kompetenzen als Dispositionen von Schülerinnen und Schülern inhaltsbezogen konkretisiert.

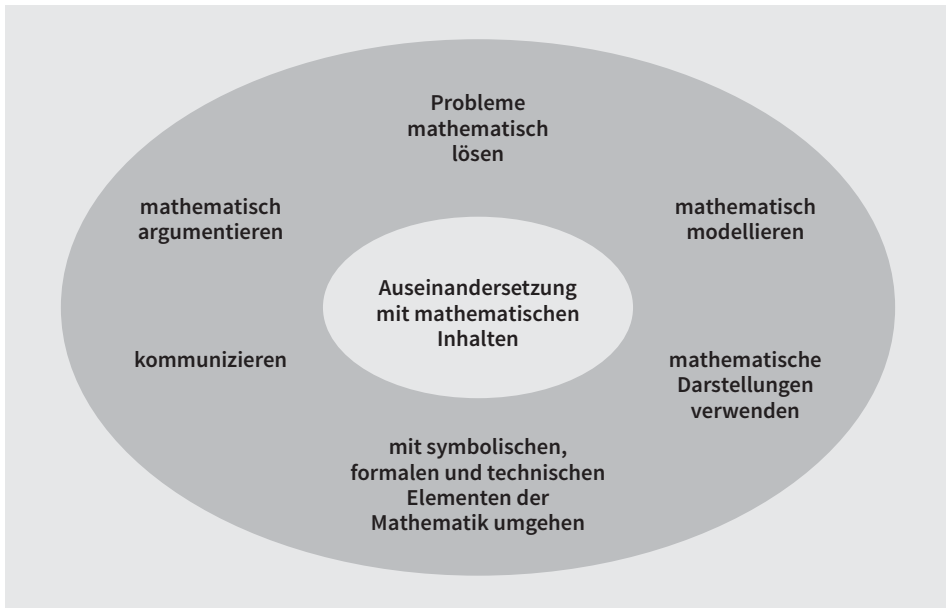


Abb. 1.5: Allgemeine mathematische Kompetenzen

Solche Standards für inhaltsbezogene mathematische Kompetenzen sind jeweils ausgewählten mathematischen Leitideen zugeordnet, um Verständnis von grundlegenden mathematischen Konzepten zu erreichen, Besonderheiten mathematischen Denkens zu verdeutlichen sowie Bedeutung und Funktion der Mathematik für die Gestaltung und Erkenntnis der Welt erfahren zu lassen. Folgende mathematische Leitideen sind zugrunde gelegt:

- Zahl,
- Messen,
- Raum und Form,
- funktionaler Zusammenhang,
- Daten und Zufall.

Eine Leitidee vereinigt Inhalte verschiedener mathematischer Sachgebiete und durchzieht ein mathematisches Curriculum spiralförmig.

Die Zuordnung einer inhaltsbezogenen mathematischen Kompetenz zu einer mathematischen Leitidee ist nicht in jedem Fall eindeutig, sondern davon abhängig, welcher Aspekt mathematischen Arbeitens im inhaltlichen Zusammenhang betont werden soll.

Dies wird wiederum für jede Leitidee verdeutlicht:

(L 1) Leitidee Zahl

Die Schülerinnen und Schüler

- nutzen sinntragende Vorstellungen von rationalen Zahlen, insbesondere von natürlichen, ganzen und gebrochenen Zahlen entsprechend der Verwendungsnotwendigkeit;
- stellen Zahlen der Situation angemessen vor, unter anderem in Zehnerpotenzschreibweise;
- begründen die Notwendigkeit von Zahlbereichserweiterungen an Beispielen, nutzen Rechengesetze, auch zum vorteilhaften Rechnen, nutzen zur Kontrolle Überschlagsrechnungen und andere Verfahren;
- runden Rechenergebnisse entsprechend dem Sachverhalt sinnvoll;
- verwenden Prozent- und Zinsrechnung sachgerecht;
- erläutern an Beispielen den Zusammenhang zwischen Rechenoperationen und deren Umkehrungen und nutzen diese Zusammenhänge;
- wählen, beschreiben und bewerten Vorgehensweisen und Verfahren, deren Algorithmen bzw. Kalküle zugrunde liegen;
- führen in konkreten Situationen kombinatorische Überlegungen durch, um die Anzahl der jeweiligen Möglichkeiten zu bestimmen;
- prüfen und interpretieren Ergebnisse in Sachsituationen unter Einbeziehung einer kritischen Einschätzung des gewählten Modells und seiner Bearbeitung.

Für jeden Standard werden – wie in den internationalen Studien TIMSS und PISA – Kompetenzniveaus definiert. Bei den internationalen Studien geschah das aufgrund empirischer Daten mithilfe anspruchsvoller Skalierungsverfahren. Dabei ergaben sich zumeist vier oder fünf Stufen. Die KMK nahm die Stufung nicht auf empirischer Basis, sondern anhand von Einschätzungen eines Gremiums von Fachdidaktikern vor, wobei immer das gleiche Dreistufen-Schema benutzt wurde (vgl. Abb. I.6).

Die an Kompetenzstufen orientierten Standards sollen regelmäßig durch Tests überprüft werden. Die Testorientierung geht so weit, dass sich die Proponenten dieses Modells Standards ohne diesbezügliche Tests ebenso wie Tests ohne Standards gar nicht vorstellen können. Standards und Tests werden als Zwillinge gedacht (vgl. z. B. Klieme u. a. 2003). Deshalb kann auch von einem Teststandard-Modell gesprochen werden.

I. Reproduzieren	II. Zusammenhänge herstellen	III. Verallgemeinern und Reflektieren
Mathematisch argumentieren		
Dazu gehört:		
<ul style="list-style-type: none"> • Routineargumentationen wiedergeben (wie Rechnungen, Verfahren, Herleitungen, Sätze, die aus dem Unterricht vertraut sind). • Mit Alltagswissen argumentieren. 	<ul style="list-style-type: none"> • Überschaubare mehrschrittige Argumentationen erläutern oder entwickeln. • Lösungswege beschreiben und begründen. • Ergebnisse bzgl. ihres Anwendungskontextes bewerten. • Zusammenhänge, Ordnungen und Strukturen erläutern. 	<ul style="list-style-type: none"> • Komplexe Argumentationen erläutern oder entwickeln. • Verschiedene Argumentationen bewerten. • Fragen stellen, die für die Mathematik charakteristisch sind, und Vermutungen begründet äußern.

Abb. 1.6: Anforderungsbereiche/Kompetenzstufen

Qualitätsverständnis

Das Qualitätsverständnis des Teststandard-Modells ist betont fachspezifisch. Deutsch-, Mathematik- und Englischkompetenzen sowie die dazugehörigen Wissensdomänen definieren, was unter Qualität zu verstehen ist. Andere Qualitätsdimensionen, wie sie das Schulentwicklungsmodell als Referenzrahmen benutzt, sind bestenfalls unterbelichtet, kommen letztlich gar nicht vor. Schulqualität ist Fachqualität. Auch wenn dieses Qualitätsverständnis im Kern plausibel ist, birgt es doch unübersehbar die Gefahr eines pädagogischen Reduktionismus. Hartmut v. Hentig spricht in diesem Zusammenhang von einer Drei-Fächer-Schule (Hentig 2003, S. V 30). Zudem regeln Standards in erster Linie, was gelernt werden soll; das Wie, welches die Reformschulen besonders betonen, tritt wieder in den Hintergrund.

Fraglos sind Standards für die Qualitätssicherung vonnöten. Problematisch ist jedoch die Verengung der Standards auf die drei »Hauptfächer« und die Verschweigung der Standards mit hoch voraussetzungsvollen Tests. Denn die Testkonstruktion wird letztlich durch skalierungstechnische Notwendigkeiten geleitet. Nicht was Schüler lernen und Lehrer lehren sollen, sondern was sich auf Kompetenzskalen abbilden lässt, könnte über die Auslegung von Standards entscheiden. Das wirkt mehrfach restriktiv: Es fällt tendenziell heraus, was dem kognitionswissenschaftlichen Kompetenzbegriff nicht genügt, was zu wenig fachlich gerahmt ist, und letztlich sogar, was sich nicht auf Stufenskalen abbilden lässt, also beispielsweise Emotionales, Musisches und Ästhetisches. Wenn man die Stufenlogik wirklich ernst nähme, dann blieben äußerst wenige Lernbereiche für Teststandards übrig. Behauptet diese doch, dass sich Lernen in klar abgrenzbarer, aufeinander aufbauender, inkludierender, einer festgelegten Reihenfolge gehorchenden Stufung vollzieht. Das ist selbst bei

der am meisten etablierten Stufenabfolge, Piagets Modell der geistigen Entwicklung, nicht unumstritten.

Hinzu kommt, dass Lehrkräfte nicht in der Lage sind, derart anspruchsvolle Kompetenztests selbst auszuwerten. Das führt zu einer unvermeidlichen Abhängigkeit von Testinstituten. Das heißt vor allem, dass Lehrkräfte die Ergebnissicherung ihres Unterrichts nicht mehr selbst in der Hand haben. Es werden ihnen relativ abstrakte Ergebnisse zurückgemeldet, nämlich die Position von Schülergruppen auf Kompetenzniveaus. Die Frage, welche Inhalte im Unterricht vermittelt wurden und welche Methoden dabei wirksam waren, bleibt im Dunkeln. Auch Notwendigkeiten der Geheimhaltung der Testaufgaben und die riesigen Kosten für Neuentwicklung halten die Testkonstrukteure davon ab, konkrete Informationen an die Lehrkräfte zu geben. Das führt eher zu einer Deprofessionalisierung als zur versprochenen Professionalisierung. Lehrer kommen kaum umhin, sich nach den Beispielaufgaben der Tests zu richten, sodass »teaching to the test« die Regel wird.

Problematisch ist ferner, dass die Normierung von Teststandards eine Art eingebaute Tendenz zur Überhöhung hat. Qualität angesichts von Regelstandards ist hohe Qualität. Vermutlich legen Politiker auch Mindeststandards (»80 Prozent auf Niveau IV«) fest. Bayern z. B. könnte alle anderen übertreffen wollen, und niemand möchte – und kann – hinter Bayern zurückstehen. Robert Marzano hat herausgefunden, dass in den USA die Schüler neun Jahre länger zur Schule gehen müssten, wenn sie die in einigen Staaten geltenden Standards erreichen wollten (Educational Leadership 2003).

Ein weiteres Problem liegt in der Versuchung, die Ergebnisse von Tests für Ausleseentscheidungen zu nutzen, also die Schullaufbahn oder den Schulabschluss von Schülern an Testergebnisse zu binden. Tests, die für Laufbahnentscheidungen genutzt werden, heißen in den USA »High Stakes Tests«. Sie werden dort in 18 Bundesstaaten eingesetzt. Sie ähneln den empfohlenen Teststandards sehr. Erste empirische Untersuchungen (vgl. Februar-Heft 2003 von »Educational Leadership«), zeigen übereinstimmend:

- Teststandards messen nicht, was sie vorgeben (schreibt Popham),
- schwache Schulen und schwache Schüler werden noch schwächer, und
- Teststandards können mehr demotivieren als motivieren (Amrein/Berliner).

Evaluation: Primär extern, dazu Rückmeldungen

Die Teststandards werden – definitionsgemäß – regelmäßig durch Tests überprüft. Die KMK hat beschlossen, dass zu diesem Zwecke eigens eine »Testagentur der Länder« gegründet wird. Deren Hauptaufgabe besteht darin, Anker-Items für die Normierung länderspezifischer Tests auf einer gemeinsamen (nationalen) Metrik bereitzustellen. Damit hat sie im Jahre 2006 anhand der dann verfügbaren bundesweit repräsentativen PISA-III-Daten begonnen, und zwar zunächst nur für Mathematik. Andere Fächer folgten später.

Externe Evaluation hat also eindeutig Priorität. Die bei diesen bundesweit, z. T. auch landesweit repräsentativen Stichprobenerhebungen anfallenden Daten werden den Schulen zurückgemeldet, wie es bei PISA I (PISA 2000) und IGLU bereits geschehen ist. Die Frage ist, wie die Schulen mit diesen Daten umgehen und ob bzw. wie sie sie für ihr eigenes Qualitätsmanagement nutzen. Diese Frage ist durch eine Studie von Peek in Brandenburg und von einer kleineren Studie in Hessen (Schwarz/Steffens 2003) untersucht worden. Diese Studien und Auskünfte der bei der Rückmeldung einbezogenen Landesinstitute lassen vermuten, dass zwar die meisten Schulen eine Datenrückmeldung anfordern, aber nur die wenigsten (ca. 20 Prozent) die Daten auch nutzen. Das mag daran liegen, dass die Schulen in Deutschland bisher noch keine Erfahrungen mit datenbasierter Schulentwicklung gemacht haben oder dass die Daten nicht repräsentativ für die Schule und überhaupt schwer zu interpretieren sind. Die Hessen-Studie zeigt bemerkenswerterweise, dass der Einsatz von Beratern, die den Schulen bei der Dateninterpretation helfen, die Intensität der innerschulischen Verarbeitung erhöht.

Der Hauptgrund für die Rückmeldeabstinz der Schulen liegt vermutlich darin, dass sich die Schulen mit den Daten aus Large-Scale-Assessments kaum identifizieren: Sie stammen nicht – wie bei interner Evaluation – aus schuleigenen Fragestellungen, die Repräsentativität für die Schule ist nicht gesichert, und manchmal gehen die Daten nur an die betroffenen Klassenlehrer, die diese für sich behalten, sodass sie nicht schulöffentlich werden.

Steuerung und Koordination: Standard- und testorientiert

Die Steuerung erfolgt zentral: Die Standards sind zentral erarbeitet und zentral vereinbart. Sie können nicht zentral, sondern nur in den Ländern realisiert werden. Die Länder haben sich verpflichtet, ihre Lehrpläne darauf abzustimmen. Die Bildungsstandards sind nach einer längeren kontroversen Debatte nicht schulformspezifisch, sondern schulformübergreifend formuliert. Die meisten Länder von Baden-Württemberg bis Nordrhein-Westfalen setzen sie allerdings in Form schulformspezifischer Kernlehrpläne um. Diese sollen die Lehrplanentwicklung der Einzelschulen steuern.

Ein Koordinierungsproblem entsteht insofern nicht, als die Schulen bei der Konzipierung der Standards nicht beteiligt waren. Sie wurden zentral erstellt und top-down an die Schulen weitergegeben.

Ähnlich verhält es sich mit dem Bund-Länder-Kommissionsprogramm Sinus, das durch Chik ergänzt wird. Beide Programme sind im Anschluss an TIMSS entstanden und intendieren eine Qualitätsverbesserung des mathematischen und naturwissenschaftlichen Unterrichts; Chik ist auf Chemie spezialisiert. Beide Programme orientieren sich an den (impliziten) Teststandards von TIMSS. Sie stellen zentral entwickelte, unterschiedliche Module (z. B. neue Aufgabenkultur) dar, die in Form von Modellvorhaben von Schulen ausprobiert wurden. Die Entwicklung der Module, die Koordination des Einsatzes in Schulen und die Evaluation lagen beim Institut für

die Pädagogik der Naturwissenschaften. Die Evaluation hat ergeben, dass dieses top-down verbreitete Programm den Unterricht offenbar verändert (Prenzel 2001). Ob es die Fachleistungen verbessert hat, ist noch unklar.

Implementation durch Tests und zentrale Prüfungen

In der Präambel zu den KMK-Standards heißt es: Sie sollen verwirklicht werden »in einem Unterricht, der selbstständiges Lernen, die Entwicklung von kommunikativen Fähigkeiten und Kooperationsbereitschaft sowie eine zeitgemäße Informationsbeschaffung, Dokumentation und Präsentation von Lernergebnissen zum Ziel hat« (Teil Mathematik, S. 7).

Dies entspricht dem Bild modernen Unterrichts, muss aber als frommer Wunsch angesehen werden, solange nicht die Voraussetzungen für die Implementation in den einzelnen Schulen geschaffen werden, d.h. die Voraussetzungen für Unterrichtsentwicklung. In den KMK-Standards wird darauf nicht einmal hingewiesen.

Dabei ist davon auszugehen, dass die Lehrer sich nicht umstandslos mit den Standards bekannt machen, geschweige denn anfreunden werden. Empirische Untersuchungen zur Frage »Was wird durch eine Lehrplanrevision verändert« zeigen, dass revidierte Lehrpläne den Unterricht kaum erreichen und ca. 70 Prozent der befragten Lehrpersonen sie gar nicht kannten (Vollstädt/Tillmann u. a. 1999).

Dies Problem scheint den KMK-Planern bewusst gewesen zu sein. Vermutlich um sicher zu gehen, dass die Bildungsstandards von den Schulen auch ernst genommen werden, hat die KMK am 4. Dezember 2003 eine Art Dreifachsicherung beschlossen:

Die Kultusministerkonferenz vereinbart:

- »1. Die Bildungsstandards für den Mittleren Schulabschluss in den Fächern Deutsch, Mathematik, Erste Fremdsprache werden von den Ländern zu Beginn des Schuljahres 2004/2005 als Grundlagen der fachspezifischen Anforderungen für den Mittleren Schulabschluss übernommen.
2. Die Länder verpflichten sich, die Standards zu implementieren und anzuwenden. Dies betrifft insbesondere die Lehrplanarbeit, die Schulentwicklung und die Lehreraus- und -fortbildung. Die Länder kommen überein, weitere Aufgabenbeispiele zu entwickeln und in landesweiten bzw. länderübergreifenden Orientierungs- und Vergleichsarbeiten oder in zentralen oder dezentralen Prüfungen festzustellen, in welchem Umfang die Standards erreicht werden. ...
3. Die Standards und ihre Einhaltung werden unter Berücksichtigung der Entwicklung in den Fachwissenschaften, in der Fachdidaktik und in der Schulpraxis durch eine von den Ländern gemeinsam beauftragte wissenschaftliche Einrichtung überprüft und auf der Basis validierter Tests weiterentwickelt.«

Es wäre allerdings eine Illusion zu glauben, dass mit diesem dreifachen Druck die Implementationsprobleme gelöst seien. Huberman und Miles haben ja – wie erwähnt – schon frühzeitig darauf hingewiesen, dass zum Druck auch Zug (bzw. Support) hinzukommen muss, weil sonst nur Gegendruck entsteht bzw. defensive Routinen aktiviert werden.

Deshalb steht zu vermuten, dass die Schulen die Abschlussprüfungen am Ende der Pflichtschulzeit wohl an den Standards ausrichten, aber den Unterricht nicht signifikant verändern, ebenso wenig die Lernformen der Schülerinnen und Schüler. Zudem sind die Standards bisher nur für die 9. Klassen formuliert und werden demnächst für die 4. Klassen definiert: Niemand weiß, was sie für die 8., 7., 6. oder die 2. Klassen bedeuten!

3. Auf dem Wege zum integrierten Qualitätsmanagement?

Das von der Schule ausgehende Grundmodell A ist mit dem von zentralen Standards ausgehenden Modell B nicht unvereinbar. Im Gegenteil: International wird seit längerem daran gearbeitet, beide zu integrieren (vgl. z. B. Gray u. a. 1996). Auch wissen wir aus der Organisationsentwicklung, dass Bottom-up-Ansätze und Top-down-Ansätze notwendig zusammengehören, wenn (Qualitäts-)Entwicklungen gelingen sollen. Sind wir also auf dem Wege zu einem integrierten Qualitätsmanagement, das zudem Qualitätssicherung mit Qualitätsentwicklung vereint?

Situation: Rezentralisierung und Dominanz der Administration

In Deutschland steht in den letzten Jahren das Teststandard-Modell im Mittelpunkt der öffentlichen Debatte (allerdings nicht in den Schulen), des administrativen Handels und auch der Vergabe von Forschungsmitteln. Schulentwicklung spielt zwar eine Rolle, wird auch erwähnt, bleibt aber eher Rhetorik. Ein typisches Beispiel ist das bereits erwähnte Klieme-Gutachten zu nationalen Bildungsstandards (vgl. Klieme u. a. 2003, S. 110 ff.). Es enthält ein ganzes Kapitel über Schulentwicklung. Faktisch geht es dabei jedoch um kaum mehr als um Gewinnung der Akzeptanz der Lehrerschaft, wie es in diesem Kapitel auch wörtlich heißt: Bezeichnend ist auch, dass in der Gutachtengruppe kein Schulentwickler vertreten war; es war auch kein Schulpädagoge dabei.

Ernst genommene Schulentwicklung verlangt nach Dezentralisierung, Stärkung der Einzelschule, Aufbau von Agenturen zur externen Evaluation, Erarbeitung und Verabschiedung eines Orientierungs- bzw. Referenzrahmens für Qualitätsevaluation und ein intensiv ausgebautes Unterstützungssystem – wie das in etlichen Schweizer Kantonen oder in den Niederlanden der Fall ist.

Guskey betont die Wirksamkeit dezentraler Schulentwicklung und stellt für die USA fest: »Die Lernerfolgsergebnisse, die am besten geeignet sind, Verbesserungen des Lernens der Schüler zu Stande zu bringen, sind Tests, Quizze, Schreibaufgaben und andere Instrumente, die Lehrer auf reguläre Weise in ihrer Klasse anwenden« (Guskey 2003, S. 7).

Demgegenüber dominiert mit dem deutschen Weg, nationale Bildungsstandards zu vereinbaren und mit Druckmitteln durchzusetzen, die Administration und tritt eine Rezentralisierung ein, bevor die Dezentralisierung überhaupt wirklich begonnen hat.

Statt dass »schwache« Schulen besonders unterstützt wurden, bleiben sie im Schatten der Standards. Nationale Regelstandards lenken die Aufmerksamkeit auf die »höheren« Schulen und helfen den Hauptschulen am wenigsten. Und »High Stakes Tests«

verschärfen die Ungleichheit, die in Deutschland im Ländervergleich ohnehin mit am stärksten ausgeprägt ist. Hinzu kommt, dass die Hauptschulen eine Schulform darstellen, die insgesamt den »schwachen« Schulen zuzurechnen ist. PISA 2000 hat ergeben, dass Hauptschulen aus strukturellen Gründen suboptimale Entwicklungsmilieus darstellen, die Schüler gleicher kognitiver Grundausstattung bzw. Intelligenz erheblich weniger fördern als Realschulen oder Gymnasien (Baumert u. a. 2001, S. 182).

Die Frage ist, wie es zur Dominanz nationaler Teststandards kommen konnte. Empirische Evidenz spricht eher dagegen, und erfolgreiche PISA-Länder wie Finnland, Schweden oder Neuseeland arbeiten nicht mit solchen (kompetenzstufenorientierten Test-)Standards.

Der Erfolg dieses Modells in der Öffentlichkeit muss offenbar anders als durch wissenschaftliche Erkenntnisse begründet sein: Nationale Teststandards sind Symbolpolitik. Sie zeigen der allgemeinen Öffentlichkeit, dass Politik und Administration etwas tun, das Qualitätsthema ernst nehmen und beabsichtigen, die Qualität zu erhöhen – was allen nur gefallen kann. Sie versprechen vor allem, die Leistungen der deutschen Schüler zu verbessern, ohne an der leidigen Schulstrukturfrage und der damit verbundenen frühen Selektivität rühren zu müssen. Sie bringen auch nicht das Dauerthema »Reform der Schulaufsicht« auf die Tagesordnung, das mit der Ausnahme Österreichs in allen OECD-Ländern angepackt wurde. Im Gegenteil: Die Schulaufsicht in der bisherigen Form wird dringend benötigt, um die Bildungsstandards zu administrieren. Schließlich sind Bildungsstandards auch kostensparend; sie sind jedenfalls erheblich preiswerter als der sonst fällige Ausbau des Unterstützungssystems für Schulentwicklung. Die Niederlande oder die Schweiz geben heute ein Mehrfaches für solche Unterstützungssysteme aus als Deutschland.

Perspektive: Vom Cocktail zur Synthese

Bildungsstandards sind sinnvoll und nötig als Ausdruck fachlicher Bildungsziele und als Maßstab für Evaluation. Sie werden jedoch problematisch, wenn sie – wie gezeigt – mit »High Stakes Tests« verschmolzen und/oder zum dominanten, zentralen Steuerungsmittel mutieren und/oder nur mit Druck durchgesetzt werden können. Erst in einem Gesamtkonzept von Qualitätsmanagement werden Standards sinnvoll.

Einige Länder, auch Bundesländer, sind auf dem Wege zu einem Gesamtkonzept. Im Ausland sind z. B. Kanada, Schottland oder die Niederlande weit vorangeschritten, in Deutschland z. B. Niedersachsen, Baden-Württemberg, Rheinland-Pfalz und Nordrhein-Westfalen. Sie erproben zahlreiche der erwähnten Komponenten der Steuerung des Qualitätsmanagements, kombinieren auch etliche, haben sie jedoch nicht (oder noch nicht) zu einem umfassenden und konklusiven System integriert. Es handelt sich eher um Cocktails als um Synthesen.

Beispielsweise ist die externe Evaluation – wie in den Niederlanden – stark ausgebaut, aber die interne unterentwickelt. Oder es wird ein adäquat aufeinander eingespieltes Konzept von interner und externer Evaluation geplant und beschlossen –,

aber es mangelt an einer Abstimmung mit den Bildungsstandards und erst recht mit Schulleistungstests.

Wenn aus einem Cocktail eine Synthese werden soll, muss nach Medien der systematischen Verknüpfung gesucht werden. Vier solcher Medien sind zurzeit zu erkennen:

- (1) Bildungsstandards und Schulcurricula,
- (2) allgemeindidaktische und fachdidaktische Unterrichtsentwicklung,
- (3) interne und externe Evaluation,
- (4) Tests zum Zwecke des System-Monitorings und der Schulevaluation.

Zu 1: Bildungsstandards müssen nicht zentral administriert und allein durch Druck implementiert werden. Sie sind »luftig« formuliert, sie geben einen Rahmen vor; sie lassen Raum für schulspezifische Auslegungen. Umgekehrt finden Schulen mit eigenem Profil in den Standards ein Orientierungssystem und auch Maßstäbe für die eigenen Leistungsansprüche. Wenn beides systematisch aufeinander bezogen wird, entsteht ein Schulcurriculum, das sich allerdings auf ein viel breiteres Feld beziehen sollte als die Standards bisher abstecken, z. B. auf andere Fächer und auf Lehr- und Lernformen, die von den bisher vorliegenden Standards gar nicht thematisiert werden.

Adressaten der Standards sind ausdrücklich die Fachkonferenzen. Damit verbunden ist ein Verständnis von Unterrichtsentwicklung als Entwicklung von Fachunterricht. Unterrichtsentwicklung im Rahmen von Schulentwicklung bezieht sich indes eher auf Methoden- und Kommunikationstraining sowie auf kooperatives Arbeiten, also auf Lernen von Schülern (und z. T. auch von Lehrern, vgl. Bastian/Rolff 2001). Fachunterricht bewegt sich gleichsam auf einer vertikalen Achse, im Gymnasium z. B. von der 5. bis zur 13. Klasse. Die Lehrer, die ein Fach unterrichten, arbeiten meist in mehreren Klassen. Der Ort professionellen Austausches ist die Fachkonferenz. In Ermangelung eines Bezugs auf dieselben Schüler (aus derselben Klasse) stehen Fachfragen im Mittelpunkt des Interesses.

Zu 2: Unterrichtsentwicklung im Sinne von Weiterentwicklung der Lernformen und Lernmethoden bezieht sich eher auf eine horizontale Ebene, auf das Lernen von Schülern derselben Klasse oder desselben Jahrgangs. Lehrer, die in diesem Sinne den Unterricht verbessern wollen, betreiben Unterrichtsentwicklung konsequenterweise auch im Rahmen von Klassen bzw. Jahrgangsteams. Beide Formen der Unterrichtsentwicklung gehören innerlich zusammen. Dieser Zusammenhang stellt sich allerdings in fragmentierten Schulen nicht von selber her. Er muss organisiert werden. Abbildung I.7 zeigt das (Koordinaten-)Kreuz integrierter Unterrichtsentwicklung, das durch eine Steuergruppe koordiniert wird und sich letztlich schulweit ausbreiten muss. Wenn diese Koordination gelingt, entsteht integriertes Lernen, eigenverantwortliches Lernen genannt.

Der innere Zusammenhang beider Formen der Unterrichtsentwicklung wird auch in den Texten der Fachdidaktiker deutlich, in denen sich zumeist Formulierungen finden wie selbstreguliertes Lernen, kooperatives Lernen oder zielorientiertes Lernen, also allgemeindidaktische Kategorien. Auf der anderen Seite ist auch den »Lernentwicklern« klar, dass man Methoden nicht ohne Inhalte lernen kann.

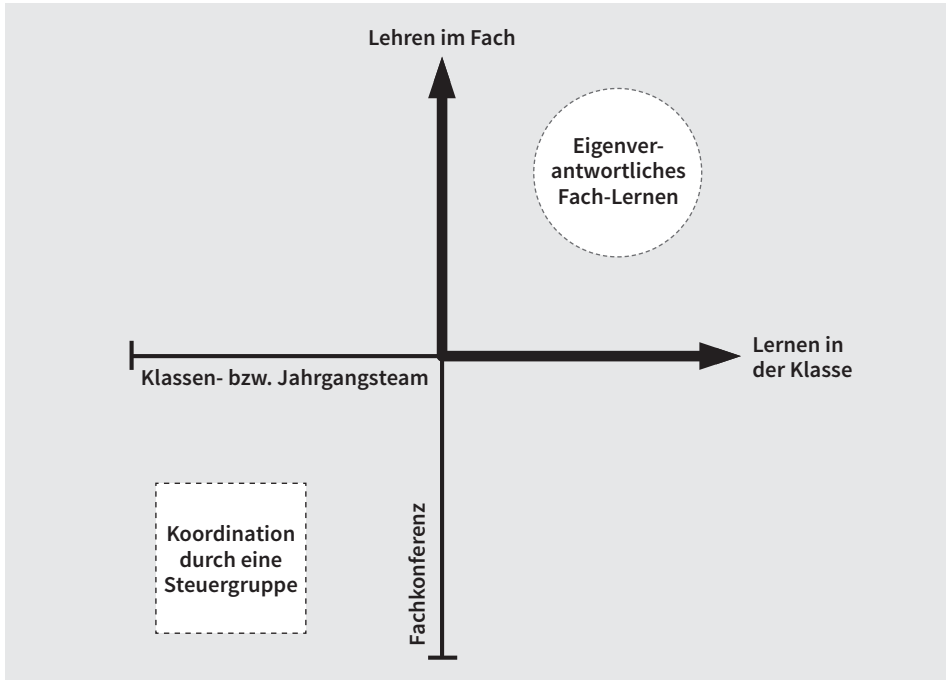


Abb. 1.7: Kreuz der Unterrichtsentwicklung (Quelle: Rolff)

Zu 3: Dass interne und externe Evaluation eng aufeinander bezogen sein müssen, um Wirksamkeit für Qualitätsentwicklung erzeugen zu können, wurde bereits ausführlich dargelegt. Die anspruchsvollste Form eines solchen Bezugssystems stellen die erwähnten Qualitätstableaus dar. Es sind auch weniger aufwendige Referenzsysteme in Form von Fragebogensystemen vorhanden, die gut erprobt und von Schulen gut handhabbar sind. Sie dienen der internen Evaluation von Einzelschulen, dem Vergleich von Schulen untereinander und zudem als Datenbasis für System-Monitoring. Sie eignen sich zur internen wie zur externen Evaluation zugleich.

Der erste in Deutschland (1987) gebräuchliche Fragebogen dieser Art war der »Guide to Institutional Learning« (GIL), den Dalin entworfen hatte (Dalin/Rolff 1990). Der GIL impliziert ein breites Qualitätsverständnis, das sich über zehn Dimensionen erstreckt:

- (1) Ziele und Werte
- (2) Unterrichtspraxis

- (3) Arbeitsklima
- (4) Normen und Erwartungen
- (5) Leitung
- (6) Entscheidungsprozesse
- (7) Einfluss und Kontrolle
- (8) Veränderungen
- (9) Zeit für Arbeitsaufgaben
- (10) Belohnung.

Als unzureichend erwies sich der Umstand, dass der GIL über kein explizites Qualitätsverständnis verfügt und sich lediglich auf die Lehrpersonen bezieht und Schüler und Eltern ausblendet.

Das zuletzt genannte Manko überwand das IFS-Barometer (1996), welches auch Schüler- und Elternfragebögen umfasst und zudem bundesweite Vergleichswerte liefert. Ähnlich aufgebaut ist die sogenannte Pädagogische Entwicklungs-Bilanz (PEB) des Deutschen Instituts für Internationale Pädagogische Forschung (Döbrich 2003), welches zunächst auch nur aus einem Lehrerfragebogen bestand, der 2002 um ein Schüler- und Elterninstrument ergänzt wurde und über Daten aus hessischen wie niedersächsischen Schulen verfügt.

Ein weiteres Instrument dieser Art hat die Bertelsmann-Stiftung erarbeitet (Stern u. a. 2003). Es ist ebenfalls mit Lehrer-, Schüler- und Elternfragebögen ausgestattet. Darüber hinaus orientiert es sich an einem expliziten Qualitätskonzept und ist auch für internationale Vergleiche geeignet.

Ein Nachteil aller vier Instrumente ist allerdings, dass es sich allein um standardisierte Fragebögen handelt, die Kapazität zur Datenerhebung sehr begrenzt ist, Fachleistungen z. B. überhaupt nicht erfasst und Schuleigenarten auch nicht.

Bei der externen Evaluation können im Prinzip zwei Varianten unterschieden werden, direkte und Meta-Evaluation. Die bisherigen Erfahrungen in den Niederlanden, in England und in einigen Schweizer Kantonen lassen vermuten, dass Meta-Evaluation eher geeignet ist, die interne Evaluation zu stimulieren, was allerdings noch einer genaueren Untersuchung bedarf. Würde sich diese Vermutung bestätigen, wäre davon abzuleiten, dass indirektere Kopplungen offenbar zu intensiveren Synthesen führen. – Meta-Evaluationen können von einer Evaluationsagentur, aber auch von Peers durchgeführt werden. Beide Ansätze sind in Deutschland noch unterentwickelt. Die Ministerien haben die Einführung von Tests als Form externer Evaluation vorgezogen.

Zu 4: Das vierte und hier zuletzt behandelte Medium der Integration ist das neuartigste, aufwendigste und möglicherweise wirksamste: die Weiterentwicklung und Kalibrierung von Vergleichsarbeiten zu Fachleistungstests mit mehrfacher Funktion. Es wurde ja bereits erwähnt, dass die Rückmeldung von Daten aus Large-Scale-Assessments, die auf Stichproben beruhen, bei den teilnehmenden Schulen nicht viel bewirkt haben.

Anders stellt sich die Situation dar, wenn eine Vollerhebung erfolgt, also Schülerdaten aus allen Klassen des Bezugsjahrgangs aller Schulen des Landes ermittelt und allen Schulen zurückgegeben werden. Dann ist eine intensivere Auseinandersetzung der Schulen mit diesen Daten zu erwarten und eine wirkungsvollere innerschulische Verarbeitung. Derartige Vollerhebungen fanden zum ersten Mal in Hamburg in Form der LAU-Studie und dann in Rheinland-Pfalz (MARKUS) statt. Über die innerschulische Verarbeitung ist auch hier bisher wenig bekannt. Schrader und Helmke haben eine erste Evaluationsstudie aus Rheinland-Pfalz vorgelegt, dabei betrug die Rücklaufquote der Schulen allerdings weniger als 5 Prozent (Schrader/Helmke 2003).

Intensiver wird vermutlich die Kopplung mehrerer Funktionen bei VERA erfolgen, einem Vorhaben, an dem sieben Bundesländer teilnehmen. VERA ist ein Kürzel für Vergleichsarbeiten (Orth 2002), die allerdings zu Tests weiterentwickelt wurden, ohne ihre Schulnähe aufzugeben. VERA wird von allen 4. Klassen aller Schulen eines Bundeslandes bearbeitet. Es existiert eine normierte Zentralstichprobe, die repräsentativ für das Bundesland ist. Jeder Schule wird die Hälfte der Aufgaben für die Vergleichsarbeiten aus dieser Zentralstichprobe vorgegeben, sodass Daten für ein zentrales System-Monitoring zur Verfügung stehen. Die ausgewerteten Daten werden jeder Schule per Internet zugespielt, sodass sie sich am Landesdurchschnitt und – wegen der Fairness – mit Schulen in vergleichbarer soziokultureller und sozioökonomischer Lage messen kann. Auf der anderen Seite (nämlich der der Schule) kann jede Schule eigene Aufgaben zur Evaluation des eigenen spezifischen Profils aussuchen, allerdings nicht frei, sondern aus einem auf Zuwachs angelegten Datenpool. Das dürfte Anlass für eine stärkere Identifikation (»ownership«) der Schule mit diesen testartigen Vergleichsarbeiten sein (vgl. Helmke/Hosenfeld). Ein ähnliches Verfahren für die 9. Jahrgänge wurde später entwickelt.

Diese Art von Vergleichsarbeiten stellt die wohl engste Verbindung zentral administrierter und schulbasierter Steuerung der Schulqualität dar. Top-down- und Bottom-up-Ansätze werden amalgamiert. Die Daten werden schulweit veröffentlicht und gehen auch an die Eltern. Datenbasierte und ergebnisorientierte Schulentwicklung wird möglich.

Druck und Zug können zusammenkommen – Zug allerdings nur, wenn die beteiligten Länder entsprechende Unterstützungssysteme für die Einzelschulen aufbauen. Qualitätsmanagement kann nur durch Schulentwicklung realisiert werden. Die Schulen müssen lernen, ihr Qualitätsmanagement in die eigenen Hände zu nehmen. Deshalb verlangt Schulentwicklung nach mehr Lehrerfortbildung und Personalentwicklung.

II. PQM – ein Modell für die Praxis

In Kapitel I haben wir gezeigt: Qualitätsmanagement von Schulen kann auf zweierlei Weise betrieben werden: von oben und von unten. Höchstwahrscheinlich ist eine Kombination von beidem das Effektivste. Über Qualität von oben herrscht eine intensive bildungspolitische Diskussion; bundes- und landesweite Standards und nationale wie internationale Tests sind die Stichwörter. Über Qualität von unten gibt es keinen öffentlichen Diskurs, und es ist auch wenig klar, was damit gemeint ist. Deshalb wollen wir in diesem Buch aufzeigen, wie Qualitätsentwicklung von unten, d. h. in der Einzelschule aussieht und wie sie gelingen kann: als Schulentwicklung zum einen und als Evaluation zum anderen. Zertifizierung könnte eine zusätzliche Komponente sein, sozusagen das i-Tüpfelchen.

