

Klieme, Eckhard [Hrsg.]; Tippelt, Rudolf [Hrsg.]

## **Qualitätssicherung im Bildungswesen. Eine aktuelle Zwischenbilanz**

Weinheim u.a. : Beltz 2008, 226 S. - (Zeitschrift für Pädagogik, Beiheft; 53)

urn:nbn:de:0111-opus-72653



in Kooperation mit / in cooperation with:

# **BELTZ JUVENTA**

<http://www.juventa.de>

### **Nutzungsbedingungen / conditions of use**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

We grant a non-exclusive, non-transferable, individual and limited right to using this document.  
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.  
By using this particular document, you accept the above-stated conditions of use.

### **Kontakt / Contact:**

**peDOCS**  
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)  
Mitglied der Leibniz-Gemeinschaft  
Informationszentrum (IZ) Bildung  
Schloßstr. 29, D-60486 Frankfurt am Main  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Zeitschrift für Pädagogik · 53. Beiheft



Zeitschrift für Pädagogik · 53. Beiheft

# **Qualitätssicherung im Bildungswesen**

Eine aktuelle Zwischenbilanz

Herausgegeben von  
Eckhard Klieme und Rudolf Tippelt

Beltz Verlag · Weinheim und Basel

Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich geschützt. Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten. Kein Teil dieser Zeitschrift darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form – durch Fotokopie, Mikrofilm oder ein anderes Verfahren – reproduziert oder in eine von Maschinen, insbesondere Datenverarbeitungsanlagen, verwendbare Sprache übertragen werden. Auch die Rechte der Wiedergabe durch Vortrag, Funk- und Fernsehsendung, im Magnettonverfahren oder auf ähnlichem Wege bleiben vorbehalten. Fotokopien für den persönlichen oder sonstigen eigenen Gebrauch dürfen nur von einzelnen Beiträgen oder Teilen daraus als Einzelkopie hergestellt werden. Jede im Bereich eines gewerblichen Unternehmens hergestellte oder genutzte Kopie dient gewerblichen Zwecken gem. § 54 (2) UrhG und verpflichtet zur Gebührenzahlung an die VG Wort, Abteilung Wissenschaft, Goethestr. 49, 80336 München, bei der die einzelnen Zahlungsmodalitäten zu erfragen sind.

©2008 Beltz Verlag · Weinheim und Basel  
Herstellung: Klaus Kaltenberg  
Gesamtherstellung: Druckhaus »Thomas Müntzer«, Bad Langensalza  
Printed in Germany  
ISSN 0514-2717  
Bestell-Nr. 41154

# Inhaltsverzeichnis

<i>Eckhard Klieme/Rudolf Tippelt</i> Qualitätssicherung im Bildungswesen. Eine aktuelle Zwischenbilanz .....	7
<b>Bereiche</b>	
<i>Wolfgang Tietze</i> Qualitätssicherung im Elementarbereich .....	16
<i>Hartmut Ditton</i> Qualitätssicherung in Schulen .....	36
<i>Olaf Köller</i> Bildungsstandards in einem Gesamtsystem der Qualitätssicherung im allgemeinbildenden Schulsystem Deutschlands .....	59
<i>Thomas Olk/Karsten Speck</i> Qualität und Qualitätsentwicklung in der Sozialpädagogik .....	76
<i>Philipp Gonon</i> Qualitätssicherung in der beruflichen Bildung. Eine Bilanz .....	96
<i>Klaus Meisel</i> Qualitätsmanagement und Qualitätsentwicklung in der Weiterbildung .....	108
<i>Lutz von Rosenstiel</i> Qualitätssicherung in der betrieblichen Weiterbildung .....	122
<i>Andrä Wolter/Christian Kerst</i> Akkreditierung als Verfahren der Qualitätssicherung von Studiengängen in Deutschland. Eine Policy-orientierte Analyse .....	135
<i>Bernhard Schmidt</i> Qualität der Lehre an Hochschulen .....	156
<i>Rudolf Tippelt</i> Qualitätsstandards und Forschungsevaluationen in der Erziehungswissenschaft .....	171

**Implementierung von Qualitätssicherung**

*Helmut Fend*

Dimensionen von Qualität im Bildungswesen. Von Produktindikatoren zu  
Prozessindikatoren am Beispiel der Schule ..... 190

*Esther Berner/Jürgen Oelkers/Kurt Reusser*

Implementation von Bildungsstandards: Bedingungen des Gelingens  
(und Scheiterns) aus internationaler Sicht ..... 210

Eckhard Klieme/Rudolf Tippelt

## Qualitätssicherung im Bildungswesen

*Eine aktuelle Zwischenbilanz*

„Qualität“ ist kein genuin pädagogischer Begriff, „Qualitätssicherung“ schon gar nicht. Trotzdem – oder vielleicht gerade deswegen – machen diese Begriffe seit nahezu 20 Jahren Karriere in der Bildungspolitik, im Alltag der Bildungseinrichtungen aller Sektoren, und auch im erziehungswissenschaftlichen Diskurs. Die Zeitschrift für Pädagogik legt hiermit – nach dem Beiheft „Qualität im Bildungswesen“ aus dem Jahr 2000 – eine zweite Zwischenbilanz der deutschsprachigen Qualitätsdebatte und ihrer internationalen Referenzen vor. Ging es in der ersten Bilanz darum, eine neue Terminologie begrifflich, auch metatheoretisch zu bearbeiten, sollen nun die verschiedenen Spielarten der „Implementation“ von Qualitätssicherungsverfahren im Durchgang durch die Bildungsstufen und -sektoren dokumentiert, ihre konzeptuellen Grundlagen offen gelegt, die damit verbundenen theoretischen und praktischen Zumutungen und Paradoxien kritisch diskutiert werden.

### 1. Aktualität und Reichweite des Qualitätsdiskurses

Alle hier versammelten Beiträge erkennen die Omnipräsenz des Qualitätsbegriffs und der damit verbundenen neuen Steuerungssysteme als Faktum an. Es ist unstrittig, dass Zielvorgaben, Strukturen und Prozesse der institutionalisierten Bildung in den deutschsprachigen Ländern in den vergangenen zehn bis zwanzig Jahren massiv verändert worden sind, ausgerichtet an Effektivität, Effizienz, Evaluation und (zumeist) Empirie. Unstrittig ist auch, dass damit Entwicklungen nachvollzogen wurden, die andernorts – vor allem in den angelsächsischen Staaten und den Niederlanden – etwa zehn Jahre früher eingesetzt hatten und inzwischen teils nachhaltig implementiert, teils schon wieder gebrochen sind, konterkariert durch unerwünschte Nebenwirkungen und mehr oder weniger subversive Strategien (z.B. „teaching to the test“).

Das deutsche Bildungswesen (und ähnlich die Systeme Österreichs und der deutschsprachigen Schweiz) hat sich durch die Einführung von Qualitätssicherungsverfahren grundlegend und auf breiter Linie verändert. Traditionell zeichnete es sich durch hohe Selbstreferenz und starke professionelle Autonomie einerseits, schwache Aufsichtsstrukturen sowie geringe Transparenz und niedrigen Rechtfertigungsdruck gegenüber Außenstehenden andererseits aus, wie Fend es in seinem Beitrag zum vorliegenden Band darstellt. Inzwischen lässt sich jedoch mit Ditton (in diesem Band) sagen: „Die Notwendigkeit von Qualitätssicherung (...) ist nicht (mehr) umstritten. Dass Zielklarheit und Erfolgskontrolle in jedem System nötig sind, das erfolgreich operieren soll, steht außer

Frage“. Während sich diese Darstellung eines neu etablierten Denkens auf das Schulsystem bezieht, spricht Meisel (in diesem Band) im Bereich der Weiterbildung bereits von einer „Verfaserung“ der Qualitätssicherungsmaßnahmen, d.h. von einer undurchsichtigen Vielfalt höchst unterschiedlicher Vorgehensweisen, die sich mit diesem Begriff verbinden.

Bei allen vermeintlichen Neuerungen darf nicht übersehen werden, dass zentrale Begriffe – auch in den Beiträgen des vorliegenden Bandes – immer wieder Anleihen bei der Curriculumentwicklung und der Evaluationsforschung der 1960er und 1970er Jahre machen. Beispielsweise ist auch in den vorliegenden Papieren von Lernzielen und Lernzielüberprüfung die Rede. Schon Mitte der 1960er Jahre plädierte Ingenkamp in einem Gutachten für den Deutschen Bildungsrat für die flächendeckende Nutzung standardisierter Tests, die an einem nationalen und mehreren Länder-Instituten entwickelt werden sollten – eine Konzeption, deren Realisierung sich heute nach Gründung des IQB und anderer Qualitätsagenturen abzuzeichnen scheint (vgl. Köller und Ditton, in diesem Band). Dennoch gibt es zwischen der heutigen Qualitätssicherung und den vierzig Jahre zuvor proklamierten Konzepten deutliche Unterschiede: Viele damalige Anläufe zu Curriculumentwicklung und Evaluierung waren relativ naiv technokratisch ausgerichtet, unterschätzten die Probleme der Implementation von Reformen, die damit verbundenen Brüche und Rekontextualisierungen (vgl. Fend, in diesem Band). Die Vermittlung systemischer Reformen bis hinein in unterrichtliche Lehr-Lern-Prozesse wurde ebenso wenig verstanden wie die Wechselwirkung von Angebot und Nachfrage – die Ko-Konstruktion des Lerngeschehens – innerhalb einer Lerngruppe. Die notwendige Fach- und Bereichsspezifität von Maßnahmen der Qualitätsentwicklung und -sicherung war weniger ausgearbeitet als heute, und die Methodik etwa der Leistungsmessung war bei weitem nicht so weit entwickelt. Umso wichtiger ist es, im aktuellen Diskurs über Qualitätssicherung immer wieder die Erfahrungen mit der „realistischen Wendung“ pädagogischer Forschung in den 1960er Jahren zu reflektieren.

## 2. Grundbegriffe

Interessanterweise wird der Begriff der Qualitätssicherung von keinem der Autoren dieses Bandes explizit definiert – sei es, weil es offensichtlich zu sein scheint, was darunter zu verstehen sei, sei es, dass schon innerhalb der einzelnen Sektoren die damit verbundenen Verfahren und Aktivitäten so vielschichtig sind, dass man sie keiner kohärenten Definition subsummieren kann. Der Begriff „Qualität“ wird eher umschrieben und durch Teilaspekte – gewissermaßen taxonomisch – ausbuchstabiert (so bei Gonon und Fend; für Definitionsansätze vgl. aber Tietze und Schmidt; alle in diesem Band).

Qualitäten werden als Eigenschaften der Kontexte sowie der Ressourcen und Ausgangsbedingungen (des Input), der Prozessabläufe sowie der Ergebnisse (unmittelbarer Output und langfristiger Outcome) von Bildungsprozessen beschrieben. Während die Qualitätsdebatte vor kurzem zumeist noch mit einer „Outputorientierung“ verbunden wurde, die als Alternative zu einer traditionellen „Inputorientierung“ galt, dokumen-

tiert die hier vorgelegte Zwischenbilanz einen wichtigen Wandel: Nahezu jeder Beitrag betont den zentralen Stellenwert der *Prozesse*<sup>1</sup> – mit der Differenzierung, dass hierunter mitunter inhaltsbezogene, d.h. bildungs- und erziehungsrelevante pädagogische Interaktionen gemeint sind (also eine *pädagogische* Prozessqualität), häufig jedoch formalisierte Verfahren und Regelungen (also eine *institutionenbezogene* Prozessqualität). Auf jeden Fall scheint sich in der erziehungswissenschaftlichen Diskussion die Erkenntnis durchgesetzt zu haben, dass Bildungsqualität zuvorderst an der Qualität von Prozessen festzumachen ist, auch wenn sich empirische Studien und Evaluationen an erzielten (Lern-)Ergebnissen als ultimativen Maßstäben orientieren.

In allen Beiträgen wird klar, dass es nicht *die* eine Skala für Bildungsqualität geben kann, sondern bereichs- und zielabhängig unterschiedliche Kriterien und Indikatoren benannt werden müssen. Im Anschluss an Heids Beitrag im Beiheft 2000 wird übereinstimmend festgehalten, dass Qualitätsfeststellungen einer normativen Grundlage bedürfen, d.h. einer zumindest impliziten *Setzung* von Gütekriterien, die nicht wahr oder falsch sein kann, sondern nur unterschiedlich gut legitimierbar oder nützlich. Zugleich gehen die Beiträge davon aus, dass gewählte Gütekriterien objektivierbar sein können und müssen, um evaluative Aussagen zum erreichten Qualitätsniveau machen zu können. Dies gilt explizit auch für die Qualitätssicherung und Evaluierung von Forschungsinstitutionen, die im Qualitätsdiskurs zwar eine Sonderstellung einnehmen, aber rechtlich und mit Blick auf die öffentliche Legitimierung von Forschungsausgaben in die regelmäßige Überprüfung nach Gütekriterien einbezogen sind (vgl. Tippelt in diesem Band).

Quer zu den Kategorien des klassischen „Bildungsproduktionsmodells“ (Input – Prozess – Output) werden die Ebenen der Systemarchitektur unterschieden, bezüglich derer Qualitätsfragen gestellt werden: bezüglich der domänenspezifischen Lehr-Lern-Prozesse innerhalb eines Kurses (Mikroebene), bezüglich der Organisationseinheit (Mesoebene, z.B. Schule) oder bezüglich des Systems insgesamt (Makroebene). Im Vordergrund steht eindeutig die Mesoebene, denn hier setzen Qualitätssicherungs- und Qualitätsentwicklungskonzepte zumeist an.

Qualitätssicherung kann – um eine weitere Gemeinsamkeit der Beiträge auszuformulieren – verstanden werden als Gesamtheit von formalisierten und institutionalisierten Verfahren, mit denen die Qualität von Organisationen im Bildungssystem, insbesondere deren Prozess- und Ergebnisqualität, beschrieben und vergleichend bewertet wird. Wenn von Qualitätssicherung die Rede ist, ist also fast immer die einzelne Bildungseinrichtung (z.B. Schule, Weiterbildungseinrichtung) als „Handlungseinheit“ gemeint.

Der Begriff Qualitätssicherung ist unscharf und umstritten, setzt er doch Assoziationen von „Kontrolle“ und „Zwang“ frei, die in den analogen englischsprachigen Be-

1 Eine Sonderstellung nimmt der Beitrag von Tietze zur Elementarpädagogik ein: Hier wird pädagogische Prozessqualität als „Output“ interpretiert, während „Effekte auf Kinder und Eltern“ als Outcome gelten. Die Herausgeber haben davon abgesehen, die Terminologie zu vereinheitlichen, um bereichsspezifische Debatten authentisch darstellen zu können.

griffen *Accountability* und *Quality Management* nicht in gleicher Weise mitschwingen. Es ist daher nicht verwunderlich, dass der Begriff zwar in wissenschaftlichen Publikationen verwendet, aber im politisch-administrativen und auch im wissenschaftlichen Alltag eher vermieden wird. Institute bzw. Abteilungen für Qualitätssicherung werden in Deutschland – jedenfalls im Schulbereich – inzwischen als „Qualitätsagenturen“ oder als Einrichtungen für „Qualitätsentwicklung“ geführt. Das scheint eher anschlussfähig an ein professionelles pädagogisches Selbstverständnis (Erziehung als Unterstützung für Entwicklung), drückt aber auch bildungspolitischen Gestaltungswillen aus. Das Konzept der Qualitätssicherung wird so von der bloßen Feststellung und Bewertung der Qualität (wie oben definiert) hin zu gezielten Veränderungen verschoben.

Damit nähert sich Qualitätssicherung dem Konzept des Qualitätsmanagements, das tatsächlich über Qualitätssicherung hinausgeht; es umfasst neben der Feststellung, und Bewertung von Bildungsqualität auch Verfahren und Maßnahmen zu deren Veränderung und Entwicklung. Hier liegt der Aspekt mehr auf der organisatorischen Verankerung und Implementierung von Qualitätssicherungs- und Qualitätsentwicklungsvorhaben. Damit kommt die Qualitätsdebatte in die Nähe von betriebswirtschaftlichen Managementlehren, wo sie nach Darstellung verschiedener Autoren, v.a. aus dem Bereich beruflicher Aus- und Weiterbildung (z.B. Gonon, in diesem Band), tatsächlich ihren Ursprung hat.

### **3. Unterschiedliche Wurzeln und Bewertungen in den Sektoren des Bildungswesens**

Vergleicht man die Darstellungen aus dem Berufs- und Weiterbildungsbereich und der Sozialpädagogik einerseits (Gonon, Meisel, von Rosenstiel, Olk/Speck), dem Schul- sowie Vorschulbereich andererseits (Ditton, Tietze; auch Fend, Berner/Oelkers/Reusser und Köller), so werden sehr unterschiedliche Wurzeln der Qualitätsdebatte erkennbar:

- Aus Sicht von Autoren, die das Schulsystem fokussieren, sind Verfahren der Qualitätssicherung eine Antwort auf Schwächen des Bildungssystems, wie sie in großen, zumal international vergleichenden Schulleistungsstudien festgestellt wurden (TIMSS; PISA, PIRLS/IGLU). Diese Studien können als Qualitätssicherungsinitiativen auf der Makro-Ebene betrachtet werden, die inzwischen, mit Hilfe von so genannten Bildungsstandards, auf die Meso- und teilweise die Mikroebene heruntergebrochen wurden (vgl. Köller und Ditton, in diesem Band). Bei der Entwicklung von Strategien der Qualitätsbewertung und -veränderung wird auf einen breiten Fundus von Wissen aus der Schul-, Lehrer- und Unterrichtseffektivitätsforschung zurückgegriffen; analog im Vorschulbereich (vgl. Tietze). Qualitätssicherung erscheint in dieser Perspektive als Konsequenz empirischer Studien und Erkenntnisse, also zumindest teilweise wissenschaftsbasiert.
- Aus Sicht von Autoren, die in außerschulischen Bildungsbereichen arbeiten, dominiert hingegen das betriebswirtschaftliche Konzept, wie es durch die Normierung

von ISO 9000ff. definiert ist. Dieses Konzept wurden zunächst im Produktionssektor im Rahmen des sogenannten „Total Quality Management (TQM)“ entwickelt. In Produktionsbetrieben ging es darum, die Produktionsabläufe verbindlich und transparent zu planen und zu beschreiben sowie systematische Kontrollen einzubauen, so dass eine gleich bleibend hohe Qualität des Produktionsprozesses gesichert werden konnte. Ziel war letztendlich eine möglichst hohe Kundenzufriedenheit. Das „Total Quality Management“ arbeitet jedoch nicht mit einer inhaltlichen Vorstellung von der Qualität des jeweiligen Produkts, sondern zielt ausschließlich auf die Optimierung der Prozessabläufe durch eine im Extremfall geradezu bürokratische Planung und Dokumentation. Das Konzept wurde auf den Dienstleistungssektor übertragen und schließlich – u.a. durch die „European Foundation for Quality Management (EFQM)“ – in den Bildungsbereich übernommen, vor allem in wirtschaftsnahen Sektoren wie der Berufs- und Weiterbildung, aber auch im Kontext der Schulevaluation. Kritisiert wird von vielen Autoren, dass Bildungsprozesse dadurch einer ökonomischen Logik unterworfen werden, und dass nur noch Prozeduren festgelegt werden, ohne wirklich nach den Inhalten von Bildung oder nach Lernergebnissen zu fragen. Qualitätssicherung erscheint hier als Übertragung betriebswirtschaftlicher Managementverfahren auf den Bildungsbereich.

- Beide vorgenannten Wurzeln von Qualitätssicherungskonzepten verbinden sich teilweise mit der Einführung moderner Verfahren zur Steuerung öffentlicher Institutionen, dem „New Public Management“, bei denen es darum geht, die operative Verantwortung für staatliche Dienstleistungen zu dezentralisieren, zugleich jedoch eine enge und möglichst präzise Ergebniskontrolle durchzuführen. Im Bildungssektor ist diese Art von Steuerung als „Output/Outcome-Orientierung“ bezeichnet worden. Nicht zuletzt supranationale Organisationen wie die Europäische Union und die OECD haben solche Strategien international verbreitet. Qualitätssicherung ist aus dieser Perspektive nichts anderes als die Implementation einer bestimmten staats- und verwaltungswissenschaftlichen Denkrichtung im öffentlichen Bildungswesen.

Diese unterschiedlichen Sichtweisen, hinter denen je spezifische Traditionen und Entwicklungswege in den Sektoren des Bildungswesens stehen, führen dazu, dass die Beiträge des vorliegenden Bandes das Prinzip der Qualitätssicherung letztlich sehr verschieden bewerten: neutral als Leitkonzept aktueller Diskurse, positiv als Chiffre für Reformdebatten oder auch negativ als bloßes Etikett für allerlei politisch oder marktwirtschaftlich inspirierte Eingriffe in pädagogische Arbeit.

#### **4. Kern kritischer Diskussion: Qualitätssicherung versus Professionalität**

Kritiker verstehen Qualitäts(sicherungs)konzepte als dezidierte Gegenkonzepte zu „einheimischen“ Begriffen wie Bildung, Erziehung und Professionalität. Pädagogisches Denken und Handeln wird nach diesem Verständnis durch das Qualitätsmanagement ausgegrenzt und eingeengt. Spitzt man die Tendenz der Argumentation u.a. bei Olk/

Speck und Gonon entsprechend zu, lässt sich die These formulieren: *Qualitätssicherung ist Deprofessionalisierung*. Auf der anderen Seite stehen Versuche (u.a. bei Tietze und Meisel), den Qualitätsbegriff pädagogisch zu besetzen oder (um-)zudeuten und „pädagogische Qualität“ von rein betriebswirtschaftlichen Gütekriterien abzugrenzen, so dass „pädagogische Qualitätssicherung“ geradezu als ein Königsweg der Weiterentwicklung pädagogischer Arbeit erscheint: *Qualitätssicherung ist Professionalisierung*.

Gerade am Verhältnis zwischen Qualitätssicherung und Professionalität brechen auch in der Praxis Konflikte auf. Qualitätssicherung als evaluierende Aufgabe von Organisationen (Schulen, Kindergärten usw.) gerät in Konflikt mit dem Selbstverständnis von Pädagogen, die ihre Aufgabe darin sehen, professionelles Wissen in eigener Verantwortung zur Lösung konkreter pädagogischer Probleme und zur Gestaltung pädagogischer Prozesse einzusetzen. Die Autonomie professionellen pädagogischen Handelns stellt eine zentrale Herausforderung für die Idee der systematisierten Qualitätssicherung dar. Die Eigenlogik und der Fallbezug des pädagogischen Handelns können nie voll aufgehen in standardisierten Kriterien, Bewertungsverfahren und Zielmarken, wie sie beispielsweise in Bildungsstandards festgelegt werden. Dies gilt umso mehr, wenn man anerkennt, dass pädagogische Prozesse von Pädagogen und Teilnehmern ko-konstruktiv bestimmt werden. Die Zuschreibung von Verantwortung *allein* an Organisationen und darin tätige Professionelle muss von diesen mit gewissem Recht als Zumutung empfunden werden.

## 5. Reflexions- und Forschungsbedarf

Solche Konfliktlinien zu rekonstruieren, ist eine zentrale Aufgabe der erziehungswissenschaftlichen Reflektion zum Thema Qualitätssicherung, auch im vorliegenden Band. Zumindest wenn man das Anliegen der Qualitätssicherung nicht grundlegend zurückweist, stellen sich zudem praktische Fragen, die einer wissenschaftlichen Klärung bedürfen. Auch sie ziehen sich durch die Beiträge dieses Heftes: Wie kann man der Vielfalt von Qualitätsperspektiven und -kriterien gerecht werden? Welche Rolle nehmen Leitung, Professionelle und Teilnehmer von Bildungseinrichtungen im Prozess der Qualitätssicherung ein? Inwiefern ist das Konzept „Kunde“ auf Bildungseinrichtungen übertragbar und sollen unmittelbare Teilnehmer oder deren sozialen Kontexte als Kunden gelten? Richten sich die Handlungsmuster und Normen der Professionellen – sogar ohne entsprechende Intention – an den (wahrgenommenen) Qualitätskriterien aus? Wie bestimmt man die „Qualität der Qualitätssicherung“? Was kann man tun, um unerwünschte Nebenwirkungen der Qualitätssicherungsmaßnahmen zu vermeiden? Kann Bildungsqualität durch Akkreditierung von Bildungseinrichtungen bzw. Bildungsgängen – also gewissermaßen im Stadium der Planung, in dem mögliche Lernergebnisse nur vorausgedacht sind – hergestellt werden, oder bedarf es einer Ergänzung dieser Akkreditierung durch Evaluierung und Ergebnismessung (vgl. die Beiträge von Wolter/Kerst und Schmidt)? Wie kann man eine objektive, vergleichende Feststellung von Lernergebnissen (Kompetenzmessung) theoretisch und methodisch absichern?

Grundsätzlich sind Qualitätssicherungsmodelle darauf angewiesen, dass Erkenntnisse zur Effektivität bzw. Wirkung und zur Effizienz von pädagogischen Prozessen im jeweiligen Bildungsbereich vorliegen, auf deren Basis Kriterien und Indikatoren bestimmt und Befunde bewertet werden können (als Beispiel vgl. Tietze und Schmidt, in diesem Band). Solche Erkenntnisse sollten sich nicht nur auf Querschnittsdaten (Vergleichsstudien, Benchmarks, Rankings) stützen, sondern die Chancen und Hürden für Entwicklungsprozesse in pädagogischen Organisationen aufdecken. Zugespitzt formuliert: Qualitätssicherung ist letztlich nur dann wissenschaftlich seriös, wenn abgesicherte Erkenntnisse darüber vorliegen, (a) welche Beobachtungen und Messungen tatsächlich als Indikatoren für mehr oder weniger hohe (pädagogische) Qualität gelten können und (b) welche Konsequenzen (Strukturveränderungen, konkrete Maßnahmen usw.) gezogen werden können, um Qualität weiter zu verbessern. Nicht in allen Bildungssektoren ist aber die Wirkungs- und Interventionsforschung weit genug fortgeschritten, um solche Aussagen abzusichern. Vielmehr werden Erhebungen aus der Qualitätssicherung an Bildungseinrichtungen und aus dem Bildungsmonitoring (also z.B. die PISA-Studien oder die Daten aus Vergleichsarbeiten und Schulinspektion) häufig benutzt, um Faktoren der Bildungsqualität zuallererst zu erforschen. Grundlagenwissen über Qualität bzw. Qualitätsentwicklung einerseits, konkrete Qualitätssicherung in der Praxis andererseits entwickeln sich – im optimalen Fall – durch- und miteinander.

Die Herausgeber sind zuversichtlich, dass mit der vorliegenden Zwischenbilanz Anstöße für weiterführende erziehungswissenschaftliche Forschungen gegeben werden können.

## **Literatur**

Helmke, A./Hornstein, W./Terhart, E. (Hrsg.) (2000): *Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule*. 41. Beiheft der Zeitschrift für Pädagogik. Weinheim: Beltz.



# Bereiche

*Wolfgang Tietze*

## Qualitätssicherung im Elementarbereich

### 1. Einleitung

Die Frage der Entwicklung und Sicherung pädagogischer Qualität im Elementarbereich (darüber hinaus im gesamten Bereich öffentlich verantworteter Bildung, Betreuung und Erziehung vor der Schule) stellt sich für diese Stufe anders und zugespitzter als für die anschließende Schule. Die Schule ist ein lange etabliertes System, sie steht zudem unter der Randbedingung quantitativer Schrumpfung, so dass aufgrund der „demografischen Rendite“ freiwerdende Ressourcen für Qualitätsverbesserungen eingesetzt werden können, Qualitätsverbesserungen also tendenziell ohne zusätzliche Ressourcenmobilisierung erreicht werden können. Für den vorschulischen Bereich ergeben sich andere Randbedingungen. Die Anerkennung des Kindergartens als Elementarbereich und grundlegende Stufe des Bildungssystems ist vergleichsweise jungen Datums (Deutscher Bildungsrat 1970) und im Falle des Krippenbereichs gesellschaftlich noch umstritten. Das System ist weniger straff organisiert als die Schule, die staatlichen Steuerungsmöglichkeiten sind aufgrund der dominierenden freien Trägerschaft eingeschränkter.

Was die demografische Rendite anbelangt, so gilt diese für den Vorschulbereich zwar auch: So schätzt der Zwölfte Kinder- und Jugendbericht (12. KJB 2006, S. 196f.), dass auf der Grundlage der abgestimmten Bevölkerungsschätzung im Jahr 2010 gegenüber 2002 rund 15% weniger Kinder im Elementarbereich zu versorgen sind und rechnerisch damit 650.000 Plätze im Elementarbereich freiwerden. Damit stünden – *ceteris paribus* – im Jahr 2010 annähernd 2 Mrd. Euro zur Verfügung, die bei sonst konstanten Aufwendungen gezielt für Qualitätssicherung eingesetzt werden könnten. Allerdings steht dieser Schrumpfung eine quantitative Expansion gegenüber, die in dieser Weise auf anderen Bildungsstufen nicht gegeben ist. Im Elementarbereich fehlen Ganztagskindergartenplätze in den westlichen Flächenländern, die deutlich teurer sind als die traditionellen Halbtagskindergartenplätze und potenziell freiwerdende Mittel binden, ganz zu schweigen von der erforderlichen Expansion des öffentlichen Früherziehungssystems für die unter Dreijährigen: Bei einem Rechtsanspruch für die Jüngeren und einem angemessenen Ganztagsangebot müssten selbst bei einem Verbleib der demografischen Rendite im System rund 3 Mrd. Euro zusätzlich pro Jahr für den quantitativen Ausbau aufgebracht werden (12. KJB 2006, S. 215). Unter den Bedingungen quantitativer Expansion ist es schwierig, überhaupt Ressourcen für Qualitätsverbesserungen zu mobilisieren. Die Erfahrung verweist in die gegensätzliche Richtung. So wurden im Zuge des Rechtsanspruchs auf einen Kindergartenplatz in den 1990er Jahren bis dahin geltende qualitätsrelevante Standards (z.B. Gruppengrößen, Baustandards) dereguliert, um auf

diese Weise neue Plätze aus den finanziellen Ressourcen des bestehenden Systems zu finanzieren (vgl. Reidenbach 1996). Solche Rahmenbedingungen gilt es im Blick zu behalten, wenn im Folgenden Ansätze der Qualitätsentwicklung und Qualitätssicherung im Elementarbereich bzw. im Vorschulbereich insgesamt diskutiert werden.

## 2. Zum Begriff pädagogischer Qualität im Vorschulbereich

Der Qualitätsbegriff im Vorschulbereich ist schillernd und je nach Perspektive und Bezugsgruppe kann Unterschiedliches darunter verstanden werden (Katz 1996; Dahlberg/Moss/Pence 1999). Die Qualität einer Einrichtung stellt sich aus Sicht von Eltern, die aufgrund ihrer Erwerbstätigkeit auf eine lange Öffnungszeit für ihr Kind angewiesen sind, anders dar als für Erzieherinnen, denen erweiterte Öffnungszeiten Schichtdienst und eigene familienunfreundliche Arbeitszeiten abverlangen. Ein Träger, dem es gelingt, die pädagogische Arbeit eng an seiner weltanschaulichen Ausrichtung zu orientieren, wird eine entsprechende Einrichtung qualitativ anders bewerten als Eltern mit einer anderen Weltanschauung, die mangels einer Alternative auf Nutzung dieser Einrichtung angewiesen sind, oder auch als die dort tätigen Erzieherinnen, die sich eine weniger enge Anbindung an die Leitbilder ihres Trägers und einen größeren Gestaltungsspielraum für die eigene pädagogische Arbeit wünschen mögen (Dippelhofer-Stiem/Kahle 1995). Diese Sichtweisen, die durch die Perspektiven weiterer „stakeholder“ zwanglos erweitert werden können, haben ihre eigene Legitimität. Sie können jedoch kaum als Begründungen für „pädagogische Qualität“ herangezogen werden. *Pädagogische Qualität* sollte das aktuelle wie das (stellvertretend wahrgenommene) zukünftige Interesse des Kindes auf Wohlbefinden, Förderung und Bildung zum Ausgangspunkt und Fluchtpunkt nehmen. In diesem Sinne soll hier definiert werden: Pädagogische Qualität in einem Kindergarten (oder einer anderen pädagogischen Umwelt wie z.B. Familie, Krippe, Kindertagespflege) ist dann gegeben, wenn die jeweiligen pädagogischen *Orientierungen, Strukturen* und *Prozesse* das körperliche, emotionale, soziale und intellektuelle Wohlbefinden und die Entwicklung und Bildung der Kinder in diesen Bereichen aktuell wie auch auf Zukunft gerichtet fördern und die Familien in ihrer Betreuungs- und Erziehungsaufgabe unterstützen (vgl. Tietze et al. 1998, S. 20).

Dieser Begriff pädagogischer Qualität hat mehrere Implikationen:

1. Pädagogische Qualität ist ein mehrdimensionales und quantifizierbares Konstrukt. Die einzelnen Qualitätsdimensionen, und damit dann auch die pädagogische Qualität insgesamt, können in unterschiedlichen Ausprägungsgraden gegeben sein.
2. Pädagogische Qualität ist nicht länger eine „gefühlte“ Entität, sondern ein multidimensionales Konstrukt, das in seinen Einzelkomponenten vom Anspruch her messbar ist.
3. Pädagogische Qualität und ihre Einzelkomponenten sind validierbar, indem behauptete oder vermutete Effekte auf Entwicklungs- und Bildungsergebnissen der Kinder in den verschiedenen Bereichen untersucht werden können.

4. Der damit unterlegte Begriff pädagogischer Qualität kann mehr oder weniger unmittelbar für Qualitätsverbesserungen genutzt werden. Über Bildungsauscomes validierte Qualitätsdimensionen bilden Stellschrauben zur konkreten Qualitätsverbesserung und -sicherung.<sup>1</sup>

Die verschiedenen in dieser Definition angesprochenen pädagogischen Qualitätsbereiche und ihr Zusammenspiel können im Anschluss an eine grundlegende Konzeptualisierung, wie sie auch im Zwölften Kinder- und Jugendbericht (12. KJB 2006, S. 199f.) enthalten ist, so dargestellt werden (vgl. Abb. 1):

*Orientierungsqualität:* Dieser Qualitätsbereich beinhaltet die normativen Orientierungen, Leitideen und Überzeugungen, unter denen das pädagogische Handeln (Prozessqualität) stattfindet bzw. stattfinden soll. Mit den Curricula für den Vorschulbereich, die die Bundesländer (mit unterschiedlicher Reichweite und unterschiedlichem Verbindlichkeitsgrad) in den letzten Jahren implementiert haben (vgl. 12. KJB 2006, S. 205ff.), ebenso mit den Kriterien der Nationalen Kriterienkataloge, wie sie im Rahmen der Nationalen Qualitätsinitiative des BMFSFJ entwickelt wurden (vgl. Tietze/Viernickel 2007), liegen solche Orientierungen vor. Zu den Komponenten der Orientierungsqualität gehören auch die einrichtungsspezifischen pädagogischen Konzeptionen, die zunehmend als Orientierungsrahmen für das konkrete pädagogische Handeln in den Einrichtungen gefordert werden, ebenso wie eine hinreichende berufsbegleitende Fortbildung des pädagogischen Personals, die den Anschluss der vor Ort gegebenen Orientierungen an den fachlichen Stand sichern helfen soll.

*Strukturqualität:* Dieser Qualitätsbereich beinhaltet die von außen vorgegebenen räumlich-materialen, sozialen und personalen Rahmenbedingungen in den Einrichtungen, die Möglichkeiten für pädagogische Prozesse eröffnen, aber auch Grenzen setzen. Platz innen und außen pro Kind, Erzieher-Kind-Schlüssel, Gruppengrößen, Qualifikation des pädagogischen Personals wie auch rechtlich gesicherte Vor- und Nachbereitungszeiten für die pädagogische Arbeit bilden zentrale Komponenten der Strukturqualität.

*Prozessqualität:* Pädagogik ist ein prozesshaftes Geschehen. Prozessqualität beschreibt die realisierte Pädagogik, die Anregungen, die die Kinder in den einzelnen Bildungs- und Entwicklungsbereichen erhalten, die Art der Interaktionen zwischen dem pädagogischen Personal und den Kindern wie auch den Kindern untereinander. Es geht um einen Umgang mit dem Kind, der seiner Sicherheit und Gesundheit verpflichtet ist, um Interaktionen, die für entwicklungsgemäße Anregungen sorgen, abgestellt auf den individuellen Entwicklungs- und Interessenshorizont des Kindes, um Aktivitäten, die seine emotionale Sicherheit und sein Lernen unterstützen, und um ein pädagogisches Handeln, das zugleich um eine Erziehungspartnerschaft mit den Eltern bemüht ist.

1 Dieses Qualitätskonzept steht im Gegensatz zu sog. relativistischen Ansätzen, die Qualität als ein Verständigungskonstrukt fassen, das individuell im Dialog der beteiligten Akteure jeweils neu zu bestimmen ist (vgl. Dahlberg/Moss/Pence 1999). Eine kritische Auseinandersetzung mit diesem relativistischen Konstrukt ist nicht Gegenstand dieses Beitrags.

*Outcomes:* Von den drei Qualitätsbereichen Orientierungen, Strukturen, Prozesse bzw. den einzelnen Komponenten dieser Qualitätsbereiche werden Auswirkungen auf den Bildungs- und Entwicklungsstand der Kinder in den verschiedenen Bereichen (Domänen) erwartet. Ebenso können Auswirkungen bei Eltern erwartet werden, und über diese bei den Kindern (indirekte Effekte). Die Effekte bei den Eltern und die Bildungs- und Entwicklungsergebnisse bei Kindern, die kurz-, mittel- oder langfristiger Art sein können, werden hier als Outcomes bezeichnet. Solche Outcomes dürfen nicht als deterministische Konsequenzen gegebener Qualitätsindikatoren missverstanden werden. Bildung bleibt auch in dieser Konzeptualisierung ein durch Selbstbeteiligung des Individuums charakterisierter Prozess, der zudem durch andere pädagogische Erfahrungsfelder (besonders die Familie des Kindes) maßgeblich mitbestimmt wird. Es wird jedoch davon ausgegangen, dass der Bildungs- und Entwicklungsstand der Kinder über eine Wahrscheinlichkeitsfunktion mit der pädagogischen Qualität der Kindertageseinrichtung verbunden ist.

*Organisation und Management:* Die Annahme eines nicht deterministischen, sondern probabilistischen Zusammenhangs gilt auch für die Beziehung zwischen den (vorgelagerten) Bedingungen der Orientierungs- und Strukturqualität mit der Prozessqualität. Dieselben Bedingungen der Struktur- und Orientierungsqualität können je nach Organisation und Management in einer Einrichtung unterschiedlich verarbeitet werden. Sie determinieren damit die Prozessqualität nicht; sie machen jedoch – je nach Ausprägungsgrad – eine bestimmte Prozessqualität mehr oder weniger wahrscheinlich.

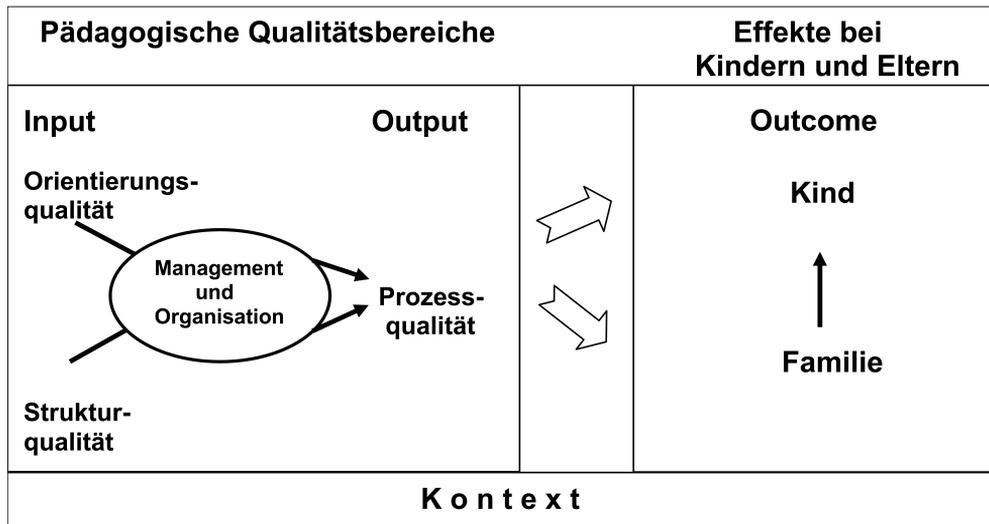


Abbildung 1: Bezugsrahmen zur Spezifizierung pädagogischer Qualität und ihrer Effekte in Kindertageseinrichtungen (in Anlehnung an Tietze et al. 1998, S. 30ff. und 12. KJB 2006, S. 415).

*Kontext:* Gleiches gilt für den Kontext, innerhalb dessen die einzelnen Kindertageseinrichtungen mit ihrer jeweiligen Qualität der Orientierungen, Strukturen und Prozesse

arbeiten. Von externen Stützbedingungen für Kindertageseinrichtungen wie Vorhandensein oder Dichte von Fachberatung, Unterstützung durch die Trägerorganisation u.ä. können Impulse für die Qualität erwartet werden.

Zusammengenommen können damit Kindertageseinrichtungen als pädagogische Dienstleistungsorganisationen betrachtet werden, die unter Nutzung eines bestimmten *Inputs* an strukturellen (*Strukturqualität*) und normativen Bedingungen (*Orientierungsqualität*) – vermittelt über Management- und Organisationsformen – einen bestimmten *Output* pädagogischer Prozesse, d.h. die konkrete Dienstleistung der Bildung, Betreuung und Erziehung der Kinder hervorbringen. Intern stellt sich dabei die Prozessqualität (die realisierte Pädagogik) als abhängig von den Bedingungen der Struktur- und Orientierungsqualität dar. Nach außen bilden alle drei Qualitätsbereiche Bedingungen für Outcomes bei Eltern und Kindern, bei letzteren in Form von direkten und indirekten Effekten.

Die hier skizzierte Konzeptualisierung wird für die weitere Erörterung der Qualitätsfragen zum Ausgangspunkt genommen. Dies geschieht unter drei Gesichtspunkten:

1. In einem ersten Abschnitt soll der Frage nachgegangen werden, inwieweit die vorherrschenden Bildungsforschungsansätze im Vorschulbereich und besonders die in Deutschland vorliegenden empirischen Untersuchungen wissenschaftliche Grundlagen für systematische Qualitätsverbesserungen liefern (können).
2. In einem zweiten Schritt soll geprüft werden, wo in dem zugrunde liegenden Ordnungsrahmen die gegenwärtig in Deutschland vorfindbaren Qualitätsverbesserungsansätze anzusiedeln sind, welche Qualitätsbereiche bzw. Qualitätsdimensionen thematisiert werden, welche nicht.
3. In einem abschließenden dritten Teil soll mit dem Deutschen Kindergarten Gütesiegel (DKGS) ein trägerunabhängiges Prüfverfahren vorgestellt werden, das die zugrunde gelegte Qualitätskonzeption zum Ausgangspunkt nimmt und der Sicherung pädagogischer Qualität bei der einzelnen Kindertageseinrichtung ebenso dient wie einem umfassenden Qualitätsmonitoring des Früherziehungssystems und damit eine Grundlage für eine politisch verantwortbare Qualitätssteuerung bilden könnte.

### 3. Forschungsansätze

Zur Bedeutung von Kindertageseinrichtungen überhaupt und ihrer pädagogischen Qualität im Besonderen für die Entwicklung und Bildung von Kindern existiert mittlerweile eine rund vier Jahrzehnte umfassende Forschungstradition. Allerdings haben sich in diesem Zeitraum Fragestellungen grundlegend verändert und Untersuchungsdesigns vielfach differenziert. Während wir heute vorwiegend nach dem positiven Beitrag von Kindertageseinrichtungen (oder auch Kindertagespflege) für die Bildung und Entwicklung von Kindern fragen (bzw. positive und potenziell problematische Effekte als Grundhypothesen nicht ausschließen), rückte eine frühe *erste Welle* empirischer Untersuchungen potenziell *schädigende* Effekte in den Mittelpunkt der Betrachtung. Hayes

et al. (1990, S. 47) sprechen von einer „Alarmphase“, mit der Forschung auf die neue gesellschaftliche Herausforderung einer zunehmenden „Fremdbetreuung“ besonders von jungen Kindern reagiert habe. Für den Kindergartenbereich in Deutschland scheint diese Frage obsolet geworden zu sein, seit der Deutsche Bildungsrat (1970) den Elementarbereich zur grundlegenden Stufe des gesamten Bildungssystems erklärte, für den Bereich der unter Dreijährigen ist diese Betrachtungsweise allerdings durchaus präsent.

Forschungsmethodisch waren und sind die Untersuchungen dieser *ersten Welle* für die Qualitätsfrage nur bedingt ergiebig. Es werden die Bildungs- und Entwicklungsergebnisse der Kinder einer Betreuungsform (z.B. Kindergarten), mit denen von Kindern einer anderen Form (z.B. Vorklasse), oder denen von Nur-Familienkindern verglichen. Nicht nur besteht die Gefahr, dass Unterschiede schlicht durch Selektionseffekte (vgl. Burchinal 1999) und nicht durch ein unterschiedliches pädagogisches Programm bedingt sind. Selbst wenn man Effekte bei guter statistischer Kontrolle als „wahr“ anerkennen kann, bleibt das Programm weitgehend ein „schwarzer Kasten“, weil unklar und unbeantwortet ist, auf welche Programmkomponenten (Qualitätsaspekte) im Einzelnen die Fördereffekte zurückgeführt werden können. Die im Rahmen der Bildungsreform der 1970er Jahre groß angelegten Untersuchungen zu „Kindergarten vs. Vorklasse“ (vgl. Ewert/Braun 1978; Dollase 1978, 1979; Winkelmann et al. 1977) sind Beispiele für diese Forschungswelle in Deutschland.

Eine *zweite Welle* von Forschungen nimmt einen Teil dieser Defizite auf, indem nicht mehr nur globale Typen (z.B. auch in Form von Modellprogrammen) untersucht werden, sondern Programme in verschiedenen Qualitätsdimensionen spezifiziert werden, die es im Einzelnen zu erfassen gilt. Die soziale Wirklichkeit der Bildungs- und Betreuungseinrichtungen mit ihrer großen Variationsbreite in bildungsrelevanten Einzelmerkmalen bildet hier den Ausgangspunkt der Untersuchungen. Spezifiziert werden in solchen Untersuchungen sowohl Merkmale der *Strukturqualität* wie Gruppengrößen, Erzieher-Kind-Schlüssel, räumlich-materiale Gegebenheiten oder Ausbildung des pädagogischen Personals als auch Merkmale der *Prozessqualität* wie Bildungsanregungen für die Kinder in verschiedenen Bereichen, Sensitivität des Umgangs mit den Kindern und andere Merkmale des pädagogischen Interaktionsklimas (vgl. Vandell/Wolfe 2000). Die zentrale Frage dabei lautet, ob und in welchem Umfang bestimmte Qualitätsmerkmale einen Einfluss auf Bereiche kindlicher Bildung und Entwicklung haben. Vom Typus her handelt es sich um Korrelationsstudien in einem „naturalistischen“ Feld mit seiner Variabilität der Bedingungen. Allerdings ist dieser Forschungsansatz prinzipiell auch offen für auf Innovation gerichtete experimentelle Designs, wie eine frühe Untersuchung von Ruopp, Travers, Glantz and Coelen (1979) zeigt, in der die Faktoren Gruppengröße und Erzieher-Kind-Schlüssel systematisch variiert wurden.

Die Qualität in den Betreuungsformen wie auch ihre Auswirkungen auf Kinder können nicht losgelöst von den familiären Bedingungen der Kinder betrachtet werden. Es spricht manches dafür, dass familiäre Bedingungen die Auswahlentscheidungen für Betreuungsformen beeinflussen (vgl. z.B. Burchinal 1999) und auch Interaktionseffekte zwischen Familie und Einrichtung auf Bildung und Entwicklung der Kinder anzunehmen sind, auch wenn sich die Forschungslage hierzu uneinheitlich darstellt (Rossbach

2005). Abgesehen davon bildet die Familie für sich genommen eine wirkmächtige Instanz im Sinne eines Haupteffekts für Bildung und Entwicklung des Kindes. Dies verweist auf eine aktuelle *dritte Forschungswelle*, in der die bildungsrelevanten Qualitäten des familienexternen Bildungs- und Betreuungssettings *und* die der Familie jeweils in ähnlicher Breite und Tiefe spezifiziert werden. Solche Untersuchungsansätze ermöglichen die Identifizierung wirkmächtiger Bedingungen in Familie *und* Einrichtung wie auch vergleichende Abschätzungen von Effektgrößen. Die erste (und bislang einzige veröffentlichte) Studie im deutschsprachigen Raum im Rahmen dieses Konzepts findet sich bei Tietze et al. (1998), wobei mittlerweile auch eine vierjährige längsschnittliche Ausweitung vorliegt (Tietze et al. 2005).

Zu den aktuellen Forschungskonzeptionen dieser dritten Welle gehört neben der Zunahme längsschnittlicher Betrachtungen bei allen größeren Studien (vgl. Roßbach 2005), dass die Förderung der Kinder in vorschulischen Einrichtungen als ein dynamischer Ablauf betrachtet wird, in dem es neben der Form und der pädagogischen Qualität mit ihren verschiedenen Dimensionen auch das Eintrittsalter des Kindes, tägliche Dauer, Wechsel von Settings und kumulierte Betreuungsumfänge zu berücksichtigen gilt (NICHD 2005, S. 3; Roßbach 2005, S. 58ff. u. 69ff.).

Das Zusammenspiel von Einrichtungen der Bildung, Betreuung und Erziehung von Kindern und der Familie für den Bildungserfolg von Kindern kann sich über die genannten Zusammenhänge hinaus noch komplexer darstellen. Dies gilt dann, wenn die Förderintention der Einrichtungen nicht nur *direkt* auf Bildungs- und Entwicklungsförderung der Kinder gerichtet ist, sondern eine explizite Elternbildungskomponente (z.B. in Form von Hausbesuchen oder sonstigen Formen der Unterstützung und Beratung von Eltern bei ihrer Erziehungsaufgabe) mit enthalten ist, um auf diese Weise über die Familie eine zusätzliche *indirekte* Bildungs- und Entwicklungsförderung der Kinder zu erreichen.

Mit der Übertragung des Konzepts der englischen Early Excellence Centres auf Deutschland (Bertram et al. 2002; Gerhold/Hebenstreit-Müller/Kühnel 2001), den verschiedenen, bundesweit streuenden Ansätzen zur Etablierung sog. Familienzentren (Peucker/Riedel 2004), hier besonders dem großflächigen Ausbau von Kindertageseinrichtungen zu Familienzentren wie gegenwärtig in NRW (MGFFI 2007) deutet sich die Notwendigkeit einer *vierten* Forschungswelle an, die mögliche *indirekte Effekte* der Einrichtungen auf die Bildungs- und Entwicklungsförderung von Kindern thematisiert, die über die von Einrichtungen ausgehende Bildung, Beratung und Unterstützung von Familien vermittelt sind. Ansätze einer diesbezüglichen Wirkungsforschung sind in Deutschland allerdings bisher nicht erkennbar.

Die hier skizzierten Forschungsansätze sind teilweise im Sinne einer *chronologischen Abfolge* zu verstehen, teilweise existieren sie aber auch – in Abhängigkeit von der jeweiligen Fragestellung – parallel nebeneinander. Die Unterschiede in der Fruchtbarkeit für die Qualitätsfrage sind evident.

#### 4. Forschungsergebnisse in Deutschland

In Deutschland fehlt es weitgehend an einer kritischen Masse an Studien, um differenzierte, durch verschiedene Studien begründete Aussagen zu den Bildungseffekten der Elementarerziehung bzw. institutionell-vorschulischer Erziehung überhaupt machen zu können. Am ehesten finden sich in Deutschland Studien, die der *ersten Forschungswelle* zuzurechnen sind, indem gefragt wird, ob dem Besuch eines Kindergartens, unabhängig von seiner Qualität, ein Bildungseffekt bei den Kindern zugeschrieben werden kann (Black Box Paradigma).

Tietze (1987) fand in einer älteren, für Nordrhein-Westfalen repräsentativen Untersuchung, dass – bei Kontrolle anderer Faktoren – aus Grundschulklassen mit einem höheren Anteil kindergartengeförderter Kinder die Quoten der Zurückstellungen bei Schulbeginn, der Sitzenbleiber sowie der Übergänger zur Sonderschule für Lernbehinderte während der Grundschulzeit geringer ausfielen. In der aktuellen IGLU-Studie (Bos et al. 2003, S. 127ff.) ergab sich, dass die Schulleistungen bei Viertklässlern, die den Kindergarten über ein Jahr besucht hatten, höher ausfielen als bei den anderen Kindern. Büchner und Spieß (2007) fanden in einer aktuellen Studie auf der Grundlage des SOEP, dass die Wahrscheinlichkeit, eine Hauptschule zu besuchen, mit zunehmender Dauer des Kindergartenbesuchs abnimmt. Anger, Plünnecke und Tröger (2007) schätzen in einer Studie des Instituts der deutschen Wirtschaft auf der Grundlage der PISA-Daten ab, dass bei einem „Gesamtkonzept“ frühkindlicher Bildung mit qualitativen Verbesserungen der Anteil der jungen Erwachsenen ohne abgeschlossene Berufsausbildung um 6 Prozentpunkte auf 10 Prozent gesenkt und der Anteil der Hochqualifizierten um gut 6 Prozentpunkte auf 37 bis 38 Prozent erhöht werden könne. Zusammengekommen sprechen diese Befunde dafür, dass ein Kindergartenbesuch kurz- und langfristige Effekte in robusten Bildungsindikatoren zeitigt.

Die genannten Studien sind Repräsentanten der ersten Forschungswelle. Institutionelle Früherziehung geht als „Schwarzer Kasten“ in die Untersuchungen ein. Die Untersuchungsergebnisse liefern kaum Ansätze für Qualitätsverbesserungen; denn qualitätsrelevante Einzelindikatoren von Einrichtungen und Programmen werden in diesem Forschungsparadigma nicht thematisiert.

Untersuchungen, die Grundlagenwissen für Qualitätsverbesserungen und Qualitätssicherung liefern sollen, setzen Designs wenigstens der zweiten, besser noch der dritten Forschungswelle voraus. Mit der Studie „Wie gut sind unsere Kindergärten“ und ihrer längsschnittlichen Ausweitung „Kinder von 4 bis 8 Jahren (Tietze et al. 1998, 2005) als Teile der European Child Care and Education Study (ECCE Study Group 1999), liegt die bislang einzige breit angelegte deutsche Untersuchung zu den Effekten *pädagogischer Qualität* im Kindergarten vor, die sich diesem Anspruch stellt.

Im Rahmen eines Untersuchungsdesigns, das nach den Qualitätsbereichen *Orientierungs-, Struktur- und Prozessqualität* in Einrichtung und Familiensetting differenziert, entsprechende Qualitätsdimensionen spezifiziert und damit der *dritten Forschungswelle* zuzurechnen ist, werden interne Abhängigkeiten der Prozessqualität von vorgelagerten Bedingungen der Struktur- und Orientierungsqualität analysiert und es werden Effekte

pädagogischer Qualität auf verschiedene Bildungs- und Entwicklungsmaße im Vorschul- und Grundschulalter (unter Einschluss von Effektstärken) abgeschätzt. Nach den Ergebnissen dieser Studie können je nach Maß zwischen 25% und 50% der Varianz der pädagogischen Prozessqualität durch vorgelagerte Bedingungen der Struktur- und Orientierungsqualität erklärt werden, ein Befund, der sich in ähnlicher Weise auch in anderen Ländern zeigt (vgl. Cryer et al. 1999). Der Befund lässt sich so interpretieren, dass Verbesserungen der Prozessqualität zum Teil durch verbesserte Inputbedingungen erreicht werden können, zu einem anderen Teil aber von einer direkten auf Prozessqualität gerichteten Intervention abhängig sind. Zu den weiteren Ergebnissen dieser Studie gehören, dass Bildungs- und Entwicklungsunterschiede von bis zu einem Jahr bei Kindergartenkindern auf die *Qualitätsunterschiede* in den Einrichtungen zurückgeführt werden können, und dass auf unterschiedliche Kindergartenqualität zurückgehende Bildungsunterschiede auch am Ende der zweiten Grundschulklasse zu verzeichnen sind. Diese bewegen sich in derselben Größenordnung, wie die Effekte, die auf unterschiedliche Grundschulqualität der Zweitklässler zurückgehen. Effekte finden sich dabei im sprachlichen, kognitiv-schulleistungsbezogenen und im sozialen Bereich. Die Ergebnisse zeigen aber auch, dass der Effekt der pädagogischen Qualität der Familie rund zwei- bis dreimal so groß ist wie der des institutionellen Settings (vgl. hierzu auch Vandell/Wolfe 2000).

Aktuelle Forschungsansätze in Deutschland sind auf weitere Differenzierungen einer solchen Forschungskonzeption gerichtet, indem *domainspezifische* Merkmale von Kindertagesstättensettings mit *domainspezifischen* Bildungs- und Entwicklungsergebnissen, im Rahmen einer kleingliedrigen längsschnittlichen Modellierung von Prozessverläufen, miteinander in Beziehung gesetzt werden (BiKS) (vgl. Universität Bamberg 2007). Allerdings sind hier noch keine Befunde publiziert.

## 5. Prozessqualität messen

Eine nicht unerhebliche Hürde bei Untersuchungen mit differenziertem Qualitätskonzept besteht darin, die Qualität der pädagogischen *Prozesse* zu erfassen. Die Prozessqualität stellt die direkte Schnittstelle zum Kind (und auch zu den Eltern) dar, sie ist von besonderer Bedeutung, erfordert andererseits einen beträchtlichen Erhebungsaufwand. Während sich die meisten Dimensionen der Struktur- und Orientierungsqualität über vergleichsweise unaufwändige Verfahren wie Befragung, Interview und Dokumentenanalyse erheben lassen, setzt die Erfassung der Prozessqualität direkte Beobachtungen voraus. Klassische und international vielfach genutzte Instrumente zur Erfassung pädagogischer Prozessqualität sind die Early Childhood Environment Rating Scale (ECERS) sowie die Infant/Toddler Environment Rating Scale (ITERS) von Harms, Clifford und Cryer. Für den Bereich der Kindertagespflege wurde die Family Daycare Rating Scale von Harms und Clifford entwickelt. Unter der Bezeichnung Kindergarten-Skala (KES-R), Krippen-Skala (KRIPS-R und Tagespflege-Skala (TAS) liegen auch deutsche Adaptionen dieser Instrumente vor (Tietze et al. 2007 a, b, c). Die Skalen erfassen neben

Platz und Ausstattung sowie Kooperation mit Eltern die Bereiche Betreuung und Pflege der Kinder, sprachliche und kognitive Anregungen, Aktivitäten, Interaktionen und Strukturierung der pädagogischen Arbeit. Sie erfordern eine wenigstens dreistündige Beobachtung durch einen gut trainierten Beobachter. Eine im engeren Sinne auf den Bildungsgehalt der Kindertagesumwelt ausgelegte und als Ergänzung zur ECERS konzipierte Skala wurde unter der Bezeichnung ECERS-E von Sylva et al. vorgelegt (2006, deutsche Forschungsversion beim Autor).

Darüber hinaus sind weitere Verfahren vorhanden, um unterschiedliche Aspekte der Prozessqualität zu erfassen (vgl. Clarke-Stewart/Allhusen 2005; NICHD 2005, S. 3ff., 78ff.; Tietze 2006, Vandell/Wolfe 2000, S. 3ff.). Gemeinsam ist den Instrumentarien, dass sie auf direkten Beobachtungen beruhen.

## 6. Ansätze für Qualitätsverbesserungen

Bildungs- und entwicklungsprädiktive Qualität ist ein komplexes Konstrukt bestehend aus unterschiedlichen Qualitätsbereichen mit jeweils verschiedenen Qualitätsdimensionen. Die bisherige Forschung zeigt, dass es offensichtlich nicht *die* Qualitätsvariable gibt, die als mehr oder weniger alleinige Stellschraube zur Verbesserung von Outcomes bei Kindern (und Eltern) benutzt werden könnte. Die frühe Einsicht in der frühpädagogischen Qualitätsforschung „Good things tend to go together“ (Phillips 1987) legt nahe, dass nachhaltige Qualitätsverbesserungen mit der Wahrscheinlichkeit verbesserter Bildungsergebnisse bei den Kindern nur über die Betätigung mehrerer Stellschrauben unter Berücksichtigung aller Qualitätsbereiche möglich sind. Zugleich ist dabei der oben genannte steuerungsrelevante Zusammenhang zwischen den Qualitätsbereichen zu beachten (vgl. Abb. 1), wonach die vorgelagerten Bedingungen des Inputs (Orientierungs- und Strukturqualität) bis zu 50% der Varianz der Prozessqualität bestimmen, die restlichen 50% aber offenbar davon unbeeinflusst sind und ihre Varianzquellen innerhalb der Einrichtungen selbst haben.

Bezieht man die aktuell praktizierten Qualitätsverbesserungs- und -steuerungsansätze auf den dargestellten Zusammenhang, so lässt sich Folgendes feststellen:

*Staatliche Inputsteuerung:* Die gegenwärtigen staatlichen Qualitätsverbesserungsansätze sind schwerpunktmäßig auf eine Verbesserung des Inputs gerichtet. Für den Bereich der *Orientierungsqualität* sind hier die Bildungspläne der Bundesländer zu nennen (vgl. 12 KJB 2006, S. 205ff.), die im Umkreis des Beschlusses der Jugendministerkonferenz „Gemeinsamer Rahmen der Länder für die frühe Bildung in Kindertageseinrichtungen“ (JMK 2004) erstmals in Deutschland entwickelt wurden, mittlerweile in allen Bundesländern zum Standard gehören. Im Bereich der *Strukturqualität* sind – unterstützt durch eine breite Förderinitiative der Bosch-Stiftung – Anstrengungen zur Verbesserung der Ausbildung des pädagogischen Personals durch Anhebung auf Fachhochschulniveau zu verzeichnen (Robert Bosch Stiftung 2007). Mittlerweile sind an einer beachtlichen Anzahl von Hochschulen entsprechende Ausbildungsgänge eingerichtet worden (Deutscher Bildungsserver 2007).

Eine kritische Wertung dieses inputorientierten Ansatzes der Qualitätsverbesserung wird in Rechnung stellen müssen, dass (mit Bezug auf die berichteten Befunde zur Abhängigkeit der Prozessqualität vom vorgelagerten Input) auf diese Weise die Qualität pädagogischer Prozesse, also die konkrete Pädagogik vor Ort, nur bedingt angehoben werden kann. Zudem fallen die Ergebnisse einer aktuellen, mehrere Bundesstaaten umfassenden Studie in den USA zu den Auswirkungen unterschiedlicher Qualifikationsniveaus des pädagogischen Personals auf die Prozessqualität und auf Outcomes bei Kindern ausgesprochen desillusionierend aus (Early et al. 2007): Die erwarteten Auswirkungen konnten nicht gefunden werden. Vor diesem Hintergrund mag es erstaunen, dass weder staatliche Stellen noch die in diesem Zusammenhang sehr engagierte Bosch-Stiftung an der Überprüfung möglicher Ausbildungseffekte für verbesserte Prozessqualität oder verbesserte Bildungsergebnisse bei den Kindern interessiert zu sein scheinen. Bislang jedenfalls sind keine Absichten für entsprechende Evaluationen bekannt geworden. Nach dem gegenwärtigen Stand muss offen bleiben, ob von den unternommenen Anstrengungen etwas bei den Kindern im Alltag der von ihnen erfahrenen Prozessqualität in den Einrichtungen oder bei ihren Bildungsergebnissen ankommt.

*Organisation und Management der Träger:* Staatlich regulierte Inputbedingungen können von Einrichtungen und ihren Trägern unterschiedlich „verarbeitet“ werden (vgl. Abb. 1). Das gilt im Hinblick auf Effizienz wie auch auf Ziele bei der Prozessqualität, für die die Inputbedingungen genutzt werden. Es nimmt daher nicht Wunder, dass gerade die großen Trägerverbände von Tageseinrichtungen die Organisations- und Managementebene als Ansatzpunkt für Qualitätsverbesserung „besetzt“ haben. Dieser Ansatz erlaubt ihnen, die vorgelagerten staatlichen Bedingungen einerseits anzuerkennen, andererseits für eigene durchaus auch partikuläre Zielstellungen durch ein darauf abgestelltes, effizientes Managementsystem zu nutzen. Pädagogische Qualität wird dabei trägerspezifisch definiert, niedergelegt in trägerspezifischen Qualitätshandbüchern. Ein von Trägerinteressen unabhängiger Begriff pädagogischer Prozessqualität hat in dieser Konzeptualisierung keinen Platz.

Die Verbesserung von Management und Organisation orientiert sich meistens am Qualitätsmanagementkonzept der DIN EN ISO 9000ff., einem Regelwerk zum Qualitätsmanagement, das in den 1980er Jahren für die Anwendung in der industriellen Produktion entwickelt wurde, später auch für Dienstleistungsorganisationen adaptiert wurde (Zollondz 2002). Die gelungene Implementierung dieses Qualitätsmanagementsystems kann durch dafür zugelassene Zertifizierer begutachtet und zertifiziert werden.

Mittlerweile haben alle großen Trägerverbände Qualitätsmanagementsysteme eingeführt, die sich mehr oder weniger eng an der Normenreihe der ISO orientieren. Für Einrichtungen in katholischer Trägerschaft wurde vom Verband Katholischer Tageseinrichtungen für Kinder (KTK)-Bundesverband (2004) das KTK-Gütesiegel entwickelt, für Einrichtungen, die dem Paritätischen Wohlfahrtsverband angeschlossen sind, das „PARITÄTISCHE Qualitätssystem®“, für Einrichtungen in Trägerschaft der Arbeiterwohlfahrt das System „AWO-QM“, für evangelische Einrichtungen liegt ein Bundes-Rahmenhandbuch als Leitfaden zur Qualitätsentwicklung vor (vgl. im Überblick Esch et al. 2006, S. 129ff.).

Die Einführung des jeweiligen Qualitätsmanagementsystems kann, muss aber nicht mit einer Zertifizierung nach ISO 9000 abschließen. Die Bedeutung einer solchen Zertifizierung wird in der Fachöffentlichkeit nicht immer richtig wahrgenommen. Es handelt sich nicht um ein Zertifikat, das die allgemeine Güte der pädagogischen Dienstleistung bescheinigt, denn eine solche wird nicht definiert, und macht auch keine Aussage darüber, ob die im Qualitätshandbuch beschriebenen Ziele konkret erreicht werden. Es liegt keine Produktzertifizierung vor. Zertifiziert wird lediglich, dass die Organisation über ein Managementsystem verfügt, nach dem die im Qualitätshandbuch beschriebene Qualität der pädagogischen Dienstleistung (das was die Einrichtung bzw. ihre Trägerorganisation darunter verstehen will), erreicht werden kann.

Es gibt gegenwärtig keine allgemein zugänglichen Daten, in wie vielen Einrichtungen nach diesen Managementsystemen bundesweit gearbeitet wird, und wie viele Einrichtungen zertifiziert sind. Ebenso lässt sich die Frage nicht beantworten, in welchem Umfang bzw. ob überhaupt die Qualität der pädagogischen Prozesse, also die Qualität der bildungs- und entwicklungsanregenden Erfahrungen der Kinder in den Einrichtungen, bei diesem Ansatz verbessert wird und ob es Effekte bei den Bildungsergebnissen der Kinder gibt. Ausländische Erfahrungen gemahnen zur Vorsicht. In einer kalifornischen Untersuchung an Kindertageseinrichtungen, die nach dem fach(!)-spezifischen System der National Association for the Education of Young Children (NAEYC) zertifiziert waren, zeigte sich eine erhebliche Anzahl von zertifizierten Einrichtungen mit nur mittelmäßiger und auch unzureichender Prozessqualität.

*Direkte Verbesserung der Prozessqualität:* Die nur indirekten und begrenzten Möglichkeiten der Verbesserung der Prozessqualität durch verbesserte Inputbedingungen und der zweifelhaften, weil ohne nachgewiesene Effekte für eine verbesserte Prozessqualität praktizierte Ansatz eines allgemein verbesserten Qualitätsmanagements, legen den Gedanken einer direkten Verbesserung der pädagogischen Prozessqualität durch In-Haus-Fortbildung und Training in den Einrichtungen nahe. Technisch gesprochen geht es darum, die nicht durch vorgelagerte Inputbedingungen der Orientierungs- und Strukturqualität bestimmten Varianzanteile der Prozessqualität aufzugreifen und diese „hausgemachten“ Anteile für Verbesserungen zu nutzen.

In einer gewissen Weise steckt diese Grundidee hinter vielen Fortbildungsmaßnahmen. Mit der vom BMFSFJ ins Leben gerufenen und geförderten Nationalen Qualitätsinitiative wurde sie systematisch entwickelt und umgesetzt (vgl. 12. KJB 2006). Im Kern galt es – bei unterschiedlicher methodischer Ausrichtung – für die an der Nationalen Qualitätsinitiative beteiligten Institute einen fachlichen Konsens herzustellen, was unter guter pädagogischer Prozessqualität zu verstehen ist und diese in Kriterienkatalogen praxisnah zu beschreiben, Verfahren und Methoden einer personalgetragenen Verbesserung pädagogischer Prozessqualität im Hinblick auf diese Kriterien zu entwickeln und zu erproben und die so entstandenen Prozeduren in einer Implementierungsphase möglichst breit im Feld zu verankern.

Über die Nationale Qualitätsinitiative konnten nach dem BMFSFJ vorliegenden Angaben mehrere tausend Kindertageseinrichtungen aus allen Bundesländern mit unterschiedlichen Intensitätsgraden erreicht werden. Die Qualitätsentwicklungskurse reich-

ten von zweitägigen zentralisierten Veranstaltungen bis zu zweijährigen systematischen Kursen unter Einbeziehung aller Mitarbeiter vor Ort und stützt sich teilweise auf Multiplikatorenmodelle. Bedauerlicherweise waren weder das BMFSFJ noch die Mehrheit der an der Nationalen Qualitätsinitiative beteiligten Institute an einer Evaluation interessiert, die diesen Namen verdient. Daher bleibt auch hier die Frage, welche Verbesserung der Qualität der pädagogischen Prozesse für die Kinder durch diese direkte Intervention erzielt wird, offen. Informelle und noch unveröffentlichte Pre-Post-Messung von PädQUIS, eines der Institute der Nationalen Qualitätsinitiative, zeigen, dass auch bei intensiven, ein- bis zweijährigen Qualitätsentwicklungskursen bei rund 10% der Einrichtungen keine substanziellen Verbesserungen der pädagogischen Prozessqualität gegeben waren, dass im Durchschnitt aber die Verbesserungen mit Effektstärken von  $d \approx 1,5$  (Cohen 1988) nach sozialwissenschaftlichen Kriterien als hoch zu bezeichnen sind.

*Kosten:* Die verschiedenen Verbesserungsansätze pädagogischer Qualität in den Kindertageseinrichtungen, unabhängig davon, ob und welche Effekte sie haben, sind mit zusätzlichen Kosten verbunden. Nach Berechnungen der Sachverständigenkommission des Zwölften Kinder- und Jugendberichts, (12. KJB 2006, S. 216ff.) würden, wenn in jeder Kindertageseinrichtung eine pädagogische Fachkraft auf Fachhochschulniveau ausgebildet und entsprechend bezahlt wird (z.B. die Leiterin), jährliche Mehraufwendungen der öffentlichen Jugendhilfe von 250 Mio. Euro anfallen. Bei einer Verbesserung des Erzieher-Kind-Schlüssels von 1:12,6 (West) bzw. 1:13,4 (Ost) auf 1:10 würden sich die jährlichen Mehraufwendungen auf ca. 2 Mrd. Euro belaufen. Qualitätsverbesserungen nach den QM-Systemen der freien Träger wie auch solche nach den Formen der Nationalen Qualitätsinitiative würden jährlich Mehraufwendungen von unter 100 Mio. Euro nach sich ziehen. Erwartungsgemäß fallen die Mehraufwendungen bei Verbesserung der Inputbedingungen am höchsten aus; sie sind realistischerweise daher nur mittel- und langfristig zu realisieren. Die Kosten für Qualitätsmanagement und direkte Qualitätsverbesserungen der Arbeit in den Einrichtungen fallen demgegenüber – ebenfalls bei einer flächendeckenden Implementierung – deutlich niedriger aus und dürften kurzfristig bewältigbar sein.

## **7. Überprüfung und dauerhafte Sicherung pädagogischer Qualität**

Die verschiedenen Ansätze, die im voran stehenden Abschnitt typisierend dargestellt wurden, verweisen darauf, dass die Frage der Qualitätsverbesserung für die Bildung, Betreuung und Erziehung von Kindern im vorschulischen Alter einen bemerkenswerten Stellenwert in der fachlichen und fachpolitischen Diskussion erlangt hat.

In internationaler Perspektive wie auch in der fachlichen Diskussion in Deutschland gewinnt die Einsicht an Boden, dass in der Verantwortung für die Bildungschancen des einzelnen Kindes wie auch in der Verantwortung vor dem Steuerzahler umfassende Prüfverfahren mit festen Standards in den verschiedenen Qualitätsbereichen (auch unter Einschluss des Outputs als der erreichten pädagogische Prozessqualität), erforderlich

werden, um auf diese Weise die pädagogische Qualität der einzelnen Kindertageseinrichtungen vor Ort überprüfen und sicherstellen zu können. Solche Verfahren sind nicht mit den Prüfverfahren für die Erteilung der Betriebserlaubnis für eine Kindertageseinrichtung nach § 45 KJHG zu verwechseln, die schwerpunktmäßig darauf gerichtet sind, eine Gefährdung des Kindeswohls auszuschließen, also minimale Qualität zu gewährleisten. Es geht vielmehr darum, im positiven Sinn einen hohen Grad an bildungs- und entwicklungsfördernder Qualität zu sichern. Deutschsprachige Überblicke über im Ausland praktizierte Verfahren finden sich bei Esch et al. (2006) sowie bei Tietze (2008). Hinweise und Anbahnungen für systematische Qualitätsüberprüfungen sind mittlerweile auch in neuen Kita-Gesetzen, Finanzierungsgrundsätzen und Förderverfahren von Bundesländern (z.B. Baden-Württemberg, Berlin, Brandenburg, Hamburg) gegeben, wobei die konkrete Ausgestaltung meist offen ist.

Im Folgenden soll abschließend mit dem Deutschen Kindergarten Gütesiegel (DKGS)<sup>2</sup> ein Prüf- und Zertifizierungsverfahren vorgestellt werden, das sich als fachwissenschaftlich basiertes Verfahren versteht und beansprucht, zur Lösung der aufgeworfenen Fragen beizutragen. Die dem DKGS zugrunde liegenden Begründungen, Aufbau- und Prinzipien und die bildungsökonomische Einbettung sind an anderen Stellen ausführlich dargestellt (vgl. Spieß/Tietze 2002; Tietze/Förster 2005). Als Charakteristika können hervorgehoben werden:

*Ausgangspunkte Qualitätsverbesserung und Qualitätsinformation:* Das DKGS nimmt zum Ausgangspunkt, dass (1) das durchschnittlich nur mittelmäßige Niveau pädagogischer Qualität, bei starken Streuungen zwischen den Einrichtungen nicht befriedigen kann, (2) eine ausschließliche Inputsteuerung sowie die Anwendung nur trägerspezifischer Qualitätsmanagementsysteme solche Defizite nicht beheben können, (3) Eltern für die Ausübung ihres ordnungspolitisch gewollten Wunsch- und Wahlrechts (KJHG, § 5) einer Qualitätsinformation bei den zur Auswahl stehenden Einrichtungen bedürfen und (4) es für Kita-Anbieter eines Anreizsystems für gute Qualität bedarf, zumindest in dem Sinne, dass gute Qualität nach außen glaubhaft angezeigt werden kann.

*Wissenschaftsbasiertes Qualitätskonzept:* Das Qualitätskonzept des DKGS versteht sich als wissenschaftsbasiertes Qualitätskonzept. Es schließt an den Stand der Qualitätsforschung an und nimmt die oben beschriebenen Qualitätsbereiche der Orientierungsstruktur und Prozessqualität sowie die Qualität des Familienbezugs explizit auf. Es spezifiziert innerhalb dieser vier Qualitätsbereiche jeweils einzelne Qualitätsdimensionen, von denen durch empirische Untersuchungen belegt oder durch Expertenurteil begründet ist, dass sie relevant sind für erwünschte Bildungs- und Entwicklungsergebnisse bei Kindern (vgl. Abb. 2).

*Qualitätsdimensionen und Standards:* Neben der wissenschaftlich begründeten Spezifizierung von Qualitätsdimensionen bedarf es der Festlegung von Standards (kritischer Niveaus) in diesen Dimensionen. Im Falle des DKGS sind die Standards so ge-

2 Das Deutsche Kindergarten Gütesiegel wurde von PädQUIS, Kooperationsinstitut der Freien Universität Berlin, in Zusammenarbeit mit dem Charlotte-Bühler-Institut, Wien, entwickelt ([www.paedquis.de](http://www.paedquis.de)).

wählt, dass sie gesetzliche Vorgaben nicht unterschreiten und sich an internationalen Empfehlungen orientieren. Die konkrete Standardsetzung ist dem Sinn eines Gütesiegels entsprechend so vorgenommen, dass Kindertageseinrichtungen Gütesiegelfähigkeit tatsächlich erreichen können.

*Konzeptübergreifende Gültigkeit:* Das DKGS begreift sich als ein konzeptübergreifendes Verfahren. Dementsprechend wird nicht überprüft, ob ein bestimmtes pädagogisches Konzept (z.B. Montessoripädagogik, Situationsansatz) angemessen realisiert wird; Überprüfung und Zertifizierung durch das DKGS beziehen sich auf fachwissenschaftlich begründete Qualitätsindikatoren, die für alle Kindertageseinrichtungen ausschlaggebend sind.

*Träger- und bundesländerübergreifende Geltung:* Die fachwissenschaftliche Basierung des DKGS beinhaltet zugleich einen träger- und bundesländerübergreifenden Geltungsanspruch. So wie kommunale, evangelische oder DRK-Krankenhäuser und Krankenhäuser in Flensburg und Füssen sich denselben fachmedizinischen Standards zu stellen haben, gilt nach den dem DKGS zugrunde liegenden Prinzipien der entsprechende Anspruch auch für Kindertageseinrichtungen in orts- und trägerübergreifender Weise.

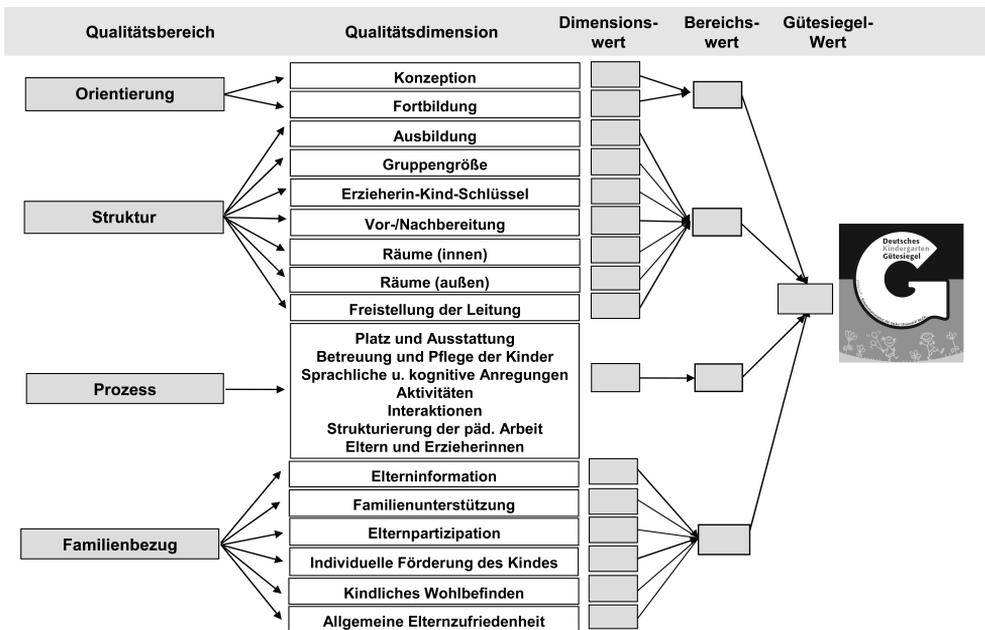


Abbildung 2: Qualitätsbereiche und -dimensionen des Deutschen Kindergarten- Gütesiegels

Nach einer mehrjährigen Entwicklungs- und Erprobungsphase ist die Entwicklung des DKGS vorläufig abgeschlossen. Die Gütesiegelüberprüfung im Feld wird von einer politik- und trägerunabhängigen Instanz vorgenommen. Das Prüfverfahren schließt direkte Beobachtungen und Ratings durch geschulte externe Evaluatoren vor Ort, Befragungen des pädagogischen Fachpersonals, Dokumentenanalysen sowie schriftliche Befragungen

der Eltern ein. Die Ergebnisse werden in jeder Qualitätsdimension auf einer 6-Punkte-Skala abgebildet, wobei ab 3 Punkten Gütesiegelfähigkeit gegeben ist. Dadurch kann im Einzelfall nachvollzogen werden, wie gut das Gütesiegel erreicht wurde bzw. in welchen Dimensionen Gütesiegelfähigkeit noch nicht gegeben ist. Mit jeder Überprüfung einer Kita auf Gütesiegelfähigkeit ist die Erarbeitung und Rückmeldung eines detaillierten Qualitätsprofils verbunden, das als Grundlage für die weitere Qualitätsentwicklung genutzt werden kann. Das DKGS ist für drei Jahre gültig, wobei nach dem ersten und zweiten Jahr ein Zwischenaudit erfolgt. Für einen weiteren Dreijahreszyklus ist eine erneute Hauptprüfung erforderlich.

Bislang wurden Gütesiegelüberprüfungen in den Bundesländern Berlin, Brandenburg, Hamburg, Nordrhein-Westfalen und Schleswig-Holstein vorgenommen. Die Ergebnisse zeigen, dass etwa ein Drittel der Kindertageseinrichtungen das DKGS „aus dem Stand heraus“ erhält. Bei der Mehrheit der Einrichtungen wird Gütesiegelfähigkeit erst nach einer intensiven Phase der Qualitätsentwicklung erreicht.

Die hier vorgestellte Form der Qualitätsüberprüfung und Zertifizierung ist wie jede andere Form nicht ohne zusätzliche Kosten zu realisieren. Die Kosten hängen dabei von der Größe der Einrichtung ab. Nach den vorliegenden Erfahrungen belaufen sich die Kosten bei einer Qualitätssicherung mit dem DKGS auf ca. zwei Euro pro Kind und Monat, ein Betrag, der bei monatlichen Betriebskosten der öffentlichen Jugendhilfe von 350 Euro bei einem Ganztagskindergartenplatz bzw. 740 Euro bei einem Ganztagsplatz für unter drei Jahre alte Kinder (vgl. 12. KJB 2006, S. 213) bewältigbar erscheint. Die Erfahrungswerte liegen in derselben Größenordnung wie der von Dohmen (2005, S. 79) in einer bildungsökonomischen Expertise geschätzte Wert, wonach bei der flächendeckenden Einführung des Deutschen Kindergarten Gütesiegels „allenfalls ein halbes Prozent“ der entsprechenden Jugendhilfekosten zu veranschlagen wäre.

## **Ausblick**

In diesem Beitrag wurde ein empirischer Qualitätsbegriff zugrunde gelegt, nach dem pädagogische Qualität in Kindergärten (wie auch in anderen vorschulischen Bildungs- und Betreuungssettings) ein mehrdimensionales und quantifizierbares Konstrukt darstellt. Das Konstrukt ist rückgebunden an den empirischen Forschungsstand und ist damit offen für zukünftige Entwicklungen. Pädagogische Qualität und ihre Einzeldimensionen sind messbar. Und sie sind validierbar, indem behauptete oder vermutete Effekte auf kindliches Wohlbefinden sowie auf Bildungs- und Entwicklungsergebnisse in den verschiedenen Bereichen untersucht werden können. Pädagogische Qualität in diesem Sinn meint nicht die Qualität eines bestimmten pädagogischen Konzepts; sie ist jedoch kompatibel mit der in der Früherziehung traditionell vorhandenen „Ansätze“- und „Konzepte“-Pädagogik, insofern als solche holistischen Konzepte als spezielle Konfigurationen im Raster der verschiedenen Qualitätsdimensionen darstellbar werden. Der hier unterlegte Begriff pädagogische Qualität kann mehr oder weniger unmittelbar für Qualitätsverbesserungen fruchtbar gemacht werden. Über Outcomes bei Kindern

validierte Qualitätsdimensionen bilden Stellschrauben für Qualitätsverbesserung und -sicherung.

Vor diesem Hintergrund wurde mit dem Deutschen Kindergarten Gütesiegel ein breit ansetzendes Qualitätsprüf- und -sicherungsverfahren vorgestellt, das nach Auffassung des Autors eine evidenzbasierte Qualitätssicherung nach festen Standards gegenwärtig am ehesten leisten dürfte.

Die positiven Ansprüche und der Nutzen des Verfahrens liegen auf der Hand. Es sichert den Kindern eine nach fachwissenschaftlichen Standards geprüfte Qualität in ihren ersten Einrichtungen öffentlich verantworteter Bildung und Erziehung zu und unterstützt über den Außenausweis die Wahlfähigkeit von Eltern, die eine qualitativ gute Einrichtung für ihr Kind anstreben.

Es liefert Referenzinformationen für die nicht wenigen pädagogischen Fachkräfte in den Einrichtungen und auch bei Trägern, die um pädagogische Qualität bemüht sind, denen aber klare Kriterien fehlen. Und nicht zuletzt dürften die bei Gütesiegeluntersuchungen anfallenden Daten die Qualitätsinformationen bereitstellen, die von den großen Trägerorganisationen wie auch und insbesondere von der Jugendhilfeadministration und -politik für ihre Aufgabe der Qualitätssteuerung benötigt werden. Es gibt gegenwärtig kein *Qualitätsmonitoring* und damit auch keine Möglichkeit für eine evidenzbasierte Qualitätssteuerung im Kindertagesstättenbereich. Steuerungsmöglichkeiten sind aber an belastbare Qualitätsinformationen gebunden. In einem weiteren Schritt sollte auch daran gedacht werden, die öffentliche Finanzierung von Einrichtungen an ein nachgewiesenes Qualitätsniveau zu binden.

Die Forderung nach einem allgemeinen, träger- und bundesländerübergreifenden Qualitätsinformationssystem, wie sie mit dem DKGS verbunden ist, scheint mit der ordnungspolitischen Wirklichkeit der Länderzuständigkeit und dem hohen Grad an Autonomie der freien (und oft auch der kommunalen) Träger in Widerspruch zu stehen. Der vermeintliche Widerspruch wäre jedoch in Fortführung der vorgetragenen Logik auflösbar, indem jedem Bundesland und jeder Trägerorganisation freigestellt wird, einen Zusatz zu einem allgemeinen Gütesiegel zu kreieren, der das jeweilige Proprium überprüft und ggf. zertifiziert. Die interessierten Instanzen können auf diese Weise ihre jeweiligen Orientierungen zum Ausdruck bringen; sie bleiben jedoch über das allgemeine Gütesiegel rückgebunden an das, was als anerkannte gute Fachpraxis auf wissenschaftlicher Grundlage gilt.

## **Literatur**

- Anger, C./Plünnecke, A./Tröger, M. (2007): Renditen der Bildung – Investitionen in den frühkindlichen Bereich. Studie im Auftrag der Wissensfabrik Köln (Inst. d. deutschen Wirtschaft Köln e.V.).
- Bertram, T./Pascal, C./Bokhari, S./Gasper, M./Holtermann, S. (2002): Early Excellence Centre Pilot Programme. Second Evaluation Report 2000-2001. Research Brief No: RR 361, <http://www.dfes.gov.uk/research/data/uploadfiles/RR361.pdf>, 27.11.2007.
- Bos, W./Lankes, E.M./Prenzel, M./Schwippert, K./Walther, G./Valtin, R. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster: Waxmann.

- Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) (Hrsg.) (2003): Auf den Anfang kommt es an! Perspektiven zur Weiterentwicklung des Systems der Tageseinrichtungen für Kinder in Deutschland. Weinheim: Beltz.
- Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) (Hrsg.) (2006): Zwölfter Kinder- und Jugendbericht. Bericht über die Lebenssituation junger Menschen und die Leistungen der Kinder- und Jugendhilfe in Deutschland. Berlin.
- Büchner, C./Spieß, C.K. (2007): Die Dauer vorschulischer Betreuungs- und Bildungserfahrungen. Ergebnisse auf der Basis von Paneldaten. Berlin. <http://www.diw.de/documents/publikationen/73/56550/dp687.pdf>, 31.01.2008.
- Burchinal, M.R. (1999): Childcare experiences and developmental outcomes. In: Helburn, S. (Hrsg.). *Annals of the American Academy of Political and Social Science*, 563, p. 73-97.
- Clarke-Stewart, A./Allhusen, V.D. (2005): *What we know about childcare*. Cambridge: Harvard University Press.
- Cohen, J. (1988): *Statistical Power Analysis for the Behavioral Sciences*. 2. Edition. Hillsdale, New Jersey: Lawrence Erlbaum.
- Cryer, D./Tietze, W./Burchinal, M./Leal, T./Palacios, J. (1999): Predicting Process Quality from Structural Quality in Preschool Programs: A Cross-Country Comparison. In: *Early Childhood Research Quarterly*, 14 (3), S. 339-361.
- Dahlberg, G./Moss, P./Pence, A. (1999): *Beyond Quality in Early Childhood Education and Care. Postmodern Perspectives*. Philadelphia, PA: Falmer.
- Deutscher Bildungsserver (2007): Bundesweite Übersicht Bachelor und Master Studienangebote (Elementar-/Frühpädagogik), <http://www.bildung-in-deutschland.de/zeigen.html?seite=3432>, 27.11.2007.
- Deutscher Bildungsrat (Hrsg.) (1970): *Empfehlungen der Bildungskommission. Strukturplan für das Bildungswesen*. Stuttgart: Klett.
- Deutsches Jugendinstitut – DJI –, München (Hrsg.) (2005): *Kosten und Nutzen eines Gütesiegels für Kindertageseinrichtungen*. München. [http://cgi.dji.de/bibs/Dohmen-Expertise\\_2005\\_Kosten\\_und\\_Nutzen\\_eines\\_Guetesiegels.pdf](http://cgi.dji.de/bibs/Dohmen-Expertise_2005_Kosten_und_Nutzen_eines_Guetesiegels.pdf), 21.10.2008.
- Dippelhofer-Stiehm, B./Kahle, B. (1995): *Die Erzieherin im evangelischen Kindergarten. Empirische Analysen zum professionellen Selbstbild des pädagogischen Personals, zur Sicht der Kirche und zu den Erwartungen der Eltern*. Bielefeld: Kleine.
- Dollase, R. (1978): Kontinuität und Diskontinuität zwischen vorschulischer und schulischer Sozialisation. In: *Bildung und Erziehung*, 31 (5), S. 412-425.
- Dollase, R. (1979): *Sozial-emotionale Erziehung in Kindergarten und Vorklasse*. Hannover: Schroedel.
- Early, D. M./Maxwell, K. L./Burchinal, M. et al. (2007): Teachers' Education, Classroom Quality, and Young Children's Academic Skills: Results From Seven Studies of Preschool Programs. In: *Child Development*, 78 (2), p. 558-580.
- European Child Care and Education (ECCE) – Study Group (1999): *School-age Assessment of Child Development: Long-term impact of Pre-school Experiences on School Success, and Family-School Relationships*. Report written by W. Tietze/J. Hundertmark-Mayser/H.-G. Roßbach. Submitted to: European Union DG XII: Science, Research and Development. RTD Action: Targeted Socio-Economic Research. Brussels. [http://www.uni-bamberg.de/fileadmin/uni/fakultaeten/ppp\\_lehrstuehle/elementarpaedagogik/Team/Rosbach/Ecce\\_Study\\_Group.pdf](http://www.uni-bamberg.de/fileadmin/uni/fakultaeten/ppp_lehrstuehle/elementarpaedagogik/Team/Rosbach/Ecce_Study_Group.pdf), 31.01.2008.
- Esch, K./Klaudy, E.K./Micheel, B./Stöbe-Blossey, S. (2006): *Qualitätskonzepte in der Kindertagesbetreuung. Ein Überblick*. Wiesbaden: Verlag für Sozialwissenschaften.
- Ewert, O.M./Braun, M. (1978): *Ergebnisse und Probleme vorschulischer Förderung. Kultusminister des Landes Nordrhein-Westfalen. Modellversuch Vorklasse in NRW – Abschlußbericht*. Köln: Greven.
- Gerhold, B./Hebenstreit-Müller, S./Kühnel, B. (2001): *Pen Green in Berlin. Die Kita als Nachbarschafts- und Familien-Zentrum*. In: *klein & groß*, Heft 2, S. 18-21.

- Hayes, C.D./Palmer, J.L./Zaslow, M.S. (1990): Who cares for America's children? Child care policy for the 1990s. Washington, D.C.: National Academy Press.
- Jugendministerkonferenz (JMK) (2004): Gemeinsamer Rahmen der Länder für die frühe Bildung in Kindertageseinrichtungen. Beschluss der Jugendministerkonferenz vom 13./14.05.2004. [www.bildungsportal.nrw.de](http://www.bildungsportal.nrw.de), 27.11.2007.
- Katz, L. (1996): Qualität der Früherziehung in Betreuungseinrichtungen: Fünf Perspektiven. In: Tietze, W. (Hrsg.). Früherziehung. Trends, internationale Forschungsergebnisse, Praxisorientierungen. Neuwied: Luchterhand. S. 226-239.
- Ministerium für Generationen, Familie, Frauen und Integration des Landes Nordrhein-Westfalen (MGFFI) (2007): Familienzentren in NRW. <http://www.familienzentrum.nrw.de/>, 27.11.2007.
- NICHD Early Child Care Research Network (Hrsg.) (2005): Child Care and Child Development. New York, London: The Guildford Press.
- Peucker, C./Riedel, B. (2004): Recherchebericht Häuser für Kinder und Familien. München: Deutsches Jugendinstitut.
- Phillips, D.A. (1987): Quality in child care: What does research tell us? Washington DC.: National Association for the Education of Young Children.
- Reidenbach, M. (1996): Kommunale Standards in der Diskussion. Setzung und Abbau von Standards am Beispiel der Kindergärten. Berlin: Deutsches Institut für Urbanistik.
- Robert Bosch Stiftung (2007): PIK – Profis in Kitas. Ein Programm zur Professionalisierung von Fachkräften in Kindertageseinrichtungen. <http://www.bosch-stiftung.de/content/language1/html/994.asp>, 27.11.2007.
- Roßbach, H.-G. (2005): Effekte qualitativ guter Betreuung, Bildung und Erziehung im frühen Kindesalter auf Kinder und ihre Familien. In: Sachverständigenkommission Zwölfter Kinder- und Jugendbericht (Hrsg.). Materialien zum Zwölften Kinder- und Jugendbericht. Band 1. Bildung, Betreuung und Erziehung von Kindern unter sechs Jahren. München: Deutsches Jugendinstitut. S. 55-174.
- Ruopp, R./Travers, J./Glantz, F./Coelen, C. (1979): Children at the center: Final results of the National Day Care Study. Boston: Abt. Associates.
- Spieß, C.K./Tietze, W. (2002): Qualitätssicherung in Kindertageseinrichtungen - Gründe, Anforderungen und Umsetzungsüberlegungen für ein Gütesiegel. In: Zeitschrift für Erziehungswissenschaft, 5 (1), S. 139-162.
- Sylva, K./Siraj-Blatchford, I./Taggart, B. (2006): The Early Childhood Environment Rating Scale – Extension (ECERS-E). Four Curricular Subscales. Stoke-on Trent: Trentham Books.
- Tietze, W. (1987): A Structural Model for the Evaluation of Preschool Effects. In: Early Childhood Research Quarterly, 2 (2), p. 133-154.
- Tietze, W. (Hrsg.) (1998): Wie gut sind unsere Kindergärten? Eine Untersuchung zur pädagogischen Qualität in deutschen Kindergärten. Neuwied: Luchterhand.
- Tietze, W./Förster, C. (2005): Allgemeines pädagogisches Gütesiegel für Kindertageseinrichtungen. In: Diller, A./Leu, H.R./Rauschenbach, T. (Hrsg.). Der Streit um Gütesiegel. Qualitätskonzepte für Kindertageseinrichtungen. München: Deutsches Jugendinstitut. S. 31-66.
- Tietze, W. (2006): Frühpädagogische Evaluations- und Erfassungsinstrumente. In: Fried, L./Roux, S. (Hrsg.). Pädagogik der frühen Kindheit. Weinheim: Beltz. S. 243-253.
- Tietze, W./Viernickel, S. (Hrsg.)/Dittrich, I./Grenner, K./Groot-Wilken, B./Sommerfeld, V./Hanisch, A. (2007): Pädagogische Qualität in Tageseinrichtungen für Kinder. Ein Nationaler Kriterienkatalog. 3. Auflage. Berlin: Cornelsen.
- Tietze, W./Schuster, K.-M./Grenner, K./Roßbach, H.-G. (2007a): Kindergarten-Skala (KES-R). Feststellung und Unterstützung pädagogischer Qualität in Kindergärten. 3. Auflage. Berlin: Cornelsen.
- Tietze, W./Bolz, M./Grenner, K./Schlecht, D./Wellner, B. (2007b): Krippen-Skala (KRIPS-R). Feststellung und Unterstützung pädagogischer Qualität in Krippen. Berlin: Cornelsen.
- Tietze, W./Knobeloch, J./Gerszonowicz, E. (2007c): Tagespflege-Skala (TAS). Feststellung und Unterstützung pädagogischer Qualität in der Kindertagespflege. Berlin: Cornelsen.

- Tietze, W. (Hrsg.) (2008): Qualitätssicherung in der Früherziehung. Internationale Ansätze. Opladen: Budrich.
- Universität Bamberg (2007): Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen im Vor- und Grundschulalter (BiKS). <http://www.uni-bamberg.de/index.php?id=2713>, 27.11.2007.
- Vandell, D.L./Wolfe, B. (2000): Child Care Quality: Does It Matter and Does It Need to Be Improved. Institute for Research on Poverty. Special Report Nr. 78. Madison: University of Wisconsin. <http://www.irp.wisc.edu/publications/specreports.htm>, 27.11.2007.
- Winkelmann, W./Holländer, A./Schmerkotte, H./Schmalohr, E. (1977): Kognitive Entwicklung und Förderung von Kindergarten- und Vorklassenkindern. Kronberg: Scriptor.
- Zollondz, H.D. (2002): Grundlagen Qualitätsmanagement. Einführung in Geschichte, Begriffe, Systeme und Konzepte. München: Oldenbourg.

*Adresse des Autors:*

Prof. Dr. Wolfgang Tietze, Fachbereich Erziehungswissenschaft und Psychologie, Arbeitsbereich Kleinkindpädagogik, Postfach 8, Habelschwerdter Allee 45, 14195 Berlin

*Hartmut Ditton*

## Qualitätssicherung in Schulen

### 1. Einleitung

Unser Schulsystem steht vor einigen Herausforderungen. Indikatoren dafür sind der vergleichsweise späte Zeitpunkt der Einschulungen, häufige Klassenwiederholungen, eine geringe Durchlässigkeit der schulischen Laufbahnen sowie die hohe Zahl der Abgänger ohne schulischen Abschluss. An internationalen Standards gemessen ist das zum Ende der Pflichtschulzeit erreichte Niveau der Kompetenzen in Lesen, Mathematik und den Naturwissenschaften bestenfalls mäßig und der Anteil der Schüler mit auffällig schwachen Kompetenzen ist überdurchschnittlich hoch. Die schulischen Leistungen streuen insgesamt und sowohl zwischen den Ländern als auch zwischen den einzelnen Schulen und Schulklassen einer Schulform innerhalb der Länder sehr erheblich. Die Kopplung der schulischen Leistungen und des Bildungserfolgs an die soziale Herkunft ist eng. Die große Streuung der Leistungen verweist auf Herausforderungen bezüglich der Vergleichbarkeit von Noten und schulischen Abschlüssen auf der einen und bezüglich der Gleichwertigkeit der Angebote und Arbeitsbedingungen an den Schulen auf der anderen Seite. Die Herausforderungen äußern sich auch in den nach wie vor hohen, wenn auch sinkenden Anteilen an Lehrkräften, die wegen Dienstunfähigkeit frühzeitig pensioniert werden. Beschwerden über schlechte räumliche Bedingungen, sanitäre Einrichtungen sowie einen gehäuften Unterrichtsausfall sind an der Tagesordnung. Ein weiteres Dauerthema sind die Klagen über die – zumindest nach Meinung der Lehrkräfte sowie der Eltern- und Lehrerverbände – zu großen Klassen. Im Zuge der Krisendiskussionen der letzten Jahre ist auch die frühe und starke Gliederung des Schulsystems erneut in die Kritik geraten. Obwohl die Systemfrage als Qualitäts- und Schulentwicklungsaspekt gesehen werden muss, wird darauf nachfolgend nicht näher eingegangen, da sich nur berichten lässt, dass einige Länder inzwischen Reformen diskutieren bzw. in Angriff nehmen, wohingegen andere Länder diesbezüglich keinen Handlungsbedarf gegeben sehen.

Das vorherrschend negative, ja geradezu deprimierende Bild, das das deutsche Schulwesen in den Schlagzeilen abgibt, wird zeitweise von Berichten über ausgewählte Fälle besonders positiver schulischer Arbeit unterbrochen. Dabei handelt es sich allerdings keineswegs um solche Ausnahmen wie es auf den ersten Blick erscheinen mag. Vielmehr ist eine große Dynamik der Entwicklungen im schulischen Bereich zu konstatieren. Die Vielzahl der Einzelinitiativen, Wettbewerbe und Modellversuche ist kaum mehr zu überschauen. Nachfolgend kann es daher auch nicht darum gehen, all das, was in den letzten Jahren an Schulentwicklung betrieben und unter das Qualitätsetikett sub-

sumiert wurde, in allen Facetten darzustellen. Im Vordergrund stehen vielmehr die als Hauptlinien erkennbaren Trends der Qualitätssicherung und -entwicklung an Schulen. Natürlich ist es gewagt, in Deutschland etwas für alle Länder gleichermaßen Gültiges aussagen zu wollen. Trotz der Länderhoheit in Bildungsfragen gibt es aber bei näherem Hinschauen überraschend viel an Übereinstimmung in den Entwicklungen zur schulischen Qualitätssicherung über die Ländergrenzen hinweg. Zumindest gilt dies für die nachfolgend im Mittelpunkt des Beitrags stehenden Aspekte.

Institutionell besonders auffällig ist die Einrichtung des von den Ländern gemeinsam getragenen Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) sowie von Qualitätsagenturen in den einzelnen Ländern, die den Auftrag haben, die Bündel der Entwicklungsmaßnahmen im schulischen Bereich zu koordinieren und Schulentwicklungskonzepte zu entwerfen bzw. fortzuschreiben. Damit zusammenhängend hervorzuheben ist die Etablierung von Bildungsstandards durch einen von allen sechzehn Ländern getragenen Beschluss. Mit diesem Beschluss geht zugleich eine Neujustierung des schulischen Lernens einher, weg von einer starren Lehrplanbindung hin zu einer Orientierung an den durch die Schule zu vermittelnden Kompetenzen. Dies wiederum ist in Bezug zu sehen zu den Beschlüssen, auch künftig an den Programmen der internationalen Vergleichsstudien teilzunehmen (PIRLS/IGLU, TIMSS, PISA) und zu den inzwischen etablierten landesweiten sowie länderübergreifend organisierten Leistungserhebungen im Rahmen von sog. Vergleichsarbeiten, Orientierungsarbeiten bzw. Lernstandserhebungen. Durchaus schon eine gewisse Tradition und Verbreitung haben Verfahren der internen sowie der extern unterstützten Evaluation an Schulen. In fast allen Ländern neu hinzugekommen ist in den letzten Jahren eine externe Evaluation im Rahmen von Schulbesuchen bzw. Schulinspektionen oder Schulvisitationen. Die Verfahren der externen Evaluation zeichnen sich zwar – nicht nur vom Namen her – durch länderspezifische Besonderheiten aus, sind aber dennoch von der Konzeption her weitgehend vergleichbar. Schließlich ist auch die Bildungsberichterstattung ein Element der Qualitätssicherung, da hiermit leicht abrufbare Daten, die ansonsten nur verstreut in unterschiedlichen Quellen zu finden sind, übersichtlich und inhaltlich strukturiert zur Verfügung gestellt und Entwicklungen transparent und nachvollziehbar gemacht werden.

Primär wird nachfolgend der Sachstand zur Qualitätssicherung an Schulen dargestellt. Die darüber hinausgehende Frage nach der Wirksamkeit resp. den erreichten Wirkungen der Maßnahmen kann derzeit nicht zuverlässig beantwortet werden. Eine Metaevaluation ist bislang noch kaum in Angriff genommen worden, eher ist selbst noch unklar, ob alle Maßnahmen in der Praxis wie intendiert umgesetzt werden. Bezüglich der Frage nach den nötigen Unterstützungssystemen und Hilfsangeboten zur Qualitätssicherung sind ebenfalls keine belastbaren Daten vorhanden. Verständlich ist diese unbefriedigende Datenlage deshalb, weil sich Vieles noch im Aufbau bzw. gerade erst in einer Konsolidierungsphase befindet.

## 2. Ansätze der Qualitätssicherung in Schulen

Die Maßnahmen zur Qualitätsentwicklung und -sicherung an den Schulen in Deutschland sind sehr vielfältig, angefangen von Initiativen einzelner Schulen in Eigenregie, über Modellversuche und Wettbewerbe bis hin zu landesweit und länderübergreifend vereinbarten Vorhaben. Üblicherweise mit zur Qualitätsinitiative gerechnet werden auch Maßnahmen, die auf eine erweiterte Selbständigkeit von Schule abzielen, und häufig im Zusammenhang damit stehen, dass in Abstimmung der Schulen mit der Schulaufsicht ein Schulprogramm vereinbart wird. Um die vielfältigen Qualitätsinitiativen zu koordinieren und eine institutionelle Basis für die weitere Entwicklung zu schaffen, wurden inzwischen in den Ländern Qualitätsagenturen eingerichtet. Die Qualitätsagenturen sind entweder eigenständige Einrichtungen, die den Qualitätsbegriff sozusagen als Signal im Namen tragen, oder es handelt sich um Abteilungen an übergeordneten Instituten oder Ministerien (vgl. Tabelle 1 im Anhang). Ein strittiger Punkt bei der Einrichtung der Agenturen war und ist, wieweit diese völlig selbständig und unabhängig von der Schulaufsicht sein dürfen oder müssen, um ihre Aufgaben wahrnehmen zu können. Wieweit ein solcher Status zugestanden wird, ist im Vergleich der Länder unterschiedlich.<sup>1</sup>

Als inhaltliche Grundlage der Arbeit in den Agenturen wurde in aller Regel ein mehr oder weniger differenzierter und konkretisierter Qualitätsrahmen erarbeitet. Überwiegend folgt dieser der Systematik von Input, Prozess und Output bzw. Outcome (Ditton 2000, 2002). Insofern besteht auch eine gewisse Übereinstimmung im Qualitätsverständnis und bezüglich der Aufgaben von Evaluationen im schulischen Bereich. Ob allerdings unter gleichen Begriffen auch jeweils das Gleiche gemeint ist, wäre erst noch genauer zu untersuchen.

Eine Institutionalisierung auf länderübergreifender Ebene ist mit der Einrichtung des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) im Dezember 2003 erfolgt. Das IQB wird nach dem Königsteiner Schlüssel durch die Bundesländer finanziert und arbeitet in enger Abstimmung mit den Ländern sowie mit nationalen und internationalen Forschungseinrichtungen, Verbänden und Institutionen. Eine Kernaufgabe des IQB besteht darin, die Bildungsstandards, die von den Ländern zum Schuljahresbeginn 2004/5 bzw. 2005/6 verbindlich eingeführt wurden, weiterzuentwickeln, sie zu normieren, ihre Erreichung zu überprüfen und ihre Implementation wissenschaftlich zu begleiten. Zu den Arbeitsfeldern des IQB gehören vor allem die

- Generierung von Aufgabensammlungen zur Operationalisierung der Standards,
- Formulierung von Kompetenzmodellen, in denen für die einzelne Fächer beschrieben wird, welche Kompetenzen die Schüler zu einem definierten Zeitpunkt erreicht haben sollen,
- Formulierung von Vergleichsaufgaben, die geeignet sind, das Erreichen der Kompetenzen durch die Schüler zu erfassen.

1 Um den Status genauer zu klären, müssten die Geschäftsordnungen der Agenturen analysiert werden.

Zentrale Arbeitsbereiche des IQB sind daher die Testentwicklung und Untersuchungen zu Kompetenzerwerb und Kompetenzentwicklung. Außerdem befasst sich das IQB mit Fragen der Evaluation und Implementation im Schulsystem sowie der Wirksamkeit von Implementationsmaßnahmen der Qualitätsentwicklung in Schulen. Längerfristig sollen im Sinne eines Bildungsmonitoring Indikatoren der Bildungsqualität gewonnen werden, die Steuerungswissen für die Bildungspolitik generieren und eine Grundlage für Reformen im Bildungssystem abgeben können. Von besonderer Bedeutung sind hierbei wiederum Schulleistungsindikatoren, die in Abstimmung mit den internationalen Erhebungsprogrammen definiert werden sollen. Schließlich hat das IQB ein Forschungszentrum etabliert, in dem die Datensätze aus den großen nationalen und internationalen Schulleistungsstudien archiviert, dokumentiert und für Re- und Sekundäranalysen zur Verfügung gestellt werden. Die Finanzierung dieses Projektes erfolgt durch das Bundesministerium für Bildung und Forschung (BMBF). Eine in Zukunft zunehmend wichtiger werdende Aufgabe wird die Erarbeitung von computergestützten Test-, Auswertungs- und Rückmeldesystemen sein, um ein zeitnahes, ökonomisch einsetzbares, flexibel anpassbares und zukunftsfähiges Qualitätssicherungssystem etablieren zu können. Mit der Einrichtung des IQB wurde sozusagen eine länderübergreifend koordinierende Schaltstelle geschaffen, die für die gemeinsamen Programme in den Ländern zuständig ist. Daneben bestehen die Agenturen der Länder und zahlreiche länderspezifische Einzelmaßnahmen. Nachfolgend geht es darum, Gemeinsamkeiten herauszustellen, auf landesspezifische Besonderheiten kann nur am Rande eingegangen werden.

## 2.1 Kompetenzorientierung und Bildungsstandards

Beginnend mit PISA hat sich in gewisser Hinsicht eine Neujustierung des schulischen Lernens weg von der Lehrplan- hin zu einer Kompetenzorientierung vollzogen. Der Versuch, schulisches Lernen durch eine Vielzahl inhaltlich detaillierter Vorgaben zu steuern, soll durch eine Zielorientierung hinsichtlich der angestrebten Kompetenzen ersetzt werden. Im Vordergrund stehen soll nunmehr die Konzentration auf Kernbereiche der Fächer sowie der Erwerb eines vertieften und anwendungsbezogenen Wissens. Mit der Kompetenzorientierung einher geht die Etablierung von zwischen den Ländern gemeinsam vereinbarten Bildungsstandards, in denen die Kompetenzen und Kompetenzanforderungen konkretisiert werden.

Nach Weinert (1999) können unter Kompetenzen im engeren Sinn die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten verstanden werden, bestimmte Probleme zu lösen. Der Kompetenzbegriff im weiteren Sinn beinhaltet darüber hinaus die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können. Kompetenz ist insofern ein Komplex aus Fähigkeit, Wissen und Verstehen, Können und Handeln sowie Erfahrung und Motivation. Eng an dieses Kompetenzverständnis lehnen sich die Ausführungen des Forum Bildung an (Arbeitsstab Forum Bildung 2001). In den Empfehlungen wird

ausgeführt, dass in einer auf Pluralität und ständigem Wandel gegründeten Gesellschaft der Weg zur Realisierung eines umfassenden Bildungsbegriffs nur über den Erwerb von Kompetenzen führen kann, die den Einzelnen zur Orientierung und zum produktiven Umgang mit Pluralität und Wandel befähigen. Zu diesen Kompetenzen zählen vor allem Lernkompetenz (Lernen des Lernens), die Verknüpfung von intelligentem inhaltlichen Wissen mit der Fähigkeit zu dessen Anwendung, methodisch-instrumentelle (Schlüssel-)Kompetenzen, insbesondere in den Bereichen Sprachen, Medien und Naturwissenschaften sowie soziale Kompetenzen und Wertorientierungen.

Auf eine solche Orientierung des schulischen Lernens an grundlegenden Kompetenzen zielen die Vereinbarungen der KMK zu Bildungsstandards ab. Eine im Vorfeld erstellte Expertise (Klieme 2004; Klieme u.a. 2003) führt dazu aus, dass Bildungsstandards Ziele des schulischen Lernens in konkrete Anforderungen umsetzen und festlegen, über welche Kompetenzen die Schüler und Schülerinnen verfügen sollen, damit die Ziele als erreicht gelten können. Systematisch geordnet werden die Anforderungen in Kompetenzmodellen, die Aspekte, Abstufungen und Entwicklungsverläufe von Kompetenzen darstellen. Bildungsstandards werden in Aufgabenstellungen konkretisiert und münden schließlich in Verfahren, mit denen die Kompetenzniveaus empirisch zuverlässig erfasst werden können. Diesbezüglich wurden erstmals in der Geschichte der Bundesrepublik Deutschland am 4. Dezember 2003 gemeinsame Bildungsstandards in den Fächern Mathematik, Deutsch und Erste Fremdsprache (Englisch, Französisch) für den Mittleren Bildungsabschluss beschlossen. Am 15. Oktober 2004 folgte die Verabschiedung von Bildungsstandards für den Hauptschulabschluss in Mathematik, Deutsch und Erster Fremdsprache und für den Primarbereich in Deutsch und Mathematik. Mit Beschluss vom 16. Dezember 2004 wurden für den Mittleren Abschluss Bildungsstandards in Biologie, Physik und Chemie vereinbart. Die Bildungsstandards wurden von den Ländern zum Schuljahresbeginn 2004/5 bzw. 2005/6 verbindlich eingeführt, was zunächst einmal heißt, dass die Standards als Leitlinien für die weitere Entwicklung in den Länder gelten. Die tatsächliche unterrichtsrelevante Umsetzung wird sicherlich noch einige Zeit in Anspruch nehmen.

Im Zusammenhang mit der Vereinbarung der Bildungsstandards wurde beschlossen, dass eine systematische und regelmäßige Überprüfung ihrer Einhaltung erfolgen soll. Mit dieser Aufgabe ist das IQB beauftragt, das derzeit Verfahren zur Messung der Kompetenzbereiche erarbeitet. Dazu wurde in 2004/2005 für den Sekundar- und in 2005/2006 für den Primarbereich mit der Entwicklung einer großen Zahl an Aufgaben begonnen. Dem folgten im April bis Juni 2005 bzw. 2006 Pilotierungen im Rahmen der Feldtests zu PISA bzw. IGLU. Die Normierung, Berichterstattung und Bereitstellung der nationalen Skalen erfolgte im Frühjahr 2008 für den Sekundarbereich und im Sommer 2008 für den Primarbereich.

Die Bildungsstandards sind nicht ausschließlich als Prüfinstrument zu verstehen, vielmehr sollen sie die Arbeit im Unterricht und an den Schulen anleiten. Damit sie Wirkung entfalten, ist die Akzeptanz und Beteiligung seitens der Schulen nötig. So erfolgt auch bereits die Operationalisierung, d.h. die Erstellung der Aufgaben zur Überprüfung der Standards in einem aufwendigen Verfahren unter Einbeziehung vielfältiger

Expertise. Beim IQB (und auch bei dem Vorgehen der Länder in der sog. VERA-Gruppe, s. unten) erfolgt die Aufgabenentwicklung durch Arbeitsgruppen, denen Lehrkräfte, Fachdidaktiker und Psychometriker angehören. In den Gruppen werden Aufgabenvorschläge entwickelt, die dann bezüglich ihrer Anwendbarkeit diskutiert und empirisch geprüft werden, bevor eine Normierung erfolgen kann.

Für die Erhebungen zu den Bildungsstandards durch das IQB werden Stichproben in den Ländern gezogen und die Schülerleistungen unter Testbedingungen mit externer Administration der Aufgaben ermittelt. Daneben besteht aber auch ein Bedarf an flächendeckenden Erhebungen, damit alle Schulen und Lehrkräfte sich über die Leistungsstände an der jeweiligen Schule und in der eigenen Klasse informieren können. Auf diese Aufgabe konzentrieren sich die Programme der Länder, die unter den Bezeichnungen Orientierungsarbeiten, Vergleichsarbeiten, Kompetenztests oder Lernstandserhebungen durchgeführt werden. Da in der Fläche eine externe Testadministration nicht möglich ist, liegt die Durchführung und Auswertung vielmehr in der Hand der Lehrkräfte. Bislang war sehr unübersichtlich, welches Land bzw. welche Ländergruppe in welcher Jahrgangsstufe und welchem Fach Leistungserhebungen durchführt. Eine Übersicht für den Bereich der allgemein bildenden Schulen der Sekundarstufe gibt Tabelle 3 im Anhang. Am häufigsten werden demnach Vergleichsarbeiten bislang in den Fächern Deutsch und Mathematik durchgeführt, daneben noch in Englisch. Vergleichsarbeiten in anderen Fächern (Französisch, Latein) sind dagegen selten. Die Vergleichsarbeiten konzentrieren sich auf die sechste bis achte Jahrgangsstufe, zudem wird in einigen Ländern die zehnte Jahrgangsstufe der Gymnasien einbezogen.

Für den Bereich der Primarstufe werden ab 2008 in einem gemeinsamen Programm aller sechzehn Ländern Vergleichsarbeiten zum Ende der dritten Jahrgangsstufe durchgeführt. Es handelt sich um die Nachfolge des „VERA-Projektes“, das von der Universität Koblenz-Landau durchgeführt und betreut wird (Helmke, Hosenfeld). Der Vorläufer dieser Länderkooperation war ein Projekt von sieben Ländern (Berlin, Brandenburg, Bremen, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Rheinland-Pfalz und Schleswig-Holstein), in dem der Erhebungszeitpunkt der Beginn der vierten Jahrgangsstufe war. Daneben hatten weitere Länder (Baden-Württemberg, Bayern, Berlin, Brandenburg, Sachsen, Thüringen) Orientierungsarbeiten in eigener Regie bzw. auf der Basis von Ländervereinbarungen durchgeführt, wobei zum Teil auch die zweiten Jahrgangsstufen einbezogen waren. Wieweit diese Arbeiten, die eine möglichst frühe Diagnose erlauben sollten, beibehalten werden, wird sich definitiv wohl erst in den nächsten Jahren zeigen. Ebenso ist denkbar, dass die Bemühungen in den Ländern um eine differenzierte Schuleingangsdiagnostik in Zukunft ebenfalls zu Kooperationen führen werden. Für den Bereich der Sekundarstufe ist dies bereits in Vorbereitung (Lernstandserhebungen bzw. Vergleichsarbeiten in Jahrgangsstufe 8), wobei davon auszugehen ist, dass das IQB hierbei eine federführende Rolle einnehmen wird.

Die Verbindung zwischen den Erhebungen des IQB und den Länderprogrammen ist dadurch gegeben, dass sich die Aufgaben in beiden Fällen auf die Bildungsstandards beziehen. Das IQB kann allerdings mit einem Multi-Matrix-Design, d.h. einer Rotation einer großen Zahl an Aufgaben in verschiedenen Testheften, die Inhaltsbereiche der Bil-

dungsstandards in der gesamten Breite abdecken. Bei den regelmäßigen (i.d.R. jährlichen) flächendeckenden Erhebungen in den Ländern ist das naturgemäß nicht realisierbar, vielmehr muss man sich hier, schon aus Kostengründen und wegen der beanspruchten Testdauer, auf eine Auswahl einzelner Inhaltsbereiche konzentrieren. Die Vergleichsarbeiten der Länder beanspruchen auch nicht den Status umfassender Monitoring-Studien, sondern sollen in erster Linie Impulse für die Unterrichtsarbeit geben. Durch die Vergleichsarbeiten erhalten die Schulen und einzelnen Lehrkräfte Aufgaben, in denen die Anforderungen in den Kompetenzbereichen auf den unterschiedlichen Niveaustufen zum Ausdruck kommen. Zudem erhalten sie Ergebnismrückmeldungen, wie die Schule und die einzelnen Schulklassen in der Schule im Vergleich zur Gesamtheit aller Schulen abgeschnitten haben. Ein Beispiel für eine solche Ergebnismrückmeldung aus dem bisherigen Projekt zu den Vergleichsarbeiten an Grundschulen (VERA) zeigt Abbildung 1. Um keine „unfairen“ Vergleiche anzustellen, erhalten die Schulen in aller Regel zusätzlich adjustierte Vergleichswerte, d.h. Darstellungen der Ergebnisse in der Schule bzw. Schulklasse im Vergleich zu Schulen bzw. Schulklassen, bei denen von ihren Arbeitsbedingungen her (Schülerzusammensetzung, soziales Umfeld) ähnliche Voraussetzungen vorliegen.

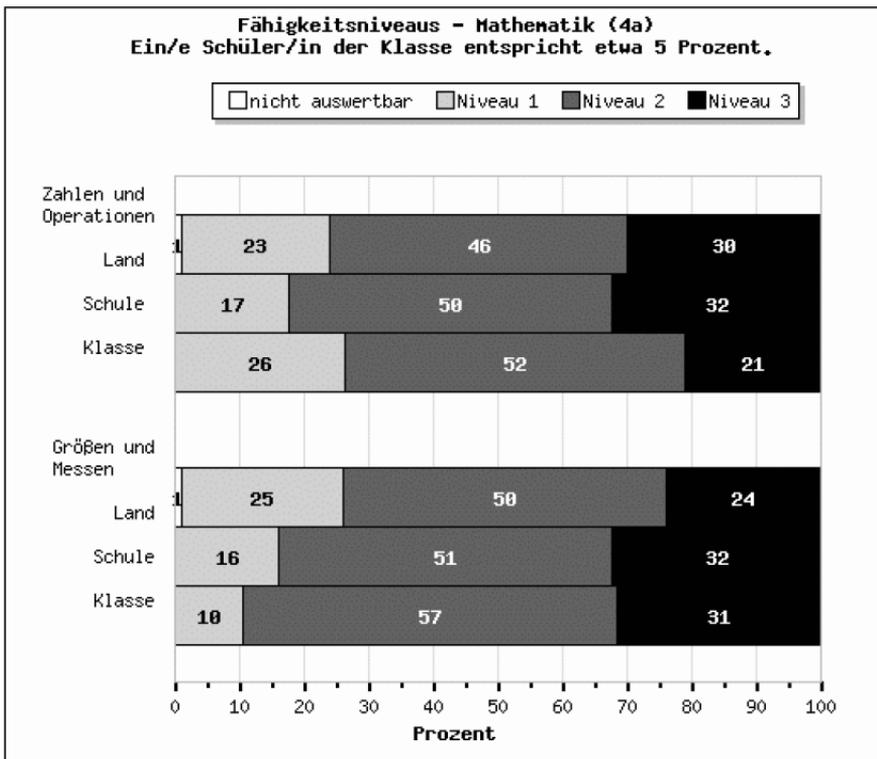
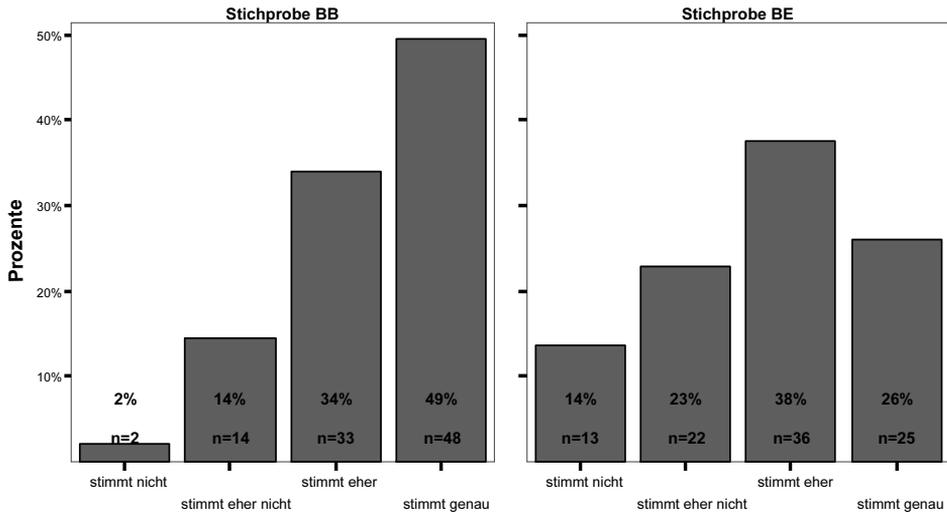


Abb. 1: Beispiel einer Rückmeldung von Leistungsergebnissen aus VERA  
(<http://www.uni-landau.de/vera/>)

Die Vergleichsarbeiten in der Grundschule und in den weiterführenden Schulen verfolgen somit eine mehrfache Zielsetzung. Sie geben den Lehrkräften einerseits Materialien und Aufgaben zur Konkretisierung der in den Bildungsstandards definierten Kompetenzbereiche und Anforderungsniveaus an die Hand und sie informieren andererseits über die erreichten Lernergebnisse im Hinblick auf die Standards. Für die konkrete Arbeit im Unterricht stehen zu den Vergleichsarbeiten didaktische Erläuterungen und Handreichungen zur pädagogischen Nutzung zur Verfügung (Blum u.a. 2006). Diese Materialien können zusammen mit den Ergebnisberichten ein Ausgangspunkt für eine differenzierte und kritische Auseinandersetzung mit den Bildungsstandards und der eigenen pädagogischen Arbeit sein. Dies impliziert bereits, dass niemand so naiv ist zu glauben, dass das Überprüfen und Testen von Leistungen an sich schon zu Verbesserungen oder überhaupt zu Reaktionen führt. Zentral ist vielmehr, wie im Vorfeld, bei der Durchführung und im Anschluss an die Erhebungen an den Schulen mit den Verfahren umgegangen und wie darauf reagiert wird. Wichtig ist auch, wie hoch die Akzeptanz der Verfahren ist und wie belastend und hilfreich sie erlebt werden. Dazu gibt es bisher erst wenige Untersuchungen und noch erheblichen Forschungsbedarf (Ditton/Arnoldt 2004; Groß Ophoff u.a. 2006; Helmke 2004; Kuper/Schneewind 2006; Rolff 2002; Schrader/Helmke 2003). Nach Ergebnissen einer Befragung zur Rezeption der Rückmeldungen im Rahmen der Orientierungsarbeiten an den Grundschulen in Berlin und Brandenburg kann zusammenfassend festgehalten werden (Ditton/Krüsken 2005), dass der Aufwand für eher hoch gehalten wird. Die Wirksamkeit für die Arbeit an der Schule wird von den Lehrkräften als vergleichsweise gering angesehen. Dagegen findet sich mehrheitlich die Einschätzung, dass die Orientierungsarbeiten wichtige Informationen über die eigene Schulklasse geben (Abbildung 2) und dass sie für den Unterricht gut nutzbar sind. Sehr eindeutig abgelehnt wird eine Veröffentlichung der Ergebnisse einzelner Schulen oder Schulklassen.

Natürlich finden sich zu Bildungsstandards und Vergleichsarbeiten auch kritische Stellungnahmen. Diskutiert werden insbesondere folgende Fragen: Bilden die Standards die Fachinhalte angemessen ab oder sind sie womöglich zu eng bzw. im Gegenteil zu weit gefasst? Handelt es sich überhaupt um die intendierten „schlanken“ Vorgaben zu Kernbereichen oder sind die Standards lediglich umetikettierte Lehrplankataloge? Wie viele Kompetenzstufen sollten (und können empirisch bestätigt) unterschieden werden? Genügen Standards für den Abschluss von Bildungslaufbahnen oder sollten sie auch für signifikante Zwischenetappen definiert werden? Werden dadurch, dass Bildungsstandards (bisher) nur für die „Kernfächer“ vorliegen, die anderen Fächer an den Rand gedrängt und implizit als weniger bis unbedeutend eingestuft? Genießen in der Wahrnehmung der Arbeit einer Schule die nicht „standardbezogenen“ schulischen Aktivitäten (z.B. besondere Angebote im sportlich, musisch, künstlerischen Bereich) kein besonderes Ansehen mehr? Kann eine zu eng verstandene Orientierung an Standards zu einem „Schubladendenken“ verführen und fachübergreifendes Arbeiten sowie die Vermittlung fachübergreifender Kompetenzen und Wertorientierungen in den Hintergrund treten lassen?



**Der Leseverständnistest gibt mir wichtige Informationen über meine Klasse.**

Abb. 2: *Einschätzung der Lehrkräfte zu den Orientierungsarbeiten in Brandenburg (BB) und Berlin (BE) (Ditton/Krüsken 2005)*

Kompetenzorientierung, Bildungsstandards und systematische Erhebungen zu ihrer Überprüfung sind ein wichtiger Baustein der Qualitätssicherung, weil sie Kernanforderungen an schulischen Lernerfolg transparent und leichter kommunizierbar machen und weil sie verdeutlichen, worauf ganz besondere Aufmerksamkeit zu legen ist. Das impliziert keineswegs, dass alles andere vernachlässigt werden darf. International ist eine Ausrichtung an Standards eher die Regel als die Ausnahme. Bildungsstandards und daran orientierte Leistungserhebungen erlauben allerdings noch keine Rückschlüsse auf die Ursachen größeren oder geringeren Lernerfolgs. Dazu müss(t)en Beziehungen der Ergebnisse zu den Bedingungen und Prozessen in Schule und Unterricht hergestellt werden. Standards und Vergleichsarbeiten sind insofern nur ein, wenn auch wesentliches Element in einem Gesamtsystem der Qualitätssicherung an Schulen.

## 2.2 Interne und externe Evaluation

Neben den Bildungsstandards und den Vergleichsarbeiten haben Verfahren der internen und externen Evaluation eine wichtige Bedeutung für die Qualitätssicherung und Schulentwicklung. Traditionell werden im schulischen Bereich unterschiedliche Evaluationsansätze verfolgt, u.a. besteht auch die Möglichkeit, Angebote zur extern unterstützten Evaluation zu nutzen (z.B. Bertelsmann: „SEIS“; Qualitätsagentur in Bayern: „bilanz ziehen“). Alles in allem ist davon auszugehen, dass an den Schulen eine eher bunte Evaluationskultur besteht.

Empirische Ergebnisse zur internen Evaluation an Schulen liegen aus einer Befragung der Qualitätsagentur in Bayern vor (Staatsinstitut für Schulqualität und Bildungsforschung – Qualitätsagentur 2006a). An der schriftlichen Befragung haben sich N=3.110 von 5.323 Schulen beteiligt (Rücklauf: 58%). Von den Schulen geben 21% an (N=568), regelmäßig eine interne Evaluation durchzuführen. Etwa die Hälfte hat bisher zweimal evaluiert, die anderen Schulen häufiger. Ebenfalls etwa die Hälfte der Schulen gibt an, jährlich zu evaluieren. Fragebogen scheinen am häufigsten und nahezu von allen Schulen, die evaluieren, verwendet zu werden. Dabei nutzen 55% der Schulen einen Fragebogen, den sie selbst entwickelt haben, 20% verwenden das Verfahren der Qualitätsagentur („bilanz ziehen“). Unterrichtsbeobachtungen scheinen seltener zu sein (N=200 Antworten), noch seltener Interviews (40 Antworten). Als Modell, an dem sie sich orientieren, nennen 64% der Schulen ein selbst entwickeltes Konzept, EFQM und DIN/ISO werden von 5% genannt, das Konzept der bayerischen Qualitätsagentur von 14% der Schulen.

In allen Ländern ist inzwischen auch eine externe Evaluation der Schulen vorgesehen (Tabelle 2 im Anhang). Die Verfahren, die sich zum Teil noch in der Erprobungsphase, zum Teil aber auch schon in der Anwendung befinden, firmieren unter den Bezeichnungen Schulbesuch, Schulvisitation, Schulinspektion oder externe Evaluation. Unabhängig von den unterschiedlichen Benennungen zeichnet sich auch hier sehr viel an Gemeinsamkeit zwischen den Ländern bezüglich der Vorgehensweisen ab. Das grundlegende Muster besteht darin, dass externe Experten im Auftrag der Schulaufsicht die einzelnen Schulen besuchen, sich über deren Arbeit informieren und als Ergebnis der Evaluation eine Einschätzung zu den Stärken und Schwächen der Schulen abgeben. Auf Grundlage der Evaluation werden Vereinbarungen zu Entwicklungszielen getroffen, die dann in späteren Evaluationen auf die Zielerreichung bzw. die zwischenzeitlichen Aktivitäten hin überprüft werden sollen. Unverkennbar lehnt man sich damit an die Schulinspektionen in anderen Ländern (Großbritannien, Niederlande) an, in denen dieses Vorgehen eine längere Tradition hat (kritisch dazu z.B. Hopkins/West/Skinner 1995).

Der Ablauf einer externen Evaluation ist im Vergleich der Länder zwar nicht völlig einheitlich, dennoch zeigen sich aber starke Übereinstimmungen. Eine Evaluation beansprucht in etwa einen Zeitraum von 10 Wochen bis 4 Monaten. Üblicherweise werden die zu evaluierenden Schulen per Zufall gezogen, in den Phasen zur Erprobung der externen Evaluation konnten sich demgegenüber Schulen freiwillig zur Teilnahme melden. Die ausgewählten Schulen werden in einem ersten Schritt über die anstehende Evaluation informiert und stellen zur Vorbereitung des Schulbesuchs geeignete Unterlagen (z.B. Schulprogramm, Jahresbericht) zur Verfügung. Auf der Basis dieser Unterlagen und weiterer Basisdaten (z.B. Schülerzahlen usw.), die von der Schulaufsicht zur Verfügung gestellt werden, können sich die Evaluations-Teams schon vor dem Schulbesuch ein erstes Bild der Schule verschaffen. Zusätzlich können den Evaluations-Teams auch Ergebnisse aus den Vergleichsarbeiten und ggf. vorliegende Ergebnisse aus internen Evaluationen bzw. im Vorfeld durchgeführten Befragungen (der Lehrkräfte, Eltern, Schüler) an der Schule zur Verfügung stehen. Die Evaluations-Teams bestehen aus 2-8 Personen, in den meisten Fällen sind es drei Personen. Mitglieder in den Teams können

Lehrkräfte, Schulleiter, Vertreter der Schulaufsicht, der Eltern und externe Experten (z.B. Vertreter der Wirtschaft, von Verbänden) sein. Der Schulbesuch im Rahmen der externen Evaluation vor Ort dauert zwischen zwei bis vier Tagen und beinhaltet eine Begehung der Räumlichkeiten, Gespräche bzw. Interviews mit der Schulleitung, mit Vertretern der Lehrkräfte sowie den Schülern und Elternvertreter. Üblicherweise werden auch Unterrichtsbesuche durchgeführt. Diese können angekündigt oder unangekündigt sein und sich entweder auf einzelne (ganze) Unterrichtsstunden oder stichprobenartige Besuche von mehreren Stunden (Teile der Stunden) beziehen. Den Abschluss der Evaluation bildet ein Gespräch des Evaluations-Teams mit der Schulleitung und dem Kollegium. In diesem Gespräch werden erste Eindrücke mitgeteilt und die Schule erhält die Gelegenheit, Informationen zu ergänzen. Das Evaluationsteam erstellt auf der Basis der vorab eingereichten Unterlagen und der Ergebnisse der Evaluation vor Ort einen Bericht, der in aller Regel zunächst einmal der Schule zugeht. Die Schule hat Gelegenheit, zu diesem Bericht Stellung zu nehmen. Der Bericht und die Stellungnahme (bzw. ein revidierter und ergänzter Bericht) geht dann der Schulaufsicht und dem Schulträger zu. Auf der Basis des Berichts und der Empfehlungen trifft die Schulaufsicht eine Zielvereinbarung mit der Schule, die daraufhin einen Aktionsplan entwickelt, in dem die Empfehlungen aufgegriffen werden. Bei einer wiederholten Evaluation sollen die Ergebnisse der vorangegangenen Evaluation, die Zielvereinbarungen und Aktionspläne die Grundlage bilden. Hervorzuheben ist noch, dass sich die externe Evaluation nicht auf die Tätigkeit einzelner Personen bezieht, sondern die Schule als Ganzes in den Blick nimmt. Hierbei gibt der Qualitätsrahmen des jeweiligen Landes die Orientierung dafür, worauf die Evaluationsteams bezüglich der Rahmenbedingungen, der schulischen und unterrichtlichen Prozesse und Wirksamkeit einer Schule besonderes Augenmerk zu legen haben. Eine Übersicht zur externen Evaluation in den Ländern gibt Tabelle 3 im Anhang.

Die Mitglieder in den Evaluationsteams arbeiten zum Teil ehrenamtlich, teils sind auch Mitarbeiter der Qualitätsagenturen in den Teams vertreten. Alle Evaluatoren werden auf ihre Tätigkeit vorbereitet und nehmen an Kursen zu Themen der Evaluation sowie zu Fragen der Schul- und Unterrichtsqualität bzw. -entwicklung teil. Sie werden ebenfalls mit dem Qualitätsrahmen und den Aspekten, die bei der Evaluation im Vordergrund stehen sollen, vertraut gemacht. Die Vorbereitung beinhaltet auch die Offenlegung der inhaltlichen und methodischen Standards für die Evaluation, um sicher zu stellen, dass sich die Teams auf vergleichbare Maßstäbe beziehen. Dies ist nicht leicht zu gewährleisten, da für die Umsetzung in der Fläche bei einer großen Zahl der zu evaluierenden Schulen zahlreiche Evaluations-Teams benötigt werden. In der Regel ist vorgesehen, dass alle Schulen eines Landes in einem zeitlichen Rahmen von ca. drei bis fünf Jahren einmal evaluiert werden. Wieweit dies praktikabel ist, wird sich noch zeigen müssen.

Zu den Wirkungen der externen Evaluation gibt es bislang kaum empirische Daten. Eine Befragung der Schulleitungen aller bisher evaluierten Schulen in Bayern, die Aspekte des Ablaufs, der Akzeptanz und der seitens der Schulen wahrgenommenen Wirkungen beinhaltet, hat die bayerische Qualitätsagentur durchgeführt (Staatsinstitut für Schulqualität und Bildungsforschung – Qualitätsagentur 2006b). Die Ergebnisse der

Studie beziehen sich auf 311 verwertbare Fragebögen von der Gesamtzahl der 320 bereits an der externen Evaluation beteiligten Schulen. Zum Zeitpunkt der Befragung befanden sich die Schulen in unterschiedlichen Phasen nach Abschluss der Evaluation. Bei 61% der Schulen waren seit der Evaluation zwischen drei und zehn Monate vergangen, bei 14% war der Zeitraum kürzer und bei den restlichen Schulen länger. Hinsichtlich der derzeitigen Arbeitsphase ordneten sich 20% in die Phase „Abschluss der Evaluation“, 30% in „Zielfindung / Zielvereinbarung“ und 32% in „Umsetzung von Maßnahmen“ ein. Weitere 4,8% gaben an, sich bereits in der Phase der Überprüfung des Erfolgs von Maßnahmen zu befinden.

Nach Angaben der Schulleitungen wurden die Ergebnisse der Evaluation in 80% der Fälle im Rahmen einer Lehrerkonferenz behandelt, in 40% der Fälle (zudem) in einer gemeinsamen Konferenz von Lehrern, Eltern und Schüler. In ebenfalls 40% der Fälle wurde für Eltern und Schüler eine gesonderte Informationsveranstaltung angeboten. Zielvereinbarungen auf der Grundlage der Evaluation wurden im Durchschnitt nach ca. sieben Monaten getroffen, für die Umsetzung gezielter Entwicklungsmaßnahmen ist im Mittel ein Zeitraum von ca. 13 Monaten anzusetzen. Bezüglich des zeitlichen Ablaufs ist die Streuung allerdings erheblich: Während der größere Teil der Schulen eher zügig an die Planung und Umsetzung von Maßnahmen heranzugehen scheint, wirkt das Vorgehen des übrigen Teils der Schulen dagegen als recht zögerlich. Die Auseinandersetzung mit den Ergebnissen der Evaluation erfolgt in erster Linie in Arbeitsgruppen innerhalb des Kollegiums (53%), bei einem pädagogischen Tag (51%), im Rahmen von Fachschaftssitzungen (34%), zum Teil aber auch in Arbeitsgruppen vom Lehrern, Eltern und Schülern (19%). Der Prozess der Zielfindung geschieht hauptsächlich unter Beteiligung der Schulleitung und Lehrkräfte, in 36% der Fälle werden die Eltern und in 25% der Fälle die Schüler einbezogen. Zielvereinbarungen werden im Regelfall zwischen der Schule und Schulaufsicht getroffen, nur in Ausnahmefällen werden Vorschläge der Schule oder Schulaufsicht einfach übernommen. Die vereinbarten Maßnahmen richten sich hauptsächlich auf die Prozessqualitäten auf der Ebene der Schule. Maßnahmen zu Rahmenbedingungen oder zum Unterricht sind demgegenüber seltener. Eher eine Ausnahme sind Vereinbarungen bezüglich der Ergebnissicherung und des Monitoring an der Schule. Eine Beratung und Unterstützung im Anschluss an die Evaluation erfolgt in erster Linie durch Mitglieder der Schulaufsicht (41,7%), andere Formen (z.B. Beratung durch Kollegen anderer Schulen) sind nicht sehr häufig (ca. 11% der Fälle).

Die Gesamtbeurteilung der externen Evaluation durch die Schulleitungen fällt in der Studie positiv aus. Die weit überwiegende Mehrheit (80%) bejaht, dass die Evaluation Anstöße für die Schulentwicklung gegeben hat. Immerhin 36% der Schulleitungen haben sich allerdings mehr von der Evaluation erhofft. Eine Profileinschätzung zeigt, dass die externe Evaluation eher als „hilfreich, anregend, motivierend“ und deutlich weniger als „enttäuschend, störend, belastend, ärgerlich“ erlebt wurde. Gemessen an der im Vorfeld verbreiteten Skepsis gegenüber externen Evaluationen verweisen die Ergebnisse eher auf positive Erfahrungen und darauf, dass Bedenken oftmals zerstreut werden konnten. Dennoch zeigt sich auch in den Bewertungen eine fast ähnlich große Streuung wie in den Reaktionsmustern und zeitlichen Abläufen an den Schulen. Dies deutet da-

rauf hin, dass es vermutlich nicht unerheblich von der jeweiligen Evaluation, bzw. dem Evaluations-Team und der Schule selbst abhängt, wieweit die Evaluation auf Akzeptanz stößt und zu hilfreichen Entwicklungsimpulsen führt.

### 2.3 *Bildungsberichterstattung*

Zu den Maßnahmen der schulischen Qualitätssicherung ist auch die Darstellung und Offenlegung des Sachstandes und der Entwicklungen im Bildungswesen zu zählen. Auf nationaler Ebene ist die Bildungsberichterstattung in Deutschland inzwischen ein etabliertes Programm. Der erste Bildungsbericht wurde im Auftrag der Ständigen Konferenz der Kultusminister erstellt und war noch auf den schulischen Bereich konzentriert (Avenarius u.a. 2003). Inzwischen ist die Berichterstattung auf das Bildungswesen im Ganzen gerichtet und als Kooperationsprojekt von Bund und Ländern angelegt. Für den Bericht zuständig ist ein Konsortium bestehend aus mehreren Forschungsinstituten sowie dem Statistischen Bundesamt und den Statistischen Ämtern der Länder. Kennzeichnend ist die indikatorengestützte Form des Berichtsystems. Es handelt sich also um Darstellungen zu zentralen Bereichen und Entwicklungen im Bildungssystem, die an Hand ausgewählter Indikatoren verdeutlicht werden, und nicht um Stellungnahmen, Wertungen oder Empfehlungen zu Reformmaßnahmen.

Der Bildungsbericht gliedert sich nach den Bereichen des Bildungswesens (Frühkindliche Betreuung, Schule und non-formale Lernwelten im Schulalter, Berufliche Ausbildung, Hochschule, Weiterbildung und Lernen im Erwachsenenalter) und bezieht außerdem die Rahmenbedingungen und Erträge von Bildung ein. Zudem wird in den Berichten ein jeweils wechselnder thematischer Schwerpunkt behandelt. Im Bericht des Jahres 2006 war dieses Schwerpunktthema Migration (Konsortium Bildungsberichterstattung 2006), für den Bericht im Jahr 2008 wird das Schwerpunktthema Übergänge sein. Überwiegend greift die Bildungsberichterstattung auf verfügbares Datenmaterial der statistischen Ämter sowie aus überregionalen empirischen Erhebungen (PIRLS/IGLU, PISA, SOEP, ALLBUS) zurück, teils werden Daten auch bezüglich der Bereiche und Fragestellungen des Berichts neu ausgewertet. Die Indikatoren werden, je nach Verfügbarkeit der Daten, international vergleichend, im Vergleich der Länder sowie gegliedert nach Geschlecht und Merkmalen der sozialen Herkunft vorgestellt. Neben dem nationalen Bericht liegen inzwischen auch Bildungsberichte einzelner Länder (Staatsinstitut für Schulqualität und Bildungsforschung 2006) und Städte (Schul- und Kultusreferat der Landeshauptstadt München 2006) vor, die sich in ihrer Konzeption an die indikatorengestützte Form des nationalen Bildungsberichts anlehnen.

Auch wenn die Bildungsberichterstattung als etabliertes Programm bezeichnet werden kann, bedeutet das nicht, dass alle damit verbundenen Herausforderungen bereits bewältigt wären. Oftmals sind Daten, die von Interesse wären, entweder gar nicht verfügbar oder nicht in dem gewünschten Differenzierungsgrad (z.B. nach regionalen Einheiten oder Subgruppen). Außerdem liegen nicht alle nötigen Daten auf einer über die Zeit oder Länder hinweg einheitlichen Erhebungsbasis vor und können als ausreichend

belastbar angesehen werden. Insbesondere besteht ein erhebliches Defizit an Daten zu Bildungsverläufen (Kristen u.a. 2005). Mit der Bildungsberichterstattung bewegt man sich von daher wenigstens partiell auf unsicherem Boden. Unklar ist auch, wer die Bildungsberichte nutzt und worin ihr spezifischer Beitrag zur Qualitätsentwicklung genau besteht. Zwar ist das Interesse an den Berichten seitens der Politik und Öffentlichkeit groß, ihre Wirkung ist jedoch schwer einzuschätzen. Interessant ist von daher der Ansatz in Bayern, den landeseigenen Bildungsbericht in regionalen Konferenzen der Qualitätsagentur mit Vertretern der Schulaufsicht zu besprechen und dabei u.a. auch nicht ohne Weiteres erklärable Unterschiede zwischen den Regionen zu thematisieren (Differenzen in der Häufigkeit von Klassenwiederholungen, Verteilungen der Schulübertritte und Schulabschlüsse).

### 3. Diskussion und Ausblick

Die Notwendigkeit von Qualitätssicherung an Schulen ist nicht (mehr) umstritten. Dass Zielklarheit und Erfolgskontrolle in jedem System nötig sind, das erfolgreich operieren soll, steht außer Frage. Der Aus- bzw. Umbau der Maßnahmen zur Qualitätssicherung an den Schulen in Deutschland steht dabei mit mehreren anderen Neu- resp. Umorientierungen in Beziehung und zielt auf eine Optimierung, höhere Erträge und eine größerer Effektivität des Systems. Schon diese Begriffe lösen im Bildungsbereich mitunter Befremden und eine mehr oder weniger stark ausgeprägte Reserviertheit aus (Böttcher 2002; Fend 2000). Dies beruht mit auf der Befürchtung, dass von den Schulen erwartet werden könnte, trotz weiterhin gleich bleibend suboptimaler Bedingungen bessere Ergebnisse erreichen zu müssen. Von daher kann auch kaum davon ausgegangen werden, dass alle Initiativen zur schulischen Qualitätssicherung auf reine Gegenliebe oder einhellige Akzeptanz stoßen. Es dürfte eher so sein, dass selbst die tatsächliche Implementierung der Maßnahmen im System noch Zeit in Anspruch nehmen wird. Ebenso wenig kann unterstellt werden, dass die Maßnahmen im Einzelnen schon in eine endgültige Form gebracht sind. Ausgehend von den Erfahrungen in der Anwendung wird man ihre Wirksamkeit überprüfen und entsprechend adaptieren und ggf. revidieren müssen. Überdies ist zu klären, wieweit einzelne Maßnahmen in einem sinnvollen Bezug zueinander stehen und bestmöglich koordiniert werden können, um unnötige Belastungen und Doppelungen zu vermeiden. Somit erscheint auch die Erwartung, dass nun alles sehr schnell sehr viel besser wird, als zwar verständlich aber kaum erfüllbar.

Bildungsstandards und Kompetenzorientierung, interne und externe Evaluationen sowie die Bildungsberichterstattung stellen aufeinander zu beziehende Elemente eines Systems der Qualitätssicherung dar. Sie bieten jedoch noch keine Gewähr für Qualitätsverbesserungen, sondern zielen zunächst einmal darauf ab, Wissen über das schulische System zu gewinnen und Hinweise auf Möglichkeiten für Ziel führende Veränderungen zu generieren. Qualitätssicherung findet dabei nicht einfach von oben nach unten statt, sondern im Zusammenspiel der verschiedenen Ebenen (Systemebene, Schulebene, Schulklassenebene). Das schulische System ist wie jedes andere System eigensinnig und

lebt von der Anpassungsfähigkeit der einzelnen Teilsysteme und Akteure. Ohne Akzeptanz und aktive Mitwirkung auf allen Ebenen sind Veränderungen nicht zu bewirken. Teilweise muss diese Akzeptanz erst noch geschaffen und der Sinn, Zweck und Nutzen von Qualitätssicherung verdeutlicht werden.

Die Initiativen zur Qualitätssicherung stellen nicht nur eine Herausforderung für die Schulen dar, sondern für das schulische System als Ganzes. Handlungsbedarf kann auf allen Ebenen und bezüglich unterschiedlichster Aspekte bestehen, nicht zuletzt auch hinsichtlich der Rahmenbedingungen für die schulische Arbeit. Da Qualität immer auch eine normative Komponente beinhaltet (Heid 2000), wird man dazu Stellung beziehen müssen, was die übergeordneten bildungspolitischen Leitlinien sind, an denen sich die Qualitätsentwicklung ausrichtet. Das deutsche Schulsystem hat den Ruf, ein stark selektives System zu sein, das Lerner mit ungünstigeren Voraussetzungen frühzeitig aussondert bzw. nach „unten“ weitergibt. Eine Leitidee könnte daher sein, dem durch frühzeitige Intervention entgegenzuwirken. Dann wäre es allerdings angezeigt, nicht erst am Ende oder gegen Ende der schulischen Laufbahnen die Einhaltung von Bildungsstandards zu überprüfen, sondern diagnostische Verfahren frühzeitig systematisch anzuwenden, um bei ungünstigen Prognosen gegenzusteuern zu können.

Abschließend zu warnen ist vor einer simplen Wettbewerbs- und Rankingmentalität, die sich im Zuge von Leistungsüberprüfungen und Evaluationen zunehmend ausbreiten scheint. Wettbewerb lebt bekanntlich davon, dass es neben den strahlenden Siegern immer auch Verlierer gibt. Nun sind Schulen aber keine Wettbewerber auf einem freien Markt und ihr Ziel ist es nicht, sich gegenseitig die Kunden abzujagen. Wissenssharing, solidarische Zusammenarbeit in Netzwerken und der Ausbau von Unterstützungssystemen in der Fläche könnten daher Ziel führender sein als Wettbewerb. Schließlich sollen ja alle Schulen gute Schulen für alle Schüler sein.

## Literatur

- Arbeitsstab Forum Bildung (2001): Empfehlungen des Forum Bildung. Bonn: BLK.
- Avenarius, H./Ditton, H./Döbert, H./Klemm, K./Klieme, E./Rürup, M./Tenorth, H.-E./Weishaupt, H./Weiß, M. (2003): Bildungsberichterstattung für Deutschland. Erste Befunde. Opladen: Leske + Budrich.
- Blum, W./Drüke-Noe, C./Hartung, R./Köller, O. (Hrsg.) (2006): Bildungsstandards Mathematik: konkret; Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen. Berlin: Cornelsen.
- Böttcher, W. (2002): Kann eine ökonomische Schule auch eine pädagogische sein? Schulentwicklung zwischen Neuer Steuerung, Organisation, Leistungsevaluation und Bildung. Weinheim und München: Juventa.
- Ditton, H. (2000): Qualitätskontrolle und -sicherung in Schule und Unterricht – ein Überblick zum Stand der empirischen Forschung. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): Qualität und Qualitätssicherung im Bildungsbereich. Weinheim: Beltz, S. 73–92.
- Ditton, H. (2002): Evaluation und Qualitätssicherung. In: Tippelt, R. (Hrsg.): Handbuch Bildungsforschung. Opladen: Leske + Budrich, S. 775–790.
- Ditton, H./Arnoldt, B. (2004): Schülerbefragungen zum Fachunterricht – Feedback an Lehrkräfte. In: Empirische Pädagogik, 18(1), S. 115–139.

- Ditton, H./Krüskens, J. (2005): Orientierungsarbeiten Jahrgangsstufe 2 – Schuljahr 2003/04. Bericht über die Ergebnisse der Begleitstudie in Berlin und Brandenburg. Berlin: Senatsverwaltung für Bildung, Jugend und Sport.
- Fend, H. (2000): Qualität und Qualitätssicherung im Bildungswesen: Wohlfahrtsstaatliche Modelle und Marktmodelle. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): Qualität und Qualitätssicherung im Bildungsbereich. Weinheim: Beltz, S. 55–72.
- Groß Ophoff, J./Koch, U./Hosenfeld, I. (2006): Ergebnisrückmeldungen und ihre Rezeption im Projekt VERA. In: Kuper, H./Schneewind, J. (Hrsg.): Rückmeldung und Rezeption von Forschungsergebnissen – Zur Verwendung wissenschaftlichen Wissens im Bildungssystem. Münster: Waxmann, S. 19–40
- Heid, H. (2000): Qualität: Überlegungen zur Begründung einer pädagogischen Beurteilungskategorie. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): Qualität und Qualitätssicherung im Bildungsbe- reich. Weinheim: Beltz, S. 41–51.
- Helmke, A. (2004): Von der Evaluation zur Innovation: Pädagogische Nutzbarmachung von Ver- gleichsarbeiten in der Grundschule. In: Seminar, 2, S. 90–112.
- Hopkins, D./West, M./Skinner, J. (1995): "Improvement through Inspection?". A Critique of the OFS- TED Inspection System. School Evaluation in England and Wales. In: Zeitschrift für Sozialisations- forschung und Erziehungssoziologie, 15, S. 337–350.
- Klieme, E. (2004): Begründung, Implementation und Wirkung von Bildungsstandards: Aktuelle Dis- kussionslinien und empirische Befunde. In: Zeitschrift für Pädagogik, 50.Jahrgang (2004)(5), S. 625–635.
- Klieme, E./Avenarius, H./Blum, W./Döbrich, P./Gruber, H./Prenzel, M./Reiss, K./Riquarts, K./Rost, J./Tenorth, H.-E./Vollmer, H. (2003): Zur Entwicklung nationaler Bildungsstandards – Eine Exper- tise. Berlin: BMBF / DIPF.
- Konsortium Bildungsberichterstattung (2006): Bildung in Deutschland. Ein indikatorengestützter Be- richt mit einer Analyse zu Bildung und Migration. Bielefeld: Bertelsmann.
- Kristen, C./Römmel, A./Müller, W./Kalter, F. (2005): Längsschnittstudien für die Bildungsberichter- stattung - Beispiele aus Europa und Nordamerika. Gutachten im Auftrag des BMBF. Berlin: Bun- desministerium für Bildung und Forschung (BMBF).
- Kuper, H./Schneewind, J. (Hrsg.). (2006): Rückmeldung und Rezeption von Forschungsergebnissen - Zur Verwendung wissenschaftlichen Wissens im Bildungssystem. Münster: Waxmann.
- Rolff, H.-G. (2002): Rückmeldung und Nutzung der Ergebnisse von großflächigen Leistungsuntersuchung. Grenzen und Chancen. In: Rolff, H.-G./Holtappels, H. G./Klemm, K./Pfeiffer, H./Schulz- Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung Vol. 12). Weinheim: Juventa, S. 75–98
- Schrader, F.-W./Helmke, A. (2003): Evaluation - und was danach? Ergebnisse der Schulleiterbefragung im Rahmen der Rezeptionsstudie WALZER. In: Schweizerische Zeitschrift für Bildungswissen- schaften, 25(1), S. 79–110.
- Schul- und Kultusreferat der Landeshauptstadt München (2006): Erster Münchner Bildungsbericht. Bildung zu einer Angelegenheit vor Ort machen. München: Landeshauptstadt München.
- Staatsinstitut für Schulqualität und Bildungsforschung - Qualitätsagentur (2006a): Bestandsaufnahme zur internen Evaluation in Bayern. München.
- Staatsinstitut für Schulqualität und Bildungsforschung - Qualitätsagentur (2006b): Konsequenzen aus der externen Evaluation an Bayerns Schulen. München.
- Staatsinstitut für Schulqualität und Bildungsforschung (2006): Bildungsberichterstattung 2006. Mün- chen: Kastner.
- Weinert, F. E. (1999): Concepts of Competence. Definition and Selection of Competencies: OECD.

*Anschrift des Autors:*

Prof. Dr. Hartmut Ditton, Ludwig-Maximilians-Universität München, Institut für Pädagogik, Lehr- stuhl für Allgemeine Pädagogik, Erziehungs- und Sozialisationsforschung, Leopoldstr. 13, 80802 Mün- chen.

**Anhang**

<b>Tab. 1: Qualitätsagenturen in den Bundesländern</b>	
<b>Bundesland</b>	<b>Qualitätsagentur</b>
<b>Baden-Württemberg</b>	Evaluation durch Landesinstitut für Qualitätsentwicklung (LS)
<b>Bayern</b>	Qualitätsagentur Staatsinstitut für Schulqualität und Bildungsforschung
<b>Berlin/Brandenburg</b>	ISQ e. V. - Institut für Schulentwicklung der Länder Berlin und Brandenburg
<b>Bremen</b>	LIS - Landesinstitut für Schule
<b>Hamburg</b>	Landesinstitut für Lehrerbildung und Schulentwicklung
<b>Hessen</b>	IQ Hessen - Institut für Qualitätsentwicklung
<b>Mecklenburg-Vorpommern</b>	L.I.S.A. - Landesinstitut für Schule und Ausbildung
<b>Niedersachsen</b>	NIILS - Niedersächsisches Landesamt für Lehrerbildung und Schulentwicklung
<b>Nordrhein-Westfalen</b>	Lfs/QA - Landesinstitut für Schule/ Qualitätsagentur (bis 01/2007)
<b>Rheinland-Pfalz</b>	AQS - Agentur für Qualitätssicherung
<b>Saarland</b>	Ministerium für Bildung, Kultur und Wissenschaft; Abt. Qualitätssicherung
<b>Sachsen</b>	Comenius-Institut - Sächsisches Staatsinstitut für Bildung und Schulentwicklung
<b>Sachsen-Anhalt</b>	Landesverwaltungsamt Unterrichtsversorgung, Datenerhebung, Schulentwicklungsplanung
<b>Schleswig-Holstein</b>	IQHS - Institut für Qualitätsentwicklung an Schulen
<b>Thüringen</b>	Thüringer Kultusministerium; Abt. Schulentwicklung

Tab. 2: Externe Evaluation in den Bundesländern		
Bundesland	Inspektion	
<b>Baden-Württemberg</b>	<b>Schulinspektion</b>	
	Umfang:	zweijährige Pilotphase auf freiwilliger Basis
	Team:	2-3 Personen, Lehrkräfte (mind. eine aus der zu evaluierenden Schulform) und eine assoziierte Fachkraft (außerschulisch)
	Weiterleitung des Berichts:	Schule und Schulaufsicht
	Zeitraumen der Evaluation:	2-3 Tage vor Ort
	Unterrichtsbeobachtung:	ja
<b>Bayern</b>	<b>Schulinspektion</b>	
	Umfang:	504 Schulen, Ausweitung geplant
	Team:	64 Teams bestehend aus externem Experten, Sprecher und zwei schulischen Vertretern
	Weiterleitung des Berichts:	an Schulaufsicht für mögliche Handlungs- und Zielvereinbarungen, bzw. weitere Maßnahmen
	Zeitraumen der Evaluation:	ca. 3 Monate
	Unterrichtsbeobachtung:	–
<b>Berlin</b>	<b>Externe Evaluation durch Inspektion</b>	
	Umfang:	Zyklus von 5 Jahren, derzeit 160 Schulen, bis zum Jahr 2009/10 sollen alle 800-850 Schulen evaluiert werden
	Team:	Schulaufsicht, Schulleitungsmitglieder und Lehrkräfte, insgesamt 9 Inspektionsteams
	Weiterleitung des Berichts:	Schulaufsicht
	Zeitraumen der Evaluation:	ca. 3 Monate
	Unterrichtsbeobachtung:	nein
<b>Brandenburg</b>	<b>Schulvisitation</b>	
	Umfang:	jährlich 150 Schulen
	Team:	3 Personen, haupt- und nebenberufl. Visitatoren
	Weiterleitung des Berichts:	an Schule und Schulaufsicht
	Zeitraumen der Evaluation:	ca. 10 Wochen (ohne Vorlaufphase), 2 Tage vor Ort
	Unterrichtsbeobachtung:	ja, ca. 20-25 min.
<b>Bremen</b>	<b>Schulinspektion</b>	
		Infos im Aufbau
<b>Hamburg</b>	<b>Schulinspektion</b>	
	Umfang:	alle 402 staatl. Schulen im Zyklus von 4 Jahren
	Team:	8 Inspektoren, 2 wiss. Mitarbeiter, Leitung, 2 Verwaltungskräfte
	Weiterleitung des Berichts:	Schule und ggf. Weiterleitung an andere
	Zeitraumen der Evaluation:	ca. 16 Wochen
	Unterrichtsbeobachtung:	ja, ca. 20 min.

Tab. 2: <b>Externe Evaluation in den Bundesländern</b> (Fortsetzung)		
<b>Bundesland</b>	<b>Inspektion</b>	
<b>Hessen</b>	<b>Schulinspektion</b>	
	<i>Umfang:</i>	innerhalb von 4-5 Jahren 2000 Schulen
	<i>Team:</i>	2-5 Inspektoren; Pädagogen und Psychologen
	<i>Weiterleitung des Berichts:</i>	an Schule und Schulaufsicht
	<i>Zeitraumen der Evaluation:</i>	10 Wochen (ohne Vorlaufphase), 2-4 Tage vor Ort
	<i>Unterrichtsbeobachtung:</i>	ja, 20 min.
<b>Mecklenburg-Vorpommern</b>	<b>Externe Evaluation</b>	
	<i>Umfang:</i>	–
	<i>Team:</i>	Vertreter der L.I.S.A. (Dezernat für Qualitätssicherung) als Teamleiter, Schulrat, Schulleiter einer anderen Schule
	<i>Weiterleitung des Berichts:</i>	Schule, Schulaufsicht, Schulrat
	<i>Zeitraumen der Evaluation:</i>	ca. 12 Wochen, ca. 1 Woche vor Ort
	<i>Unterrichtsbeobachtung:</i>	ja
<b>Niedersachsen</b>	<b>Schulinspektion</b>	
	<i>Umfang:</i>	alle staatl. Allgem. Und berufl. Schulen
	<i>Team:</i>	externe Fachleute, Beschäftigte im gehobenen und mittleren Dienstes
	<i>Weiterleitung des Berichts:</i>	Schule, Schulaufsicht, Eltern- und Schülervertretung, Personalvertretung
	<i>Zeitraumen der Evaluation:</i>	ca. 3 Monate
	<i>Unterrichtsbeobachtung:</i>	ja, ca. 20 min.
<b>NRW</b>	<b>Qualitätsanalyse</b>	
	<i>Umfang:</i>	alle allgemeinbildenden Schulen
	<i>Team:</i>	3 Personen mit mind. 1 Experte aus der jeweiligen Schulform
	<i>Weiterleitung des Berichts:</i>	Schule, Schulaufsicht, Schulträger
	<i>Zeitraumen der Evaluation:</i>	ca. 3 Tage vor Ort, insgesamt ca. 14 Wochen
	<i>Unterrichtsbeobachtung:</i>	ja
<b>Rheinland-Pfalz</b>	<b>Externe Evaluation</b>	
	<i>Umfang:</i>	alle staatl. Schulen (1615) innerhalb von 3 Jahren
	<i>Team:</i>	–
	<i>Weiterleitung des Berichts:</i>	an Schule, Schulaufsicht und Ministerium für Bildung
	<i>Zeitraumen der Evaluation:</i>	ca. 16 Wochen
	<i>Unterrichtsbeobachtung:</i>	ja

Tab.2: Externe Evaluation in den Bundesländern (Fortsetzung)		
Bundesland	Inspektion	
<b>Saarland</b>	<b>Schulbesuche</b>	
	Umfang:	in der Erprobungsphase derzeit nur alle Grundschulen
	Team:	–
	Weiterleitung des Berichts:	Schulleitung, Kollegium, Schulaufsicht, Schulträger, Schul-Elternvertretung
	Zeitraumen der Evaluation:	ca. 10 Wochen
	Unterrichtsbeobachtung:	ja, eher Unterrichtsbesuch
<b>Sachsen</b>	<b>Externe Schulevaluation</b>	
	Umfang:	derzeit 50 Schulen, ab 2007/08 sollen alle allgem. Schulen im Zyklus von 5 Jahren evaluiert werden
	Team:	–
	Weiterleitung des Berichts:	Schule, Kontaktgruppe (wird von der Schule eigens für die Evaluation gebildet)
	Zeitraumen der Evaluation:	ca. 12 Wochen, 3 Tage vor Ort
	Unterrichtsbeobachtung:	ja, ca. 45 min.
<b>Sachsen-Anhalt</b>	<b>Landesweite Erhebungen und Schulinspektion</b>	
	Umfang:	derzeit ca. 200 Schulen, alle Schulen geplant in 2007 im Zyklus von 5 Jahren
	Team:	–
	Weiterleitung des Berichts:	Schule, Schulträger, schulfachliche Referenten, Grundlage für Landesbericht
	Zeitraumen der Evaluation:	ca. 2 Monate, 2-3 Tage vor Ort
	Unterrichtsbeobachtung:	ja, ca. 20-25 min.
<b>Schleswig-Holstein</b>	<b>EVIT-Externe Evaluation im Team</b>	
	Umfang:	landesweit
	Team:	3-er Team: Vertreter der Schulaufsicht, Vertreter des IQHS, Schulleiter einer vergleichbaren Schule
	Weiterleitung des Berichts:	Schule, Schulträger
	Zeitraumen der Evaluation:	Zeitraumen 4 Jahre, 1-3 Tage vor Ort
	Unterrichtsbeobachtung:	ja
<b>Thüringen</b>	<b>Entwicklungsvorhaben "eigenverantwortliche Schule"- EVS</b>	

Tab. 3: Vergleichsarbeiten in den 16 Bundesländern

Bundesland	Fächer	Grundschule					HS / GS / MS **					Realschule					Gymnasium																
		Jahrgang					Jahrgang					Jahrgang					Jahrgang																
		2	3	4	5	6	6	7	8	9	10	5	6	7	8	9	10	5	6	7	8	9	10	5	6	7	8	9	10	Koll.			
<b>Baden-Württemberg</b>	Deutsch	X				X						X								X													
	Mathematik	X				X						X								X													
	Englisch					X																											
	Sonstige *																			X													
<b>Bayern</b>	Deutsch	X	X			X					X								X														
	Mathematik	X	X			X					X								X														
	Englisch										X																						
	Sonstige *																																
<b>Berlin</b>	Deutsch	X		X							X								X														
	Mathematik	X		X							X								X														
	Englisch										X																						
	Sonstige *																																
<b>Brandenburg</b>	Deutsch	X		X							X								X														
	Mathematik	X		X							X								X														
	Englisch										X																						
	Sonstige *																																
<b>Bremen</b>	Deutsch																																
	Mathematik					X					X								X														
	Englisch					X					X								X														
	Sonstige *																																
<b>Hamburg</b>	Deutsch		X			X					X								X														
	Mathematik		X			X					X								X														
	Englisch					X					X								X														
	Sonstige *																																

\* i.d.R.: Französisch, Latein \*\* Hauptschule / Gesamtschule / Schule mit mehreren Bildungswegen

Tab. 3: Vergleichsarbeiten in den 16 Bundesländern (Fortsetzung)

Bundesland	Fächer	Grundschule					HS / GS / MS **					Realschule					Gymnasium													
		Jahrgang					Jahrgang					Jahrgang					Jahrgang													
		2	3	4	5	6	7	8	9	10	5	6	7	8	9	10	5	6	7	8	9	10	5	6	7	8	9	10	Koll.	
<b>Mecklenburg-Vorpommern</b>	Deutsch		X		X		X		X			X																		
	Mathematik		X		X		X		X			X																		
	Englisch																													
	Sonstige *																													
<b>Niedersachsen</b>	Deutsch			X																										
	Mathematik		X						X																	X				
	Englisch																													
	Sonstige *																													
<b>NRW</b>	Deutsch		X	X					X																	X				
	Mathematik		X	X					X																	X				
	Englisch																													
	Sonstige *																													
<b>Rheinland-Pfalz</b>	Deutsch	X																												
	Mathematik	X																												
	Englisch																													
	Sonstige *																													
<b>Saarland</b>	Deutsch												X																	
	Mathematik																									X				
	Englisch																													
	Sonstige *																													
<b>Sachsen</b>	Deutsch			X																										
	Mathematik		X																							X				
	Englisch																													
	Sonstige *																													

\* i.d.R.: Französisch, Latein    \*\* Hauptschule / Gesamtschule / Schule mit mehreren Bildungswegen

Tab. 3: Vergleichsarbeiten in den 16 Bundesländern (Fortsetzung)

Bundesland	Fächer	Grundschule						HS / GS / MS **				Realschule						Gymnasium							
		Jahrgang						Jahrgang				Jahrgang						Jahrgang							
		2	3	4	5	6	7	6	7	8	9	10	5	6	7	8	9	10	5	6	7	8	9	10	Koll.
<b>Sachsen-Anhalt</b>	Deutsch		X	X			X	X													X	X			
	Mathematik		X	X			X	X													X	X			
	Englisch					X															X				
	Sonstige *						X														X				
<b>Schleswig-Holstein</b>	Deutsch		X			X							X								X				
	Mathematik		X			X		X					X								X				
	Englisch					X							X								X				
	Sonstige *																								
<b>Thüringen</b>	Deutsch		X			X							X								X				X
	Mathematik		X			X		X					X								X				X
	Englisch		X			X		X					X								X				X
	Sonstige *																								X

\* i.d.R.: Französisch, Latein \*\* Hauptschule / Gesamtschule / Schule mit mehreren Bildungswegen

Olaf Köller

# **Bildungsstandards in einem Gesamtsystem der Qualitätssicherung im allgemein bildenden Schulsystem Deutschlands**

## **1. Überblick**

Als Folge der Publikationen zu den Ergebnissen der *Programme of International Student Assessment-Studie* (PISA) 2000 (vgl. u.a. Deutsches PISA-Konsortium 2001, 2002) haben sich in den deutschsprachigen Ländern Österreich, Schweiz, Luxemburg und Deutschland erhebliche Veränderungen in der Steuerung des allgemein bildenden Schulsystems ergeben, die gern als Paradigmenwechsel von der Input- zur Outputsteuerung beschrieben werden (vgl. Klieme u.a. 2003). Erreichte fachliche Kompetenzniveaus auf Seiten der Schülerinnen und Schüler sollen Rückschlüsse auf den Erfolg des Bildungssystems zulassen. Ausdruck findet diese Outputsteuerung unter anderem in der langfristigen Beteiligung an internationalen *Large-Scale-Assessments* wie PISA (vgl. z.B. Deutsches PISA-Konsortium 2007) oder der Progress in Reading Literacy Study (PIRLS) / Internationalen Grundschul-Lese-Untersuchung (IGLU) (vgl. z.B. Bos u.a. 2007), aber auch in der Verabschiedung von Bildungsstandards, in denen für ausgewählte Fächer definiert ist, über welche Kompetenzen Schülerinnen und Schüler zu festen Zeitpunkten ihrer schulischen Karriere verfügen sollten.

Mit ihren Beschlüssen von 2003 und 2004 hat die Ständige Konferenz der Kultusminister (KMK) für alle 16 Länder der Bundesrepublik Deutschland solche verbindlichen Bildungsstandards für die Grundschule und das Ende der Sekundarstufe I verabschiedet. Trotz Föderalismus sind die darin vereinbarten Ziele schulischer Lehr-Lernprozesse für alle 16 Länder bindend, dementsprechend haben sich auch alle Länder verpflichtet, Sorge zu tragen, dass die Standards und die mit ihnen verbundenen Ziele im Unterricht umgesetzt werden. Für die Sekundarstufe I wurden bewusst keine schulformspezifischen, sondern abschlussbezogene Standards für den Hauptschul- und Mittleren Schulabschluss definiert. Damit wurde dem Umstand Rechnung getragen, dass in allen Ländern Deutschlands derselbe Abschluss an unterschiedlichen Schulformen vergeben wird (vgl. z.B. Köller u.a. 2004).

Mit den KMK-Beschlüssen vom Oktober 2007 ist der Auftrag dahingehend erweitert worden, auch Standards für das Ende der gymnasialen Oberstufe zu erarbeiten, und zwar erneut für die Fächer Deutsch, erste Fremdsprache (Französisch, Englisch), Mathematik und die drei Naturwissenschaften (Biologie, Chemie, Physik). Damit werden in absehbarer Zeit in den Kernfächern abschlussbezogene Standards für das gesamte allgemein bildende Schulsystem in Deutschland vorliegen. Dies fasst noch einmal die Tabelle 1 im Überblick zusammen.

Tab. 1: **Fächer, für die laut KMK-Beschluss Bildungsstandards vorliegen bzw. erarbeitet werden, nach Abschluss**

	Ende der 4. Jahrgangsstufe	Ende der Sekundarstufe I		Ende der gymnasialen Oberstufe
		Hauptschulabschluss	Mittlerer Schulabschluss	Allgemeine Hochschulreife
Deutsch	■	■	■	■
Mathematik	■	■	■	■
Englisch		■	■	■
Französisch		■	■	■
Biologie			■	■
Chemie			■	■
Physik			■	■

An die Verabschiedung der Bildungsstandards hat die KMK in ihren Beschlüssen vom Juni 2006 (vgl. KMK 2006) ein breites System der Qualitätssicherung gekoppelt, das neben stichprobenbasierten internationalen und nationalen Vergleichen auch regelmäßige flächendeckende Vergleichsarbeiten vorsieht und dazu dienen soll zurückzumelden, ob Schülerinnen und Schüler die in den Standards formulierten Leistungserwartungen einlösen können. Flankiert wird dieses System durch ein indikatorengestütztes Monitoring auf gesamtstaatlicher Ebene im Rahmen der nationalen Bildungsberichterstattung (vgl. Klieme u.a. 2007) und durch die systematische Evaluierung von Einzelschulen, wofür nahezu alle Länder inzwischen Schulinspektionsteams eingerichtet haben.

Das Ziel dieses Beitrages besteht darin, das Gesamtsystem der Qualitätssicherung im deutschen Schulsystem, wie es sich aus den Bildungsstandards ergibt, vorzustellen und mit seinen Implikationen und Nebenwirkungen zu diskutieren. Darüber hinaus soll im Einklang mit den Ausführungen von Klieme u.a. (2003) präzisiert werden, welches Potenzial die in Deutschland verabschiedeten Bildungsstandards für die Unterrichtsentwicklung haben und dass sie ihre Wirkung auch nur dann entfalten werden, wenn sie die Grundlagen sowohl von qualitätssichernden als auch qualitätsentwickelnden Maßnahmen sind.

## 2. Bildungsstandards im allgemeinbildenden Schulsystem Deutschlands

Die KMK hat als Reaktion auf PISA 2000 im Mai 2002 beschlossen, länderübergreifende Bildungsstandards in den Fächern Deutsch und Mathematik sowie für die erste Fremdsprache (Englisch, Französisch) für den Mittleren Schulabschluss und für den Hauptschulabschluss zu erarbeiten. Die anschließende konzeptionelle Rahmung um die Ein-

führung der Bildungsstandards wurde maßgeblich durch die Expertise „Zur Entwicklung nationaler Bildungsstandards“ (Klieme u.a. 2003) vorangebracht. Sie beschreibt die Funktionen, konzeptuellen Grundlagen und die Entwicklung von Bildungsstandards. In ihrer Expertise haben Klieme u.a. (2003) ausgeführt, wodurch sich Standards auszeichnen sollen. Danach benennen sie fachspezifisch die wesentlichen Ziele pädagogischer Arbeit, ausgedrückt als erwünschte Lernergebnisse bzw. Kompetenzen der Schülerinnen und Schüler zu bestimmten Zeitpunkten ihrer Bildungsbiographien. In den entsprechenden KMK-Dokumenten wurde dies dahin gehend umgesetzt, dass Standards als sogenannte *Can do Statements* formuliert sind. Beispielsweise heißt es im Bereich der ersten Fremdsprache für den Mittleren Schulabschluss für den Kompetenzbereich Lesen: „Die Schülerinnen und Schüler können Korrespondenz lesen, die sich auf das eigene Interessengebiet bezieht und die wesentliche Aussage erfassen (B2); klar formulierte Anweisungen, unkomplizierte Anleitungen, Hinweise und Vorschriften verstehen (B1/ B2) (...)“ (KMK 2004, S. 14). Kompetenzen sind hier als gezeigtes Verhalten definiert. Folgt man einem stärker psychologisch geprägten Kompetenzkonzept (vgl. z.B. Weinert 2001), so wird man nicht das Verhalten selbst, sondern die bei den Schülerinnen und Schülern verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, die notwendig sind, um bestimmte Probleme bzw. Aufgaben lösen zu können, als Kompetenzen verstehen. In diesem Fall gewinnen Kompetenzen den Charakter psychologischer Konstrukte, welche mit Hilfe von Messinstrumenten operationalisiert werden können.

### *2.1 Von Bildungsstandards zu nationalen Skalen – Aufgabenentwicklung und Normierung*

Bildungsstandards sind in den entsprechenden Dokumenten der KMK zunächst verbal formuliert (vgl. z.B. für die Fremdsprachen KMK 2004). Sie benennen mehr oder weniger konkret die fachspezifischen Kompetenzen, die Schülerinnen und Schülern in der Regel erreichen sollten (sogenannte Regelstandards). Was genau „in der Regel“ heißt, ist dabei unklar und nicht ganz zu Unrecht verschiedentlich kritisiert worden (vgl. z.B. Bremerich-Vos 2003; Spinner 2005). Man wird dieser Kritik allerdings erst begegnen können, wenn auf der Basis der Standards normierte Aufgabenpools und damit verbunden nationale Skalen entwickelt wurden, die eine Verortung von Schülerinnen/ Schülern und Aufgaben (Items) erlauben. Dann wird es möglich sein, Leistungsbereiche zu definieren, die Schülerinnen und Schüler in der Regel (Regelstandards) oder mindestens (Mindeststandards; vgl. zu dieser Debatte auch Klieme u.a. 2003) am Ende der 4. Jahrgangsstufe oder am Ende der Sekundarstufe I erreicht haben sollten.

Mit der Gründung des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) hat die KMK im Jahr 2004 die Voraussetzung geschaffen, entsprechende Aufgabenpools und nationale Skalen zu entwickeln (vgl. zu den Aufgaben des IQB Köller 2005). Der Prozess – von den Standards zu nationalen Skalen – soll im Folgenden allgemein beschrieben werden. Für das Fach Mathematik in der Sekundarstufe I findet sich eine aus-

fürliche Beschreibung des Prozesses bei Ehmke u.a. (2006). Rupp u.a. (2008) beschreiben den Aufgabenentwicklungsprozess für das Fach Englisch.

Die Itementwicklung auf der Basis der Bildungsstandards folgt weitgehend dem *State of the Art* (vgl. z.B. Rost 2004), hat allerdings einige Besonderheiten zu berücksichtigen. Hierzu zählt unter anderem in der Sekundarstufe I, dass Aufgaben für unterschiedliche Jahrgangsstufen (8, 9 und 10) und Abschlüsse (Hauptschulabschluss und Mittlerer Schulabschluss) generiert werden müssen, die sich dennoch auf einer gemeinsamen Metrik abbilden lassen. Im Wesentlichen vollzieht sich der Prozess in sieben Schritten:

1. *Fachdidaktisch und lernpsychologisch fundierte Konkretisierung der Kompetenzen in den jeweiligen Fächern*: Mit Blick auf die Definition von Messmodellen ist es unumgänglich zu präzisieren, was unter den jeweiligen Kompetenzen (Konstrukten) zu verstehen ist, konkret: Wie ist beispielsweise das Hör- und Leseverstehen in der Fremdsprache theoretisch verankert? Die dafür notwendigen konzeptionellen Arbeiten am IQB finden in enger Kooperation mit Fachdidaktikerinnen und Fachdidaktikern statt.
2. *Erarbeitung von Richtlinien zur Konstruktion von Testaufgaben/Items (Item- und Testspezifikationen)*: Hier geht es um die Erstellung von Trainingsmaterial, das bei der Schulung von Aufgabenentwicklern eingesetzt werden kann. In diesen Materialien werden die Konstrukte/Kompetenzen erläutert, noch wichtiger sind allerdings die Ausführungen, wie Tests üblicherweise konstruiert werden und welche Fehler man bei der Itementwicklung vermeiden sollte. Dies erfolgt im Einklang mit gängigen Werken zur Testkonstruktion (vgl. z.B. Rost 2004).
3. *Initiierung und Aufrechterhaltung der Itementwicklung durch erfahrene Lehrkräfte*: Mit Hilfe der in Schritt 2 erarbeiteten Materialien werden im dritten Schritt Lehrkräfte aller 16 Länder geschult. Anschließend generieren diese Lehrkräfte unter fachdidaktischer Betreuung die Testaufgaben in regionalen Arbeitsgruppen. Lehrkräfte im Bereich der Aufgabenentwicklung zu professionalisieren, erfüllt hier auch den Zweck, die so erreichte Expertise später in die jeweiligen Länder zu tragen, um sie dort bei Arbeiten zur Qualitätssicherung anzuwenden.
4. *Begutachtung der Items durch Experten*: Bildungsforscher, Psychometriker und Fachdidaktiker beurteilen und kommentierten in diesem Schritt die Items hinsichtlich ihrer psychometrischen und fachdidaktischen Güte. Dazu zählt auch die Klassifikation der Aufgaben hinsichtlich der durch sie gemessenen Kompetenzen und Anforderungsbereiche.
5. *Empirische Erprobung der entwickelten Aufgaben in großen Schülerstichproben*: Sämtliche Items werden in großen Stichproben erprobt. Die Stichprobengröße wird dabei so gewählt, dass hinreichend robuste Schätzungen der Itemparameter auf der Grundlage von Modellen der Item-Response-Theory möglich sind. Zusätzlich werden kleinere empirische Untersuchungen zur Validierung der Items durchgeführt.
6. *Normierung der Aufgaben*: Auf der Basis national repräsentativer Stichproben von Schülerinnen und Schülern werden die Items kalibriert und nationale Skalen wer-

den definiert. Neben den neu entwickelten Instrumenten werden auch etablierte Leistungstests eingesetzt, um so zusätzliche Validitätshinweise zu erhalten.

7. *Definition von Kompetenzstufen (Standard-Setting)*: Im Konsens zwischen Fachdidaktikern, Erziehungswissenschaftlern, Lehrkräften und Vertretern der Politik und unter Einbeziehung von empirische Befunden aus den Pilotierungs- und Normierungsstudien werden Niveaustufen definiert und festgestellt, bei welchen Werten auf den nationalen Skalen davon ausgegangen werden kann, dass die entsprechenden Schülerinnen und Schüler die in den Standards beschriebenen Leistungserwartungen erfüllt haben. Dieses konsensuelle Verfahren wird dabei einem psychometrischen Vorgehen, bei dem schwierigkeitsgenerierende Merkmale von Aufgaben zur Definition von Niveaustufen bzw. deren Grenzen verwendet werden (vgl. Hartig 2007), vorgezogen.

Die Abbildung 1 zeigt am Beispiel Deutsch und Mathematik in der Grundschule, mit welchen Zeiträumen bei der Abarbeitung der Schritte 1 bis 7 zu rechnen ist. *Cum grano salis* vergehen vier Jahre vom Beginn der Aufgabenentwicklung bis zum Vorliegen nationaler Skalen.

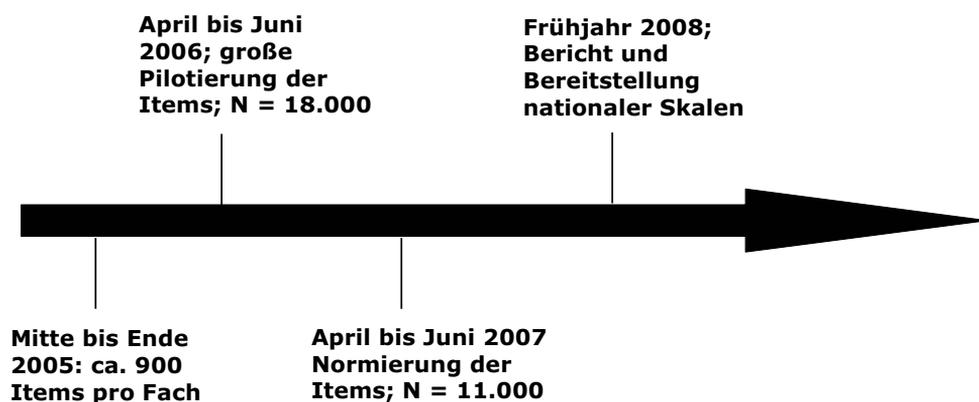


Abb. 1: Zeitlicher Ablauf bei der Standard-basierten Aufgabenentwicklung für die Fächer Deutsch und Mathematik in der Grundschule

Im Prozess der Itementwicklung wurden pro Fach ca. 900 Items für die Jahrgangsstufen 3 und 4 erstellt, die jeweils den allgemeinen und inhaltlichen Kompetenzen sowie den Anforderungsniveaus zugeordnet wurden. Die Aufgabenentwicklung für die beiden nebeneinander liegenden Jahrgänge war nötig, da die Standards zum einen für das Ende der vierten Jahrgangsstufe definiert sind, zukünftige standard-basierte Vergleiche allerdings in der dritten Jahrgangsstufe durchgeführt werden sollen (vgl. 2.2).

Die Aufgaben wurden im Frühjahr 2006 im Rahmen von IGLU (vgl. Bos u.a. 2007) in einer national repräsentativen Stichprobe der 3. und 4. Jahrgangsstufe erprobt. Inse-

samt bearbeiteten rund 18.000 Schülerinnen und Schüler in einem äußerst komplexen *Multi-Matrix-Design* die rund 900 Items pro Fach. Konkret wurde ein Untersuchungsplan umgesetzt, bei dem jede Schülerin / jeder Schüler nur eine Teilmenge der Aufgaben vorgelegt wurde. Die Testung dauerte zwei Schulstunden. Das Ziel dieser Pilotstudie bestand allein darin, solche Teilaufgaben zu identifizieren, die für Erstellung einer nationalen Metrik und sich daran anschließende Vergleichsstudien (s.u.) geeignet sind.

Die Normierungsstudie mit rund 12.000 Schülern wurde im Jahr 2007 im Rahmen der *Trends in Mathematics and Science Study* (TIMSS) durchgeführt mit der Zielsetzung, nationale Skalen und Kompetenzstufenmodelle bereitzustellen. Dies wird in der Tat im Jahr 2008 der Fall sein (zu weiteren Details vgl. IQB im Druck a, b). Für Mathematik in der Sekundarstufe I werden die nationalen Skalen mit den dazu gehörigen Kompetenzstufenmodellen ebenfalls im Jahr 2008 vorliegen, in den Sprachen Deutsch, Englisch und Französisch werden sie 2009 folgen.

## 2.2 *Qualitätssicherung auf der Basis der Bildungsstandards – die Plöner Beschlüsse der KMK vom Juni 2006*

Die 16 Länder haben im Juni 2006 eine Gesamtstrategie der Qualitätssicherung im Bildungssystem verabschiedet, in der neben der nationalen Bildungsberichterstattung (vgl. Konsortium Bildungsberichterstattung 2006) Schulleistungsvergleiche auf der Basis der Bildungsstandards zukünftig eine herausragende Stellung einnehmen sollen. Hierauf wird im Folgenden eingegangen.

### *Stichprobenbasierte internationale und nationale Vergleiche:*

Bereits in den „Konstanzer Beschlüssen“ von 1997 hatte die KMK festgelegt, dass sich Deutschland regelmäßig an internationalen Schulleistungsstudien beteiligt, um auf Systemebene Informationen über Erträge schulischer Bildungsprozesse zu gewinnen. Die Folge war die Beteiligung an der PISA-Studie der OECD und an PIRLS/IGLU der *International Association for the Evaluation of Educational Achievement* (IEA). Die in beiden Programmen realisierten internationalen Vergleiche sollen Optimierungs- bzw. Steuerbedarf im allgemeinbildenden System Deutschlands aufdecken.

Mit dem Vergleich zwischen den Bundesländern, wie er erstmalig im Rahmen von PISA-E 2000 realisiert worden ist (vgl. Deutsches PISA-Konsortium 2002), ist dieses Systemmonitoring um eine nationale Komponente ergänzt worden. Der Ländervergleich trägt dem föderalen System Rechnung und schafft Transparenz hinsichtlich der Bildungserträge in den 16 Ländern der Bundesrepublik Deutschland. Auf der Basis der international entwickelten standardisierten Tests für die Bereiche Mathematik, Lesen und Naturwissenschaften werden Leistungsvergleiche durchgeführt, die teilweise um nationale Instrumente erweitert werden. Die Ländervergleiche wurden in PISA 2003 und PISA 2006 wiederholt.

So wichtig der auf internationalen Instrumenten beruhende Ländervergleich ist, weist er doch auch Beschränkungen auf (vgl. Köller 2007):

- Durch die internationalen Vorgaben bei der Testentwicklung können die Besonderheiten der Fächer in den einzelnen Teilnehmerstaaten kaum berücksichtigt werden.
- Die fremdsprachlichen Kompetenzen bleiben bei dieser Strategie bislang unberücksichtigt.
- Es gibt in den Naturwissenschaften keine Auffächerung in die drei Domänen Physik, Chemie und Biologie.
- Im Bereich der deutschen Sprache beschränken sich die internationalen Tests auf den Bereich Leseverständnis, weitere zentrale Teilkompetenzen (Hörverstehen, Schreiben, Sprachbetrachtung) werden ausgeklammert.

Unter anderem als Folge dieser Beschränkungen hat die KMK in ihren „Plöner Beschlüssen“ vom Juni 2006 eine Gesamtstrategie zur Qualitätssicherung im allgemeinbildenden Schulsystem vorgelegt, in der das nationale Bildungsmonitoring auf der Grundlage der Bildungsstandards durchgeführt werden soll. Auf der Basis von landesweit repräsentativen Stichproben soll in den Ländern festgestellt werden, welche Anteile der Schülerpopulation die Standards erreichen bzw. überschreiten. Das entsprechende Programm ist in Tabelle 2 im Überblick wiedergegeben.

<b>Tab 2: Internationale Vergleiche und Ländervergleiche auf der Basis der Bildungsstandards in den nächsten Jahren als Folge der Plöner Beschlüsse der KMK</b>			
<i>Jahr</i>	<i>Stufe</i>	<i>Internationale Studie</i>	<i>Fächer im Ländervergleich</i>
2009	Sekundarstufe I	PISA	Deutsch, Englisch, Französisch
2011	Primarstufe	IGLU/TIMSS	Deutsch, Mathematik
2012	Sekundarstufe I	PISA	Mathematik, Biologie, Chemie, Physik
2015	Sekundarstufe I	PISA	Deutsch, Englisch, Französisch
2016	Primarstufe	IGLU	Deutsch, Mathematik
2018	Sekundarstufe I	PISA	Mathematik, Biologie, Chemie, Physik

PISA: Programme for International Student Assessment; IGLU: Internationale Grundschul-Lese-Untersuchung; TIMSS; Trends in Mathematics and Science Study

Die Beschlüsse sehen im Grundschulbereich vor, dass die Überprüfung der Bildungsstandards im Fünf-Jahres-Rhythmus in der 3. Jahrgangsstufe geschehen soll, und zwar zeitlich gekoppelt an IGLU/PIRLS (vgl. Bos u.a. 2003), welche ebenfalls einem Fünf-Jahres-Rhythmus folgt. Neben diesem nationalen Monitoring wird sich Deutschland weiterhin an den internationalen Studien (PIRLS sowie TIMSS) beteiligen. Letztmalig in IGLU/PIRLS 2006 wurde der Ländervergleich auf der Basis internationaler Instrumente durchgeführt, der Bericht zu den Länderunterschieden wird Ende 2008 vorliegen. Erstmals in 2011 mit der nächsten PIRLS- und TIMSS-Erhebung wird es dann auf der Basis der Bildungsstandards zu einem Ländervergleich in den Fächern Deutsch und

Mathematik kommen. Dies wiederholt sich dann passend zum IGLU/PIRLS-Rhythmus alle fünf Jahre.

Im Bereich der Sekundarstufe I sehen die Plöner Beschlüsse vor, dass letztmalig 2006 der nationale Vergleich zwischen den Ländern auf der Basis der PISA-Instrumente erfolgte. Auch hier wird im Spätherbst 2008 der Bericht über Länderunterschiede erscheinen. Ab PISA 2009 wird dann der nationale Vergleich zwischen den Ländern auf Grundlage der Bildungsstandards durchgeführt, in 2009 für die Sprachen Deutsch, Englisch und Französisch, 2012 folgt der Vergleich in Mathematik und den Naturwissenschaften, 2015 folgen wieder die Sprachen, 2018 Mathematik und die Naturwissenschaften. Die Ländervergleiche in der Sekundarstufe I passen sich damit einem sechsjährigen Rhythmus an, getrennt für die Sprachen und Mathematik und die Naturwissenschaften. Berücksichtigt werden 8. Jahrgangsstufen (Bildungsgänge, die zum Hauptschulabschluss führen) und 9. Jahrgangsstufen (Bildungsgänge, die zum Mittleren Schulabschluss führen). Die für die ländervergleichenden Studien notwendigen Itempools werden aktuell unter der Ägide des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) aufgebaut, die entsprechenden Schritte wurden bereits beschrieben. Zentrale Vorarbeiten für das Fach Mathematik in der Sekundarstufe I wurden durch das Deutsche PISA-Konsortium geleistet (vgl. Ehmke u.a. 2006).

Der Gewinn gegenüber den bisherigen internationalen Studien wird vor allem auch darin liegen, dass in den Sprachen keine Beschränkung mehr auf das Leseverständnis stattfindet, sondern die Bereiche freies Schreiben und Hören (Deutsch, Englisch, Französisch) sowie Sprachbetrachtung und Orthografie (Deutsch) berücksichtigt werden. Inwieweit es möglich sein wird, mündliche Sprachproduktion und (in den Fremdsprachen) interkulturellen Kompetenzen im Large-Scale-Assessment zu testen, ist Gegenstand der aktuellen Forschungsarbeiten am IQB (vgl. Köller 2005).

#### *Flächendeckende Vergleichsarbeiten auf der Basis der Bildungsstandards:*

Über die Ländervergleiche hinaus hat sich in allen 16 Ländern Deutschlands ein System etabliert, in dem zu festen Zeitpunkten im Jahr ausgewählte Jahrgänge in Kernfächern in einer Gesamterhebung getestet werden. Dies gilt für den Grundschulbereich ebenso wie für die Sekundarstufe I. Das hier prominenteste Projekt ist ohne Frage „Vergleichsarbeiten in der Grundschule“ (VERA; vgl. Hosenfeld/Groß Ophoff/Bittins 2006; Helmke u.a. 2005), die mittlerweile für alle 16 Länder in den Fächern Mathematik und Deutsch durch die Universität Koblenz-Landau koordiniert und zukünftig am Ende der 3. Jahrgangsstufe durchgeführt wird. Im Fall von VERA ist es gelungen, länderübergreifend die Erträge schulischer Lehr-Lernprozesse in der Grundschule zu testen. Systematische Rückmeldungen gekoppelt mit didaktischen Kommentaren zu den Aufgaben sollen zur Unterrichtsentwicklung und Förderung der Schülerinnen und Schüler beitragen, wobei vielfach offen bleibt, wie dies genau geschehen soll.

Weit weniger einheitlich stellte sich bislang die Situation von flächendeckenden Vergleichsarbeiten in der Sekundarstufe I dar. Je nach Land wurden unterschiedliche Fächer in variierenden Jahrgängen und Schulformen (Bildungsgängen) einmal im Jahr überprüft. In den letzten Jahren ist es teilweise zu Länderverbänden gekommen, bei

denen die Aufgabenentwicklung zentralisiert, die jeweilige Durchführung aber in den Händen der einzelnen Länder liegt. Beispiele hierfür sind die länderübergreifenden Vergleichsarbeiten in den Fächern Deutsch, Mathematik und Englisch in der 6. Jahrgangsstufe (vgl. [www.vera.lernnetz.de](http://www.vera.lernnetz.de)), an denen sich sieben Länder beteiligen sowie die Lernstandserhebungen in der 8. Jahrgangsstufe im Fach Mathematik, an denen sich im Jahr 2008 12 Länder beteiligt haben. Inwieweit Schulen hier verpflichtet sind, an den Vergleichsarbeiten teilzunehmen, variiert je nach Land.

Die eingesetzten Tests sollen sich an den Standards bzw. daran angelehnten Lehrplänen orientieren. Die Items sollen hinsichtlich der jeweils getesteten Kompetenzen spezifiziert werden und flankierende didaktische Kommentare sollen analog zum Vorgehen in der Grundschule (s.o.) zur Unterrichtsentwicklung beitragen, wobei auch hier offen bleibt, wie dies genau geschehen soll. Hinsichtlich ihrer inhaltlichen und psychometrischen Güte variieren die Vergleichsarbeiten/Lernstandserhebungen erheblich. Beispielsweise werden in den Vergleichsarbeiten in der Grundschule nur pilotierte Aufgaben, deren psychometrische Kriterien bekannt sind, verwendet. In anderen Programmen (z.B. Orientierungsarbeiten in Sachsen) wird auf die Pilotierung der Aufgaben zugunsten größerer Unterrichtsnähe verzichtet. Gemeinsam ist allen diesen Testuntersuchungen, dass sie – im Gegensatz zu den internationalen Studien und den vom IQB verantworteten Ländervergleichen – nicht durch eigens ausgebildetes Personal administriert und ausgewertet werden, sondern vor Ort durch die Lehrkräfte der jeweiligen Schule. Anders wären flächendeckende Erhebungen gar nicht realisierbar. Dieses Vorgehen entspricht der Funktion der Vergleichsarbeiten als einer Art formativer Evaluation, impliziert aber, dass wissenschaftlich seriöse Vergleiche zwischen Schulen oder gar Ländern auf dieser Basis nicht vorgenommen werden können.

Das Kernproblem aller flächendeckenden Vergleichsarbeiten besteht darin, dass sie bislang keinen expliziten Bezug zu den Bildungsstandards in dem Sinne haben, dass auf einem nationalen Maßstab, der durch die operationalisierten Standards vorgegeben ist, Ergebnisse an die Schulen zurückgemeldet werden. Dies ist bislang unmöglich, da die entsprechenden Skalen und Kompetenzmodelle durch das IQB noch nicht bereitgestellt werden können. Stattdessen behilft man sich mit rudimentären Kompetenzstufenmodellen (im Falle der Vergleichsarbeiten in der Grundschule vgl. Helmke u.a. 2005) oder mit sozialen Vergleichen, in denen die einzelne Schule/Klasse am Landesdurchschnitt gemessen wird. Teilweise wird dabei versucht, die Eingangsselektivität der Schülerschaft zu korrigieren (sogenannte faire Vergleiche; vgl. Helmke u.a. 2005).

Mit den Plöner Beschlüssen der KMK von 2006 ist versucht worden, hier Abhilfe zu schaffen und die flächendeckenden Vergleichsarbeiten an die Bildungsstandards anzubinden. Die Beschlüsse sehen vor, dass ab 2010, wenn entsprechende normierte Items und die dazu gehörigen Kompetenzstufenmodelle auf der Basis der Bildungsstandards vorliegen, in allen Ländern flächendeckende Vergleichsarbeiten durchgeführt werden, die eine explizite Verortung der Leistungen auf den nationalen Skalen ermöglichen. In der Grundschule sollen die Vergleichsarbeiten am Ende der 3. Jahrgangsstufe in Deutsch und Mathematik stattfinden, in der Sekundarstufe I sollen sie am Ende der 8. Jahrgangsstufen in den Fächern Deutsch, Mathematik und erste Fremdsprache realisiert

werden. Damit sollen Schulen und Klassen die Gelegenheit erhalten zu überprüfen, wo die eigenen Schülerinnen und Schüler im Hinblick auf die Erreichung der in den Standards definierten Ziele stehen.

### **3. Herausforderungen und Risiken der Qualitätssicherung auf der Basis der Bildungsstandards**

Mit den Bildungsstandards und dem damit verbundenen System der stichprobenbasierten und flächendeckenden Leistungserhebungen ist die Hoffnung verbunden, sowohl auf Systemebene als auch auf der Ebene der Einzelschule oder Einzelklasse Informationen zur Verbesserung von institutionalisierten Bildungsprozessen zu gewinnen. Diese Hoffnung hat in der Vergangenheit insbesondere die Vergleichsarbeiten in der Grundschule und Sekundarstufe I getragen, ohne dass es allerdings bislang Evidenz dafür gibt, wonach sich erkennbare Veränderungen im Unterrichtsgeschehen als direkte Folge der Vergleichsarbeiten ergeben haben. Im Folgenden sollen daher Herausforderungen im Bereich der Professionalisierung von Lehrkräften kurz skizziert werden. Daran anschließend werden Probleme skizziert, die sich ergeben, wenn Bildungsstandards in ein Testsystem münden, das an den Schulen als Bedrohung wahrgenommen werden kann.

#### *3.1 Herausforderungen für die Schul- und Unterrichtsentwicklung*

In ihrer Expertise haben Klieme u.a. (2003) immer wieder betont, dass Bildungsstandards nur dann zur Unterrichtsentwicklung beitragen können, wenn auf ihrer Basis neben einem System der Qualitätssicherung auch eines der Qualitätsentwicklung im Bildungssystem etabliert wird. Der Glaube, allein die Testung würde zu einer stetigen Unterrichtsentwicklung führen – dieser Glaube legitimiert im Übrigen aktuell in vielen Ländern die Durchführung von Vergleichsarbeiten – ist unrealistisch, vielmehr können Standards und das an sie angebundene System der flächendeckenden Vergleichsarbeiten und stichprobenbasierten *Assessments* nur dann Wirkung entfalten, wenn sie durch entsprechende Unterstützungssysteme flankiert werden (vgl. z.B. Hosenfeld u.a. 2006). Und in der Tat sind mit den Bildungsstandards weit reichende Konsequenzen für die Etablierung einer neuen Aufgaben- und Unterrichtskultur verbunden. Blum (2006) schreibt hierzu: „Nur ein Unterricht, der den eigenaktiven Erwerb von Kompetenzen in lernförderlicher Arbeitsatmosphäre in den Mittelpunkt aller Lehr-/Lernanstrengungen stellt, wird Lernenden überhaupt die Chance bieten, die in den Standards formulierten Kompetenzerwartungen auch tatsächlich zu erfüllen ... Etwas konkreter bedeutet ‚*standardorientiertes Unterrichten*‘: Jede einzelne Unterrichtsstunde und jede Unterrichtseinheit muss sich daran messen lassen, inwieweit sie zur Weiterentwicklung inhaltsbezogener und allgemeiner Schüler-Kompetenzen beiträgt... Die wichtigste Frage ist nicht ‚Was haben wir durchgenommen?‘, sondern ‚Welche Vorstellungen, Fähigkeiten und Einstellungen sind entwickelt worden?‘“ (ebd. S. 15ff.). Die Entwicklung von Vorstel-

lungen, Fähigkeiten und Einstellungen ist dabei unmittelbar an den intelligenten Einsatz von Aufgaben im Unterricht gebunden. Das Erkennen, welches Potenzial eine Aufgabe beim Kompetenzaufbau, bei der Konsolidierung von Kompetenzen und beim Evaluieren von Kompetenzen hat, stellt eine zentrale professionelle Aufgabe für die Fachgruppen an den Schulen dar. Sie lässt sich eigentlich nur kooperativ erfolgreich lösen, indem Lehrkräfte, beispielsweise in den Fachkonferenzen gemeinsam über kompetenzorientierte Aufgaben und deren Einsatz im Unterricht reflektieren.

Im Sinne dieser Kompetenzorientierung ist in Abbildung 2 eine Mathematikaufgabe illustriert, bei der die allgemeine mathematische Kompetenz „Modellieren“ im Vordergrund steht (vgl. Leiß/Blum 2006):

## Tanken

Herr Stein wohnt in Trier, 20 km von der Grenze zu Luxemburg entfernt. Er fährt mit seinem VW Golf zum Tanken nach Luxemburg, wo sich direkt hinter der Grenze eine Tankstelle befindet. Dort kostet der Liter Benzin nur 1,05 €, im Gegensatz zu 1,30 € in Trier. Lohnt sich die Fahrt für Herrn Stein? Begründe deine Antwort.



Abb 2: *Beispielaufgabe zum kompetenzorientierten Unterricht im Fach Mathematik: Allgemeine Kompetenz: Modellieren (vgl. Leiß & Blum, 2006, S. 42)*

Fünf Schritte führen bei dieser Aufgabe zur richtigen Lösung (s. ausführlicher in Leiß/Blum 2006, S. 42): (1) Zunächst muss die Problemsituation verstanden werden, (2) anschließend muss das Problem strukturiert und präzisiert werden. (3) Es folgt die Mathematisierung, d.h. die Transformation in mathematische Terme, die (4) ausgerechnet werden müssen. (5) Schließlich muss das Ergebnis interpretiert und überprüft werden. Deutlich wird, dass es sich hierbei um eine anwendungsbezogene, das mathematische Verständnis fördernde Aufgabe handelt. Um diese Chancen nutzen können, sind verschiedene Schritte der Lehrerprofessionalität notwendig. Hierzu zählen:

Eigene Professionalisierung der Lehrkräfte im Hinblick auf die Philosophie und die Inhalte der Standards. Hier liegen mittlerweile vielfältige Materialien vor (vgl. z.B. Artelt/Riecke-Baulecke 2004; Blum u.a. 2006), die ergänzt werden müssen durch entsprechende Fortbildungsveranstaltungen, was im Übrigen in vielen Ländern auch der Fall ist.

Flächendeckende Angebote der dezentralen Lehrerfortbildung, wie sie im Rahmen des Modellversuchs Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts (SINUS) bzw. SINUS-Transfer umgesetzt wurden (vgl. Ostermeyer u.a. 2004). In SINUS/SINUS-Transfer wurden kooperative Arbeitsformen entwickelt und

besonderer Wert auf eine kognitiv aktivierende, kompetenzorientierte Unterrichts gelegt. Am SINUS-Transfer-Projekt, das seit August 2004 auch in Grundschulen durchgeführt wird, beteiligen sich immerhin 13 Länder, sodass für die meisten Schulleitungen auch Ansprechpartnerinnen und Ansprechpartner im jeweiligen Land vorliegen.

Forcierung der Auseinandersetzung in den Fachgruppen/Fachkonferenzen mit den Bildungsstandards. Die Standards werden nur dort in den Unterricht einfließen, wo permanent die Beschäftigung mit der Kompetenzorientierung vom Kollegium eingefordert wird. Auch dies muss in ein System der dezentralen Professionalisierung eingebunden sein.

Aufforderung an die Lehrkräfte, Eltern kompetent über die Standards und die damit verbundenen Implikationen für Qualitätssicherung und -entwicklung zu informieren. Mit den Bildungsstandards ist die explizite Aufforderung verbunden, Rechenschaft über das Erreichte abzulegen. Dies soll zukünftig in flächendeckende Vergleichsarbeiten münden, welche sich an den Standards orientieren. Solche Vergleichsarbeiten sorgen jedoch häufig auf Lehrkraft- wie Elternseite zur Verunsicherung. Eine Einbettung dieser Arbeiten in die Philosophie der Qualitätssicherung auf der Basis von Bildungsstandards muss dementsprechend in die Elternschaft kommuniziert werden, damit über die Sinnhaftigkeit des Unterfangens reflektiert werden kann.

Gelingt die breite Umsetzung dieser Maßnahmen, so steigen die Chancen, dass die mit den Bildungsstandards gekoppelten Visionen für einen erfolgreichen Unterricht in den Schulen Realität werden. Bildungsstandards werden aber weitgehend wirkungslos bleiben, wenn der Glaube, allein ihre Überprüfung steigere die Qualität von Unterricht, fortwährt.

### 3.2 *Risiken*

In diesem abschließenden Kapitel soll noch auf ein erhebliches Risiko aufmerksam gemacht werden, das bereits in der Expertise von Klieme u.a. (2003) zur Einführung von Bildungsstandards aufgegriffen wurde und dort zu dem expliziten Hinweis führte, dass auf der Basis der Bildungsstandards kein breites System des *High Stakes Testings* etabliert werden sollte. Die Folgen eines solchen Systems könnten zum Phänomen der so genannten *Inflating Gain Scores* (vgl. u.a. Green/Winters/Forster 2003; Koretz 2002) führen, die bei Vergleichsarbeiten/Lernstandserhebungen in den USA auftreten. Bei diesem Phänomen steigen die Leistungen in den flächendeckenden Vergleichsarbeiten innerhalb der jeweiligen Bundesstaaten von Jahr zu Jahr an, ohne dass sich auch entsprechende Zuwächse im stichprobenbasierten *National Assessment of Educational Progress* (NAEP) zeigen. Wo *Score Inflation* auftritt, muss die Validität der Vergleichsarbeiten/Lernstandserhebungen in Zweifel gezogen werden.

Mit der Unterzeichnung des *No-Child-Left-Behind*-Gesetzes im Jahre 2001 ist in den USA eine Grundlage geschaffen worden, die es erlaubt Schulen auf Grund der Leistungen ihrer Schülerinnen und Schüler in flächendeckenden Vergleichsarbeiten zu belohnen oder zu bestrafen. In diesem System des so genannten *High Stakes Testings* ist es

demnach für Schulen existenziell, erfolgreich in den jeweiligen Testungen abzuschneiden. In den USA hat dies (a) zu einer Welle des *Teaching to the Test* und zu Betrugsversuchen geführt (vgl. Green u.a. 2003) und (b) zu einer systematischen Erforschung der Effekte, die solche Maßnahmen haben, bei denen Schulen die Ergebnisse weniger für die eigene professionelle Entwicklung nutzen können, sondern vielmehr Sanktionen erwarten müssen.

Sehr prominent sind hier die Arbeiten von Koretz (unter anderem 2002), in denen die Leistungen in *High Stakes Tests* mit denen im stichprobenbasierten NAEP (*Low Stakes Test*), das keine Rückwirkungen auf die Einzelschule hat, auf Bundesstaatenebene verglichen werden. Sofern die erfassten Testleitungen realistische Abbilder der erreichten Kompetenzniveaus auf Seiten der Schülerinnen und Schüler sind, sollten sich (a) vergleichbare Trends und (b) hohe Korrelationen auf Schulebene zwischen den Leistungen in *Low* und *High Stakes Tests* ergeben. Dagegen sollten sich deutliche Abweichungen ergeben, d.h. divergierende Trends und geringe Korrelationen, wenn es verschiedentlich in Schulen zu Praktiken kommt, mit denen versucht wird, die Werte der Schülerinnen und Schüler in *High Stakes Tests* zu maximieren. Koretz und Barron (1998) haben zur Illustration dieser Effekte Trendanalysen für Kentucky aus dem *Kentucky Instructional Results Information System (KIRIS) Testing Program (High Stakes Tests)* mit solchen aus dem NAEP (*Low Stakes Tests*) verglichen. Berücksichtigt wurden die Kompetenzentwicklungen von Viert- und Achtklässlern in Mathematik zwischen 1992 und 1996. In KIRIS ergaben sich im untersuchten Zeitraum Leistungszuwächse, die in der 4. Jahrgangsstufe einer Effektstärke von  $d = .61$  entsprachen, in der 8. Jahrgangsstufe einer Effektstärke von  $d = .52$ . In NAEP lagen die entsprechenden Effektstärken bei  $d = .17$  und  $d = .13$ .

Green u.a. (2003) haben in verschiedenen US-Bundesstaaten auf Schulebene die Korrelationen zwischen den Leistungstrends in *Low* und *High Stakes Tests* analysiert. Im optimalen Szenario, in dem es auf Seiten der Schulen differenzielle Fördereffekte gäbe, sollten diese Korrelationen nahe 1 liegen, d.h. erfolgreiche Schulen sollten sowohl in *Low* wie in *High Stakes Tests* günstige Trends aufweisen. Die Analysen von Green u.a. (2003) ergaben je nach Bundesstaat und Distrikt sehr unterschiedliche Ergebnisse. Während in Florida die Trends in *Low* und *High Stakes Tests* sehr hoch korrelierten ( $r = .71$ ), lag die Korrelation in Virginia nur bei  $r = .17$ , im Chicago School Distrikt ergab sich gar eine Nullkorrelation, d.h. die Leistungstrends in *Low* und *High Stakes Tests* waren vollständig entkoppelt.

Koretz (2002) nennt verschiedene Ursachen, welche solche unerwarteten Befunde, die jegliche Validität von *Assessment*-Programmen in Frage stellen, erklären können:

*Umwidmung von Unterrichtszeit und -material (Teaching to the Test)*: Um Schülerinnen und Schüler auf die *High Stakes* Testung optimal vorzubereiten, werden im Unterricht primär Übungssituationen hergestellt, in denen das Aufgabenmaterial eine hohe Ähnlichkeit zu den im Test verwendeten Aufgaben aufweist. Darüber hinaus kann es auch zu einer systematischen Erhöhung der Unterrichtsstunden genau in den Fächern kommen, für die in absehbarer Zeit eine *High Stakes* Testung ansteht. Stecher und Barron (1999) untersuchten im Bundesstaat Kentucky Unterrichtszeiten in verschiedenen

Fächern und Jahrgangsstufen. Dabei zeigte sich, dass viele Schulen in der vierten Jahrgangsstufe, in der die Naturwissenschaften im *High Stakes Testing* überprüft wurden, die Wochenstundenzahlen in den Naturwissenschaften deutlich erhöhten (5.2 Wochenstunden verglichen zu 3.5 Wochenstunden in der 5. Jahrgangsstufe). In Mathematik, wo das *High Stakes Testing* in der 5. Jahrgangsstufe durchgeführt wurde, stieg die Zahl der Wochenstunden von 4.9 (Klasse 4) auf 6.4 (Klasse 5). Bei diesem Vorgehen ist zu erwarten, dass die Schülerinnen und Schüler in der Tat in den Tests höhere Kompetenzniveaus zeigen, diese aber nicht nachhaltig eingehalten werden, da Zeit und Material nach dem *High Stakes Testing* wieder zurückgefahren werden.

*Betrug (Cheating)*: In den USA besteht mittlerweile hinreichend Evidenz (vgl. Koretz 2002), wonach Lehrerinnen und Lehrer in *High Stakes Testings* Einfluss auf die Testbearbeitung nehmen, in dem sie (a) Lösungen vorsagen, (b) Lösungen verbessern, (c) Aufgaben, die später im Test auftauchen, üben lassen und (d) Aufgaben umformulieren, um die Lösungen zu erleichtern.

*Coaching*: Hierunter versteht man eine Anpassung des Unterrichts an die kognitiven Anforderungen der Aufgaben in *Assessment*-Programmen. Dies ist etwas anderes als bloßes *Teaching to the Test*, will es doch Kompetenzen fördern und kann daher durchaus wünschenswert sein. Bezogen beispielsweise auf die Bildungsstandards in der ersten Fremdsprache ist es sicherlich zu begrüßen, wenn mehr Unterrichtszeit für kommunikative Fertigkeiten in den Bereichen des Hörens und Sprechens reserviert wird. Gleichzeitig wächst dadurch die Gefahr, dass Bereiche, die später nicht Gegenstand der Testung sind, vernachlässigt werden.

Ohne Frage stellen solche Effekte Systeme der testbezogenen Qualitätssicherung in Frage und wann immer man plant, Maßnahmen des *High Stakes Testing* zu etablieren, wird man mit entsprechenden Nebenwirkungen rechnen müssen. In Deutschland beginnen einige Länder wie Nordrhein-Westfalen und Bayern damit, die Schulen, die in den flächendeckenden Vergleichsarbeiten am besten abgeschnitten haben, zu prämiieren bzw. zu veröffentlichen. Andere Länder (Brandenburg) lassen die Ergebnisse aus Vergleichsarbeiten mit einem erheblichen Gewicht in die Halbjahreszeugnisnote einfließen. Solche Maßnahmen erhöhen vermutlich die Chance für das Auftreten von *Inflating Gain Scores* und es stellt sich die Frage, ob man mit der Etablierung eines Qualitätssicherungssystems zur Förderung von Kompetenzen auf Seiten der Schülerinnen und Schüler beitragen möchte oder vor Ort in den Schulen eine Kultur der Testwertmaximierung aufbauen möchte.

#### 4. Abschluss

Auf Grundlage der in 2003 und 2004 verabschiedeten Bildungsstandards ist in Deutschland eine Gesamtstrategie der Qualitätssicherung im allgemeinbildenden Schulsystem entwickelt worden, die aktuell umgesetzt wird, indem Aufgabenpools generiert werden, welche stichprobenbasierte und flächendeckende Assessments in Schulen der Primarstufe und Sekundarstufe I erlauben werden. Längerfristig wird die gymnasiale Oberstufe in dieses System der Qualitätssicherung eingebunden werden. Ein qualitätssicherndes System im Bildungswesen auf der Basis von Bildungsstandards, dies sollte deutlich geworden sein, führt nicht automatisch zu einer Steigerung der Unterrichtsqualität und als deren Folge zu höheren Lernerträgen auf Seiten der Schülerinnen und Schüler. Vielmehr müssen Maßnahmen der Qualitätssicherung durch Unterstützungsmaßnahmen flankiert werden. Dies umfasst die dezentrale Professionalisierung von Schulleitungen und Lehrkräften auf der Basis der Standards ebenso wie zusätzliche, keineswegs kostenneutrale Angebote für benachteiligte Kinder und Jugendliche. Will man Bildungsstandards und die daran angebundene Qualitätssicherung für die Schul- und Unterrichtsentwicklung konstruktiv nutzen, so sind zum jetzigen Zeitpunkt alle Maßnahmen in Richtung *High Stakes Testing* eher kontraindiziert. Wünschenswert ist es vielmehr, Schulen Instrumente der Selbstevaluation an die Hand zu geben, mit denen sie standardbasiert Leistungsstände ihrer Schülerinnen und Schüler feststellen und bei Bedarf Fördermaßnahmen einleiten können.

#### Literatur:

- Artelt, C./Riecke-Baulecke, T. (2004): Bildungsstandards: Fakten, Hintergründe, Praxistipps. München: Oldenbourg.
- Blum, W. (2006): Bildungsstandards Mathematik: Einführung. In: Blum, W./Drüke-Noe, C./Hartung, R./Köller, O. (Hrsg.): Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen. Berlin: Cornelsen-Scriptor, S. 14–32.
- Blum, W./Drüke-Noe, C./Hartung, R./Köller, O. (Hrsg.) (2006): Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsideen und Fortbildungsmöglichkeiten. Berlin: Cornelsen-Scriptor.
- Bos, W./Lankes, E.-M./Prenzel, M./Schwippert, K./Walther, G./Valtin, R. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster: Waxmann.
- Bos, W./Hornberg, S./Arnold, K.-H./Faust, G./Fried, L./Lankes, E.-M./Schwippert, K./Valtin, R. (Hrsg.) (2007): IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster: Waxmann.
- Bremerich-Vos, A. (2003): Bildungsstandards, Kompetenzstufen, Kernlehrpläne, Parallel- und Vergleichsarbeiten. In: Deutschunterricht 5, S. 4–10.
- Deutsches PISA-Konsortium (Hrsg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.
- Deutsches PISA-Konsortium (Hrsg.) (2002): PISA 2000. Die Länder der Bundesrepublik Deutschland im Vergleich. Opladen: Leske + Budrich.

- Deutsches PISA-Konsortium (Hrsg.) (2007): PISA 2006: Der Bildungsstand der Jugendlichen in Deutschland – Die Ergebnisse der dritten internationalen Vergleichsstudie. Münster: Waxmann.
- Ehmke, T./Leiß, D./Blum, W./Prenzel, M. (2006): Entwicklung von Testverfahren für die Bildungsstandards Mathematik. In: *Unterrichtswissenschaft* 34, S. 220–238.
- Green, J.P./Winters, M.A./Forster, G. (2003): Testing high stakes tests: Can we believe the results of accountability tests? Civic Report. New York: Center for Civic Innovation.
- Hartig, J. (2007): Skalierung und Definition von Kompetenzniveaus. In: Beck, B./Klieme, E. (Hrsg.): *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Ergebnisse Band 1*. Weinheim: Beltz, S. 83–99.
- Helmke, A./Hosenfeld, I./Ophoff, J. G./Hochweber, J./Isaac, K./Koch, U./Scherthan, F./Wagner, S. (2005): VERA 2004: Erste Ergebnisse des Ländervergleichs. Landau: Universität Koblenz-Landau.
- Hosenfeld, I./Groß-Ophoff, J./Bittins, P. (2006): Vergleichsarbeiten und Schulentwicklung. *Schulmanagement-Handbuch Band 118*. München: Oldenbourg.
- IQB (im Druck a): Bildungsstandards: Kompetenzen überprüfen: Grundschule Mathematik 3./4. Schuljahr. Berlin: Cornelsen.
- IQB (im Druck b): Bildungsstandards: Kompetenzen überprüfen: Grundschule Deutsch 3./4. Schuljahr. Berlin: Cornelsen.
- Klieme, E./Avenarius, H./Blum, W./Döbrich, P./Gruber, H./Prenzel, M./Reiss, K./Riquarts, K./Rost, J./Tenorth, H.E./Vollmer, J. (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Reihe Bildungsreform, Band 1. Bonn: Bundesministerium für Bildung und Forschung.
- Klieme, E./Avenarius, H./Baethge, M./Döbert, H./Hetmeier, H.-W./Meister-Scheufelen, G./Rauschenbach, T./Wolter, A. (2007): Grundkonzeption der Bildungsberichterstattung für Deutschland. In: Krüger, H.-H./Rauschenbach, T./Sander, U. (Hrsg.): *Bildungs- und Sozialberichterstattung*. 6. Beiheft der Zeitschrift für Erziehungswissenschaft. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 129–145.
- KMK (2004): Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss. München: Luchterhand.
- KMK (2006): Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. München: Luchterhand.
- Köller, O./Watermann, R./Trautwein, U./Lüdtke, O. (Hrsg.) (2004): Wege zur Hochschulreife in Baden-Württemberg. TOSCA – Eine Untersuchung an allgemeinbildenden und beruflichen Gymnasien. Opladen: Leske+Budrich.
- Köller, O. (2005): Neue Besen kehren gut: Das Institut zur Qualitätsentwicklung im Bildungswesen soll die nationalen Bildungsstandards in Deutschland überprüfen. In: *Psychologie in Erziehung und Unterricht* 52, S. 281–286.
- Köller, O. (2007): Bildungsmonitoring und Vergleichsarbeiten in Deutschland. In: Labudde, P. (Hrsg.): *Bildungsstandards am Gymnasium*. Bern: h.e.p., S. 65–73.
- Konsortium Bildungsberichterstattung (2006): *Bildung in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann.
- Koretz, D. (2002): Limitations in the use of achievement tests as measures of educators' productivity. In: *Journal of Human Resources* 37, S. 752–777.
- Koretz, D./Barron, S.I. (1998): *The Validity of Gains on the Kentucky Instructional Results Information System (KIRIS)*. Santa Monica: RAND.
- Leiß, D./Blum, W. (2006): Beschreibung zentraler mathematischer Kompetenzen. In: Blum, W./Drückeno, C./Hartung, R./Köller, O. (Hrsg.): *Bildungsstandards Mathematik: konkret*. Berlin: Cornelsen-Scriptor, S. 33–50.
- Ostermeier, C./Carstensen, C.H./Prenzel, M./Geiser, H. (2004): Kooperative unterrichtsbezogene Qualitätsentwicklung in Netzwerken. Ausgangsbedingungen für die Implementation im BLK-Modellversuchsprogramm SINUS. In: *Unterrichtswissenschaft* 32, S. 215–237.
- Rost, J. (2004): *Lehrbuch Testtheorie/Testkonstruktion*. Bern: Huber.

- Rupp, A.A./Vock, M./Harsch, C./Köller, O. (2008): Developing standards-based assessment tasks for English as a first foreign language – Context, processes, and outcomes in Germany. Münster: Waxmann
- Spinner, K.H. (2005): Der standardisierte Schüler. In: Didaktik Deutsch 18, S. 4–13.
- Stecher, B.M./Barron, S. I. (1999): Quadrennial Milepost Accountability Testing in Kentucky. CSE Technical Report No. 505. Los Angeles: Center for the Study of Evaluation, University of California.
- Weinert, F.E. (2001): Vergleichende Leistungsmessung in Schulen – Eine umstrittene Selbstverständlichkeit. In: Weinert F.E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim: Beltz, S. 17–31.

*Anschrift des Autors:*

Prof. Dr. Olaf Köller, Humboldt-Universität zu Berlin, Institut zur Qualitätsentwicklung im Bildungswesen, Unter den Linden 6, 10099 Berlin, Email: iqboffice@iqb.hu-berlin.de

Thomas Olk/Karsten Speck

# Qualität und Qualitätsentwicklung in der Sozialpädagogik

## 1. Einleitung

Seit etwa Mitte der 1990er Jahre sieht sich die sozialpädagogische Praxis<sup>1</sup> mit dem Qualitätsdiskurs konfrontiert. Übergreifenden gesellschaftlichen Trends folgend wird nun auch von der Kinder- und Jugendhilfe eine Präzisierung ihrer Wirkungsziele, eine Offenlegung ihrer „Gütekriterien“ und Vorgehensweisen sowie eine genauere Überprüfung und ein Nachweis ihrer Ergebnisse eingefordert. Inzwischen sind in einer ersten Phase vor allem im industriellen und gewerblichen Bereich entwickelte Konzepte und Verfahren des Qualitätsmanagements (wie z.B. die Normenreihe DIN EN ISO 9000ff, TQM sowie EFQM) für die Belange sozialpädagogischer Praxisfelder adaptiert sowie in einer zweiten Phase ein breites Spektrum von Verfahren und Instrumenten der Qualitätsentwicklung und des Qualitätsmanagements für spezifische Handlungsfelder sowie Einrichtungen und Dienste der sozialpädagogischen Praxis entwickelt und eingeführt worden. Begleitet wurde dieser „Qualitätsboom“ durch eine Vielzahl von Fachtagungen und Kongressen und einer mittlerweile ausufernden Publikationstätigkeit zu Themenbereichen wie „Qualität“, „Qualitätsmanagement“, „Qualitätsentwicklung“, „Qualitätssicherung“ etc. (vgl. Merchel 2005 sowie 2006). Es scheint daher lohnend, eine erste Zwischenbilanz zu ziehen und nach den Hintergründen, den Umgangsweisen der unterschiedlichen Akteure mit dem Qualitätsthema als auch nach absehbaren Auswirkungen der Einführung von Qualitätsmanagementverfahren auf das sozialpädagogische Handeln zu fragen. Allerdings stößt ein solches Anliegen auf Grenzen. Zwar hat sich im Bereich der Kinder- und Jugendhilfe durchaus eine Wirkungs- bzw. Effektivitätsforschung (vgl. Hollenstein 1993; Schmidt u.a. 2002; vgl. für einen Überblick Otto 2007) sowie im internationalen Bereich eine Implementationsforschung (vgl. Bermann 1995; Ahire/Waller/Golhar 1996 sowie Lee u.a. 2007) etabliert, es liegen aber für Deutschland kaum umfassende empirische Untersuchungen über die intendierten wie nicht-intendierten Auswirkungen der Einführung von Qualitätsmanagementverfahren in einzelnen Handlungsfeldern bzw. Einrichtungen und Diensten der Kinder- und Jugendhilfe vor. Insbesondere fehlen – von einigen Untersuchungen abgesehen – grundlegende Studien a) zu den tatsächlichen Handlungsmustern unterschiedlicher Akteursgruppen im Umgang

1 Im Interesse einer möglichst hohen Übereinstimmung zwischen dem Gegenstandsbereich der Sozialpädagogik als einer Teildisziplin der Erziehungswissenschaft und dem sozialpädagogischen Praxisfeld konzentriert sich der folgende Beitrag auf die Kinder- und Jugendhilfe – also den Geltungsbereich des SGB VIII – obwohl ähnliche Entwicklungen und Fragestellungen auch in angrenzenden Feldern der Sozialen Arbeit beobachtbar sind.

mit der Qualitätsthematik und den eingeführten Verfahren in der Praxis, b) zu den Wirkungen und dem Nutzen der Verfahren der Qualitätsentwicklung für die Klienten<sup>2</sup> bzw. Nutzer einschließlich ihrer Beteiligungschancen, c) zu den mittel- und langfristigen Auswirkungen und Nebenwirkungen von Verfahren der Qualitätsentwicklung auf das unmittelbare sozialpädagogische Handeln der Fachkräfte (vgl. dazu neuerdings Beckmann u.a. 2006 und 2007).<sup>3</sup> Der folgende Beitrag beabsichtigt daher, das vorliegende empirische Wissen zu diesem Untersuchungsgegenstand zusammenzutragen und offene Forschungsfragen zu formulieren.

Der Beitrag ist folgendermaßen aufgebaut: Zunächst werden die Auslösebedingungen und „treibenden Kräfte“ des Qualitätsbooms im Bereich der Kinder- und Jugendhilfe rekonstruiert (vgl. Abschnitt 2), sodann die veränderte Relationierung von gesellschaftlichen Erwartungen und Steuerungsinteressen gegenüber der sozialpädagogischen Praxis herausgearbeitet (vgl. Abschnitt 3), organisationale Strategien des Umgangs mit dem Qualitätsthema problematisiert (vgl. Abschnitt 4) sowie die Auswirkungen von Verfahren und Instrumenten der Qualitätsevaluation auf die Interaktion zwischen Fachkräften und Klienten analysiert (vgl. Abschnitt 5). Ein Fazit beschließt die Argumentation (vgl. Abschnitt 6).

## **2. Hintergründe und „treibende Kräfte“ der Qualitätsdebatte in der Sozialpädagogik**

Obwohl die Beschäftigung mit der „Qualität“ sozialpädagogischer Arbeit sowie die Diskussion um „Gütekriterien“ qualitativ hochwertiger sozialpädagogischer Praxis durchaus auf eine lange Tradition zurückblicken kann<sup>4</sup>, ist die Art und Weise der Qualitätsanfrage in ihrer heutigen Form dennoch neu in der Sozialpädagogik (vgl. Hornstein 2000, Merchel 2000). Es geht nun nicht länger um die Verbesserung der Qualität durch Höherprofessionalisierung und die Entwicklung fachinterner Gütekriterien und methodischer Vorgehensweisen, sondern um ökonomische Effektivitäts- und Effizienzerwartungen, Kosten-Leistungs-Relationen und quantifizierbare Erfolgsmessungen. Da es sich hierbei in erster Linie um Erwartungen von politischen und gesellschaftlichen Akteuren handelt, die losgelöst und unabhängig von sozialpädagogischen Selbstdeutungen und disziplininternen Maßstäben formuliert werden, wird diese Debatte und die hiermit verbundenen Anspruchshaltungen gegenüber der Sozialpädagogik als in hohem Maße

- 2 Aus Gründen der besseren Lesbarkeit wird nicht durchgängig eine geschlechtergerechte Sprache verwendet. Im Falle der männlichen Form ist immer auch die weibliche Form mit gemeint.
- 3 Darüber hinaus gibt es zwar Ansätze für eine grundsätzliche theoretische Auseinandersetzung mit dem Qualitätskonzept und seiner Bedeutung für das sozialpädagogische Feld (vgl. Köpp/Neumann 2003, Honig/Joos/Schreiber 2004 sowie die Beiträge in Beckmann u.a. 2004), allerdings konnten diese zumeist auf die frühkindliche Erziehung und Betreuung bezogenen Arbeiten noch keine breite Debatte in der Sozialpädagogik auslösen.
- 4 Ein prominentes Beispiel stellen die „Strukturmaximen einer modernen lebensweltorientierten Jugendhilfe“ dar (vgl. BMJFFG 1990, S. 88ff.).

von außen aufgezwungen empfunden.<sup>5</sup> Im sozialpädagogischen Fachdiskurs wird zudem befürchtet, dass die Übertragung von aus dem industriellen Bereich stammenden Verfahren und Instrumenten des Qualitätsmanagements auf die sozialpädagogische Praxis professionelle Handlungsspielräume durch Standardisierung und Formalisierung einschränken und die Partizipation der Nutzer am Prozess der Dienstleistungserbringung eher verhindern als ermöglichen könnte (vgl. z.B. Shaw 1995 sowie Engel u.a. 1998 sowie Langer 2007).

Trotz solcher skeptischen Einschätzungen und Befürchtungen konnten sich auch die sozialpädagogischen Einrichtungen und Dienste dem Qualitätsthema und den hiermit verbundenen gesellschaftlichen Erwartungen keineswegs entziehen. Vielmehr ist sowohl in der sozialpädagogischen Praxis als auch in der Fachdisziplin eine intensive Auseinandersetzung mit der Qualitätsthematik zu beobachten. Hierzu haben nicht zuletzt die im Folgenden genannten Entwicklungen und „treibenden Kräfte“ beigetragen:

- Verschärfung der Konkurrenzsituation zwischen einer wachsenden Anzahl unterschiedlicher Typen freier Träger der Kinder- und Jugendhilfe um die knapper werdenden öffentlichen Finanzmittel, die die kontinuierliche Verbesserung der Leistungsqualität bei möglichst geringem Mittelaufwand als einen „strategischen Wettbewerbsfaktor“ auf staatlich regulierten Märkten für soziale Dienstleistungen erscheinen lässt (vgl. Volkmar 1998);
- Wandel des Selbstverständnisses des Sozialstaates vom „Versorgungsstaat“ zum „Minimalstaat“ bzw. „aktivierenden Staat“, in dessen Folge Managementkonzepte und -verfahren aus dem privatwirtschaftlichen auf den öffentlichen Sektor übertragen werden, mit dem Ziel Effektivität und Effizienz öffentlicher Leistungsproduktion zu steigern (vgl. Olk 2001);
- Einführung von gesetzlichen Vorschriften zu Vereinbarungen zwischen öffentlichen Kostenträgern (Zuwendungsgebern) und freien Leistungsanbietern (Zuwendungsempfänger) nach dem Vorbild des Krankenversicherungs-, Pflegeversicherungs- und Sozialhilferechts nun auch im Bereich der Kinder- und Jugendhilfe durch die zum 01.01.1999 in Kraft getretenen neuen Vorschriften der §§ 78a-g SGB VIII;
- die kommunale Verwaltungsreform auf der Grundlage der Ideen und Prinzipien der internationalen „New Public Management“-Bewegung nach dem Muster des von der Kommunalen Gemeinschaftsstelle für Verwaltungsvereinfachung (KGSt) entwickelten „Neuen Steuerungsmodells“ (NSM) seit der ersten Hälfte der 1990er Jahre (vgl. KGSt 1994; vgl. zu einer umfassenden Evaluation dieses Reformprozesses Bogumil u.a. 2007) sowie nicht zuletzt

5 Ganz in diesem Sinne identifiziert Kuper (vgl. 2002, S. 549) die spezifische Relevanz des Qualitätsbegriffes im Bildungsdiskurs in der hiermit möglichen Distanzierung von pädagogischen Selbstdeutungen und dem Verweis auf Fremdreferenzen, über die das Bildungssystem – und analog auch das System der Kinder- und Jugendhilfe – an seine unterschiedlichen Umweltausschnitte zurückgebunden ist. Insofern repräsentiert der Qualitätsdiskurs nichts anderes als die in der Umwelt institutionalisierten Erwartungen an die Beschaffenheit der professionellen Leistungen im sozialpädagogischen Praxisfeld.

- die Bundesinitiative „Qualitätssicherung in der Kinder- und Jugendhilfe“, in deren Rahmen insgesamt 36 Broschüren in einer Auflage von jeweils 15.000 Exemplaren publiziert worden sind (vgl. BMFSFJ 2002) und mit deren Hilfe die bestehende Praxis in der Kinder- und Jugendhilfe überprüft, kritisch bewertet und durch Umschichtungen von Mitteln finanzielle Ressourcen für neue Aufgaben und innovative Handlungsformen erwirtschaftet werden sollten (vgl. kritisch Hornstein 2000, S. 133f.).

Das Aufkommen des Qualitätsthemas ist ganz offensichtlich – ähnlich wie im Bereich von Schule und Hochschule – eine Begleiterscheinung des übergreifenden Trends, im Bereich der Privatwirtschaft entwickelte betriebswirtschaftlich begründete bzw. primär kostenorientierte Konzepte und Verfahren des Managements nun auch auf die Bereiche Bildung und sozialpädagogische Dienste zu übertragen. Möglich und plausibel wird dies, weil die quantitativen Wachstumsprozesse der letzten Jahrzehnte zu einer Bedeutungsaufwertung von Ressourcenfragen und zu einem veränderten Funktions- und Selbstverständnis sowohl des Bildungs- als auch des Sozialwesens beigetragen haben. Die Einrichtungen, Dienste und Trägerorganisationen im sozialpädagogischen Praxisfeld, die bislang überwiegend unter der Perspektive ihrer weltanschaulich-milieu-gebundenen Wertorientierungen und ihrer sozialpädagogischen Fachlichkeit betrachtet wurden, müssen nun auch als Wirtschaftsbetriebe begriffen werden, deren Leitung und organisatorische Strukturierung zunehmend betriebswirtschaftliche Kompetenzen des Managements erfordert. Auch versteht sich die moderne Kinder- und Jugendhilfe spätestens seit der Einführung des Kinder- und Jugendhilfegesetzes (SGB VIII) von 1990 nicht länger als gesellschaftliche Disziplinierungs- und Kontrollagentur, sondern vielmehr als eine fachlich begründete Sozialleistung für potenziell alle Kinder, Jugendlichen und ihre Personensorgeberechtigten (vgl. Münder u.a. 2006). Die hiermit verbundene Transformation freier Träger von „Wertgemeinschaften zu Dienstleistungsunternehmen“ (vgl. Rauschenbach/Sachße/Olk 1996) und die Umwandlung der Jugendämter in „kommunale Dienstleistungsunternehmen“ hängen also insgesamt damit zusammen, dass der Bereich der Kinder- und Jugendhilfe aufgrund des seit den 1970er Jahren beobachtbaren enormen Wachstums an Personal, Organisationen, Finanzmitteln und Adressatengruppen längst aus der geschützten Nische eines quantitativ unbedeutenden „Randbezirks“ kommunaler Sozialstaatlichkeit herausgetreten und nun verstärkt dem rauen Wind öffentlicher Aufmerksamkeit und dem wirksamkeitsbezogenen Kontrollblick sozialpolitischer Akteure ausgesetzt ist.

### **3. Implementierung von Qualitätserwartungen durch Leistungs-, Entgelt- und Qualitätsentwicklungsvereinbarungen**

Im System der Kinder- und Jugendhilfe, das durch das Zusammenwirken öffentlicher mit einer Vielzahl freier Träger und einer unüberschaubaren Vielfalt unterschiedlichster Einrichtungen, Dienste und Angebotsformen sozialpädagogischer Leistungen auf loka-

ler Ebene gekennzeichnet ist, sind es insbesondere die 1999 eingeführten Leistungs-, Entgelt- und Qualitätsentwicklungsvereinbarungen nach § 78a-g, mit deren Hilfe Konzepte und Verfahren der Qualitätsentwicklung bzw. des Qualitätsmanagements in die Praxis sozialpädagogischer Einrichtungen und Diensten implementiert werden (vgl. die Beiträge in Kröger 1999). Sie sind Ausdruck eines neuen Steuerungsverständnisses, das durch einen Wechsel von der Steuerung über quantitative Inputs (Ausweitung von Finanzmitteln, Personal, Einrichtungen etc.) zu einer Steuerung qualitativer Parameter organisatorischer Strukturen und Abläufe gekennzeichnet ist (vgl. auch Merchel 2005, S. 40). Ähnlich wie im Bereich des schulisch geprägten Bildungssystems basiert dieses neue Steuerungsverständnis auf der Einsicht, dass sich Qualität nicht allein durch die Beeinflussung übergreifender Systemstrukturen, sondern letztlich nur durch die kleinteilige Gestaltung von pädagogischen Settings und konkreten Arbeitsabläufen in (sozial-) pädagogischen Institutionen beeinflussen lässt.

Bei der Steuerung über Leistungsvereinbarungen geht es also letztlich darum, weit über die üblichen Kontrollpraktiken im herkömmlichen Zuwendungsverfahren hinaus durch die präzise Beschreibung von Leistungen, die Zuordnung von Entgelten zu Leistungsbestandteilen und die Vereinbarung von Qualitätskriterien und Verfahren zu ihrer fortlaufenden Überprüfung eine detaillierte Feinsteuerung organisationsinterner Abläufe im Verhältnis zwischen öffentlichen und freien Trägern zu erreichen. Der anvisierte Umstieg von der rein quantitativen Input-Steuerung zu einer wirkungsorientierten Qualitätssteuerung ist Ausdruck eines gestiegenen Steuerungsinteresses der öffentlichen Träger gegenüber den freien Trägern der Kinder- und Jugendhilfe.

Allerdings zeigt sich bei näherem Hinsehen, dass nicht die Gütekriterien professionellen Handelns oder die Interessen- und Bedürfnisartikulationen der Klienten im Mittelpunkt von Verhandlungen und Vereinbarungen stehen, sondern in erster Linie Kennzahlen, Entgelte und das verfügbare Finanzvolumen. Gleichzeitig steigen die Anforderungen an die öffentlichen Kostenträger. Wenn sich Mittelvergabe und Verwendungsprüfung nicht mehr – wie im herkömmlichen Zuwendungsverfahren – auf den bloßen „Verwendungsnachweis“ im Sinne einer wirtschaftlichen und formal richtigen Mittelverausgabung beschränken, sondern auch anspruchsvolle Ergebnisziele einbezogen werden, dann ist nicht nur ein erhöhter zeitlicher Arbeitsaufwand, sondern auch ein anderes Beratungsverständnis und eine höhere Controllingkompetenz erforderlich (z.B. in Bezug auf dialogisches Vorgehen, Feldkenntnis vor Ort, realistische Zieldefinitionen, Fehlerfreundlichkeit), die nicht zwangsläufig in jedem kommunalen Jugendamt und jedem Landesjugendamt in ausreichendem Umfang verfügbar sind.<sup>6</sup>

Die Probleme der qualitativen Feinsteuerung sozialpädagogischen Handelns gehen auch aus vorliegenden Analysen bestehender Leistungs-, Entgelt- und Qualitätsentwick-

6 Obwohl die Einführung von Leistungsvereinbarungen ursprünglich mit Argumenten wie Kostenreduktion und „effektiveren Mitteleinsatz“ begründet wurde, ist also zumindest in der Einführungsphase, wahrscheinlich aber auch darüber hinaus, sowohl bei den betroffenen freien Trägern als auch den öffentlichen Kostenträgern mit einem Anstieg des administrativen Arbeitsaufwands und der Bindung qualifizierten Verwaltungspersonals zu rechnen (vgl. Olk/Speck/Sattler 2004 sowie Olk/Speck/Albert 2004).

lungsvereinbarungen nach §§ 78a ff. SGB VIII (vgl. Münder/Tammen 2003) und aus Rahmenverträgen der Länder nach § 78f SGB VIII hervor (vgl. Gottlieb u.a. 2003). Münder und Tammen (vgl. 2003) konstatieren, dass in den allerwenigsten Fällen die untersuchten Qualitätsentwicklungsvereinbarungen eigenständige Aussagen zur Qualitätsbewertung, -entwicklung und -gewährleistung enthalten: „Der Aspekt der *Entwicklung* der Qualität findet in den meisten Vereinbarungen keine Berücksichtigung. Sehr viele der Regelungen in diesem Regelungspunkt haben den Charakter theoretischer, abstrakter und allgemeiner Abhandlungen zum Thema Qualität mit wenig erkennbarem Bezug zur konkret vereinbarten Leistung.“ (Münder/Tammen 2003, S. 37; Hervorh. i. Org.) Zudem ließen sich in den Vereinbarungen kaum Verknüpfungen zwischen Entgeltvereinbarungen einerseits und Leistungs- und Qualitätsentwicklungsvereinbarungen andererseits feststellen: So sehen die Verträge beispielsweise in keinem Fall „ein erhöhtes Entgelt für den Fall besonders guter Leistung oder besonders hochwertiger Qualität vor“ (Münder/Tammen 2003, S. 36). Gottlieb u.a. (vgl. 2003) bestätigen die Analyseergebnisse für die übergeordneten Rahmenverträge der Länder nach § 78f SGB VIII. Demnach sind die Leistungs- und Qualitätsentwicklungsvereinbarungen deutlich weniger präzise formuliert als die Entgeltvereinbarungen. Zudem fungieren die Qualitätsentwicklungsvereinbarungen eher als Prüfvereinbarungen denn als Vereinbarungen zur Qualitätsentwicklung.

## **4. Qualitätsmanagement in sozialpädagogischen Organisationen**

### *4.1 Verfahren und Modelle des Qualitätsmanagements*

Im Hinblick auf die Umgangsweisen von Einrichtungen, Diensten und Trägerorganisationen der Kinder- und Jugendhilfe mit der Qualitätsthematik ist festzustellen, dass pragmatische Strategien dominieren (vgl. Merchel 2005, S. 41f.; Flösser/Oechler 2004). Die generelle Debatte über Möglichkeiten und Wege der Qualitätsverbesserung hatte sich rasch auf die Frage nach den für die eigene Organisation geeigneten Modellen und Verfahren des Qualitätsmanagements verengt, während grundsätzliche Fragen nach der Bestimmung der Qualität sozialpädagogischen Handelns als bereits beantwortet oder aber als irrelevant behandelt wurden. Insofern kann kaum überraschen, dass inzwischen eine kaum noch überschaubare Vielzahl an Konzepten und Modellen zur Qualitätsentwicklung, zum Qualitätsmanagement bzw. zur Qualitätssicherung angewandt und erprobt werden (vgl. Peterander/Specck 1999; Specck 2006 sowie Gerull 2000).

Gerull hat auf der Basis von über 700 Publikationen und Materialien zum Thema Qualität die Qualitätsmanagement-Systeme und -konzepte im sozialen Bereich in einem Hand- und Werkbuch zusammengetragen, systematisiert und bewertet (vgl. Gerull 2000). Unter Qualitätsmanagement wird dabei ein „Teil der Gesamtführungsaufgabe eines Unternehmens zur Entwicklung, Aufrechterhaltung und Förderung der Leistungsfähigkeit“ und unter Qualitätsmanagement-System „ein strategisches Konzept zur Umsetzung des Qualitätsmanagements“ verstanden (S. 5/Glossar). Als klassische und bun-

desweit anerkannte Qualitätsmanagement-Systeme können nach dieser Systematisierung – ähnlich wie im Bildungsbereich – das Qualitätsmanagement nach DIN EN ISO 9000ff., das Total Quality Management (TQM) sowie das Qualitätsmanagement-Modell der European Foundation für Quality (EFQM) gelten. Darüber hinaus gibt es weitere Qualitätsmanagement-Systeme, die sich durch weniger komplexe und detaillierte Konzepte und Abläufe auszeichnen und nicht vorrangig als Unternehmenskonzept für hierarchische Organisationen konzipiert sind. Auch die Wohlfahrtsverbände haben zum Teil weitere Qualitätsmanagement-Systeme und Zertifizierungssysteme entwickelt und eigene Qualitätsinstitute eröffnet.

Das Neue an der Einführung der Qualitätsmanagement-Systeme lässt sich insbesondere in den folgenden vier Aspekten zusammenfassen:

1. Während bislang vor allem die Strukturqualität im Mittelpunkt stand, treten nunmehr die Prozess- sowie vor allem die Ergebnisqualität ins Zentrum des Interesses.
2. Während es bislang üblich war, allgemeine – meist viel zu globale und kaum erreichbare – Ziele sozialpädagogischen Handelns zu formulieren, besteht nun der Anspruch, messbare Indikatoren zur Überprüfung der konkret definierten Ergebnisse bzw. Wirkungen zu entwickeln.
3. Während bislang der Nachweis einzelner Rahmenbedingungen oder Instrumente zur Aufrechterhaltung bzw. Weiterentwicklung der Qualität genügten (z.B. erforderliche Qualifikation, Fort- und Weiterbildung, Supervision), wird nunmehr gefordert, über ein umfassendes Verfahren des Qualitätsmanagements die systematische Steuerung und kontinuierliche Verbesserung der Strukturen und Rahmenbedingungen, der Abläufe des professionellen Handelns sowie der Ergebnisse nachzuweisen.
4. Während bislang allenfalls zu Beginn einer Förderung der Nachweis bestimmter Rahmenbedingungen eingefordert wurde, werden nun zunehmend a) Qualitätskriterien und Qualitätsnachweise zum Gegenstand von Aushandlungsprozessen über Finanzen erhoben und b) die Ausreichung von Finanzmitteln direkt an die Einhaltung bzw. Erreichung bestimmter Qualitätskriterien gekoppelt.

#### *4.2 Zur Implementation von Instrumenten und Verfahren des Qualitätsmanagements auf der Ebene der Einzelorganisationen*

Die Frage nach Beweggründen, Verlaufsprozessen und Auswirkungen der Einführung von Verfahren und Instrumenten des Qualitätsmanagements lenkt den Blick auf die Einzelorganisationen der sozialpädagogischen Praxis. Im Hinblick auf die prozessuale Perspektive gilt: Ob überhaupt und gegebenenfalls welche Verfahren und Instrumente des Qualitätsmanagements eingeführt werden, wie dieser Prozess binnenorganisatorisch gesteuert und umgesetzt wird und welche Auswirkungen dies auf Organisationsstruktur, -kultur und -routinen zeitigt, hängt von den konkreten Bedingungen und Akteurskonstellationen vor Ort ab. Die Einführung eines Qualitätsmanagementmodells kann als Bestandteil eines umfassenden Organisationsentwicklungsprozesses ausgestaltet wer-

den, in dessen Verlauf das reflexive Wissen der Organisation über sich selbst verbessert, eine realistische Bestandsaufnahme des eigenen Wirkens sowie eine Selbstvergewisserung über programmatische Ziele, organisatorische Ablaufroutinen und Effekte angestoßen wird. Solche Prozesse können Innovationspotenziale freilegen und Leistungsverbesserungen auf der Ebene sozialpädagogischen Handelns ermöglichen. Allerdings sind auch gegenteilige Effekte denkbar, also die hierarchische Umsetzung handlungsrestringierender Regelwerke und Kontrollinstrumentarien, die zu einer Bürokratisierung der Leistungsprozesse und zu einer Verschlechterung der Leistungsergebnisse führen. Solche alternativen Prozessverläufe machen darauf aufmerksam, dass es unterschiedliche Formen der Implementation von Modellen des Qualitätsmanagements geben kann. Es ist sowohl denkbar, dass die Leitungsebene einer Organisation die Einführung des Qualitätsmanagements als eine rein managerielle Aufgabe betrachtet, die ohne Beteiligung der unteren Leitungsebenen, Fachkräfte und Klienten top-down durchzusetzen ist. Auf der anderen Seite sind Formen der Implementation denkbar, in denen die Führungs- und Leitungsebene zwar eine besondere Verantwortung für die Gestaltung des Implementationsprozesses übernimmt, allerdings die übrigen beteiligten Gruppen in angemessener Weise an der Ausgestaltung der entsprechenden Verfahren und Instrumente beteiligt. Die Einführung von Qualitätsmanagementverfahren veranlasst also die Führungsebene in jedem Falle zur Übernahme neuer Steuerungsverantwortlichkeiten; aber die Neuverteilung von Einfluss- und Verantwortungsbereichen muss keineswegs als Nullsummenspiel gestaltet werden, sondern kann auch die Verantwortlichkeit und Partizipation sowohl des Fachpersonals als auch der Klienten an diesem Prozess stärken (vgl. auch Gerull 2004). Empirische Studien, die solche binnenorganisatorischen Implementationsverläufe detailliert rekonstruieren, liegen für den deutschsprachigen Raum bislang nicht vor. Solche Untersuchungen, die als qualitative Organisationsstudien anzulegen wären, würden die subjektiven Perspektiven, Interessen und Relevanzsetzungen der unterschiedlichen Akteursgruppen ermitteln und Prozessverläufe der Organisationsentwicklung rekonstruieren, die sich im Zusammenhang mit der Einführung formaler Qualitätsmanagementsystemen als Resultat des Zusammenwirkens strategischer Handlungen unterschiedlicher Akteursgruppen ergeben würden.

Allerdings liegt neuerdings eine empirische Studie vor, die sich mit den Auswirkungen der Einführung von Qualitätsmanagement-Systemen in Einrichtungen der Jugendhilfe auf die Arbeitsbedingungen sozialpädagogischer Fachkräfte beschäftigt (vgl. Beckmann u. a. 2006 sowie 2007). Anhand einer Befragung von 30 Führungskräften, 261 Fachkräften und 435 Familien in 30 Einrichtungen der sozialpädagogischen Familienhilfe in Nordrhein-Westfalen werden die Zusammenhänge zwischen der Einführung von Qualitätsmanagementsystemen und unterschiedlichen Formen der Formalisierung des professionalisierten sozialpädagogischen Handelns untersucht. Zumindest für die befragten Einrichtungen stellt diese Untersuchung fest, dass Arbeitsbedingungen vorherrschen, die professionelles sozialpädagogisches Handeln keineswegs einengen sondern vielmehr ermächtigen. Die Vermutung, dass die Einführung formaler Qualitätsmanagement-Systeme in jedem Falle mit Prozessen der Standardisierung und Deprofessionalisierung verbunden sei, muss in dieser Pauschalität also zurückgewiesen werden.

Allerdings gilt auch: je ausgeprägter in diesen Einrichtungen Methoden des Monitorings, der Steuerung über Misserfolgsanalysen und Leistungsindikatoren genutzt werden, umso mehr lassen sich restringierende Formen der Formalisierung des professionellen Handelns identifizieren.

Darüber hinaus liegen Studien zur quantitativen Verbreitung von Qualitätsmanagement-Systemen in Organisationen vor, die sich exklusiv oder teilweise auf den Bereich der Kinder- und Jugendhilfe beziehen. So geht etwa aus US-amerikanischen Studien hervor, dass die überwiegende Mehrheit der befragten Einrichtungen und Dienste im Bereich der Familienhilfe inzwischen Instrumente und Verfahren des Qualitätsmanagements eingeführt haben (vgl. Lee u.a. 2007). Als relevante Bedingungsfaktoren für die Einführung entsprechender Modelle und Instrumente werden insbesondere die Größe der Einrichtung, die Branche sowie die Finanzbasis der Organisation identifiziert. Als Barrieren werden – wenig überraschend, aber oft übersehen – Bedingungen wie zu starke Arbeitsbelastung und prekäre Finanzsituation hervorgehoben (vgl. Lee u.a. 2007).

Hinsichtlich der Entwicklung in Deutschland liegen ebenfalls zumeist regional begrenzt aussagefähige empirische Untersuchungen vor. So kann etwa Gerull (vgl. 2004) auf der Grundlage einer schriftlichen Befragung von 85 Einrichtungen in Nordrhein-Westfalen (einbezogen waren Jugend- und Sozialämter sowie insbesondere Einrichtungen aus dem Bereich der sozialpädagogischen Erziehungshilfen) zeigen, dass rund zwei Drittel der befragten Einrichtungen über zumindest einzelne Instrumente bzw. Verfahren des Qualitätsmanagements verfügen und dass ein weiteres Sechstel die Einführung solcher Verfahren plant. Qualitätsmanagement-Systeme sind bei den Einrichtungen und Diensten in freier Trägerschaft stärker verbreitet als in öffentlichen Ämtern und der Verbreitungsgrad nimmt (wenn die öffentlichen Einrichtungen außer Acht gelassen werden) mit der Größe dieser Einrichtungen zu. Auch eine Befragung aus dem Jahre 2001 belegt, dass neuere Verfahren und Instrumente der Qualitätsentwicklung – nicht zuletzt angestoßen durch förderrechtliche Vorgaben – auf der Ebene der Einrichtungen, Dienste und Träger der Kinder- und Jugendhilfe angekommen sind (vgl. Mamier 2003, S. 33). Erwartungsgemäß äußerten sich Jugendringe und Jugendverbände zu eigenen Verfahren und Instrumenten der Qualitätsentwicklung mehrheitlich nicht bzw. verneinten den Einsatz entsprechender Instrumente. Hervorzuheben ist auch die seltene Erwähnung der Beteiligung der Klienten bzw. Nutzer als Instrument der Qualitätsentwicklung (z.B. bei der individuellen Hilfeplanung). Dieser Befund entspricht den Ergebnissen einer kleineren Untersuchung von Peters u.a. (vgl. 2001), wonach auch nach Aussagen der befragten Landesjugendämter in den bestehenden Rahmenempfehlungen der Länder die Klienten bzw. Nutzer sozialpädagogischer Angebote (Eltern und ihre Kinder) bei der Erarbeitung von Qualitätsentwicklungsvereinbarungen keine Rolle spielen. Dies ist nicht zuletzt ein Resultat des Sachverhalts, dass in den Empfehlungen zu den Qualitätsentwicklungsvereinbarungen Regelungen und Vorschläge zur Partizipation der Klienten bzw. Nutzer fehlen (vgl. Peters/Ritter/Schollmeyer 2001).

Hervorzuheben ist zudem, dass in allen Untersuchungen bewährte und überkommene Verfahren und Instrumente der Qualitätsentwicklung wie Supervision, Aus- und Fortbildung, Fallkonferenzen sowie Konzeptionsentwicklung am häufigsten genannt

werden, während neuere Verfahren wie Selbst- und Fremdevaluation, Qualitätszirkel, Leitbildentwicklung, die Einführung eines Qualitätsbeauftragten etc. sowohl eine geringere Beliebtheit als auch eine geringere Verbreitung aufweisen. Während Merchel (vgl. 2005) dies als einen deutlichen Beleg für eine weit verbreitete Praxis der Entschärfung der Qualitätsherausforderung durch Verweis auf altbekannte Verfahren wertet, kommen sowohl Lee u.a. (vgl. 2007) für die US-amerikanische als auch Gerull (vgl. 2004) für die deutsche Situation zu einer positiveren Einschätzung: Sie bewerten entsprechende Äußerungen des Fachpersonals als ein Anzeichen dafür, dass die Einführung und Praktizierung neuer Verfahren der Qualitätsentwicklung von den Fachkräften erst dann als sinnvoll angesehen wird, wenn auch die Realisierungsbedingungen für die bisher bekannten Verfahren verbessert worden sind. Dieser Befund spricht dafür, dass die neuen Instrumente und Modelle der Qualitätssicherung die als bewährt geltenden Verfahren keineswegs verdrängen, sondern eher ergänzen.

Ist nun die aus den empirischen Befunden hervorgehende weite Verbreitung unterschiedlicher Modelle und Verfahren des Qualitätsmanagements in den Einrichtungen und Diensten der Kinder- und Jugendhilfe gleichzusetzen mit einer tief greifenden Veränderung der innerorganisatorischen Handlungsroutrinen und professionellen Vorgehensweisen? Diese Frage ist auf der Grundlage vorliegender empirischer Studien kaum schlüssig zu beantworten. Denn nachgewiesen ist bislang nur, dass sich die einzelnen Einrichtungen mit Verfahren und Instrumenten des Qualitätsmanagements beschäftigen und diese zumindest partiell in die Formalstruktur der Organisationen übernehmen. Ob der hiermit verbundene organisatorische Wandel tatsächlich die „Kernstruktur“ sozialpädagogischer Organisationen berührt, bleibt eine offene Forschungsfrage. Die Relevanz dieser Frage ergibt sich aus dem Forschungsstand: Die internationale Implementationsforschung konnte zeigen, dass die Übernahme eines formalen Qualitätsmanagement-Systems keineswegs mit seiner Verankerung in der Organisationskultur gleichzusetzen ist (vgl. etwa Zbaracki 1998 sowie Ahire/Waller/Golhar 1996). Auf solche Prozesse der „Immunsierung“ von Organisationen gegenüber Erwartungen aus der Umwelt verweist auch die neo-institutionalistische Organisationstheorie (vgl. nur Meyer/Rowan 1977 sowie 1978; vgl. Olk 1986 sowie zusammenfassend Schaefers 2002). Dieser Ansatz der Organisationstheorie richtet seine Aufmerksamkeit insbesondere auf den Umweltbezug von Organisationen und auf die Bedeutung von Fragen der Legitimierung organisationellen Handelns. Danach sind Organisationen bestrebt, durch die Übernahme von in der Umwelt verankerten „Rationalitätsmythen“ nach außen Legitimität zu gewinnen und damit ihren Ressourcenzufluss zu sichern. Damit erweisen sich viele Merkmale und Bestandteile der Formalstruktur von Organisationen – z.B. bestimmte Abteilungen, Verfahrensweisen oder Programme – keineswegs als Ausdruck „rationalen“ Organisierens, als vielmehr als Ergebnis der Ausrichtung der Formalstrukturen an unabweisbaren Erwartungen und Rationalitätsannahmen von außen. Meyer und Rowan (vgl. 1977) gehen nun von einer „losen Kopplung“ der Formal- und Aktivitätsstruktur der Organisation aus mit der Folge, dass institutionalisierte Erwartungen an die Organisation zwar in die Formalstruktur übernommen, aber gleichzeitig die eigentliche technische Ebene der Aktivitätsstruktur der Organisation von diesen Anforder-

rungen entlastet wird. Eine solche Entkopplung von Formal- und Aktivitätsstruktur einer Organisation ist umso attraktiver, je mehr die Rationalitätserwartungen der Umwelt und die Realisierungschancen innerhalb der Organisation in Konflikt geraten. Dies ist aber gerade in (sozial-)pädagogischen Organisationen wie Schulen, Jugendhilfeeinrichtungen etc. der Fall, da die Ziel-Mittel-Relationen im Interventionsfeld komplex und schwer durchschaubar und die professionellen Vorgehensweisen allenfalls als „weiche“ Technologien gelten können. Es kann also vermutet werden, dass die erfolgreiche Verbreitung von Effektivitäts- und Effizienzsteigerungen verheißenden Qualitätsmanagement-Modellen in sozialpädagogischen Organisationen eher ein Ausdruck der mimetischen und „zeremoniellen“ Übernahme als besonders erfolgreich geltender Modelle und weniger als Prozesse tief greifender Umstrukturierungen vorhandener Interventionsverfahren und Handlungspraxen in der „Kernstruktur“ der Organisationen gelten können. In welchem Ausmaß die Übernahme von Qualitätsmanagementmodellen in die Formalstruktur von Organisationen tatsächlich in der Lage ist, die „Qualitätspraxen“ (vgl. Beckmann u.a. 2006) in dieser Organisation nachhaltig zu verändern, ist also eine empirische Frage, die nur durch gründliche Fallstudien geklärt werden könnte.<sup>7</sup>

## **5. Qualitätsmanagement und professionelles Handeln**

Fasst man die Befunde und Einsichten des vorhergehenden Kapitels zusammen, dann lässt sich resümieren, dass es allenfalls lose und uneindeutige Verknüpfungen zwischen der Übernahme eines formalen Qualitätsmanagement-Systems und den hiermit verbundenen Auswirkungen auf organisationelle Handlungspraxen und technische Verfahrensweisen gibt. Qualitätsmanagement kann – muss aber nicht – weit reichende Organisationsentwicklungs-Prozesse auslösen. Ebenso unterbestimmt und vieldeutig sind auch die Wechselbeziehungen zwischen formalen Modellen des Qualitätsmanagements einerseits und den professionellen Handlungsweisen sozialpädagogischer Fachkräfte andererseits. Welche Auswirkungen die Einführung von Instrumenten und Verfahren des Qualitätsmanagements auf die unmittelbare Ebene der Interaktion zwischen sozialpädagogischen Fachkräften und Klienten haben, hängt von weiteren spezifischen binnenorganisatorischen Bedingungen ab. Dies hängt damit zusammen, dass Verfahren und Modelle des Qualitätsmanagements und die fachlichen Ansprüche an „gutes“ professionelles Arbeiten auf unterschiedlichen Ebenen ansetzen. Während sich das Qualitätsmanagement auf die Ebene der organisatorischen Rahmung sozialpädagogischen

7 Die Diskrepanz zwischen „talk“ und Realität wurde auch in der empirischen Studie von Beckmann u. a. (2006, 2007) nachgewiesen. So konnte kein Zusammenhang zwischen dem Grad der Formalisierung der Arbeitsbedingungen und der Einführung bzw. Nichteinführung eines formalen Qualitätsmanagementsystems nachgewiesen werden. Die Autoren interpretieren dies dahingehend, dass Einrichtungen der Kinder- und Jugendhilfe die Einführung eines formalen Qualitätsmanagementsystems als Konkurrenzfaktor auf dem Sozialmarkt betrachten und daher vor allem legitimatorisch mit der Qualitätsthematik umgehen.

Handelns bezieht, betreffen fachliche Gütekriterien sozialpädagogischen Handelns die Ebene der professionellen Selbstreflexion und der Interaktion zwischen sozialpädagogischen Fachkräften und Klienten. Burkhard Müller hat dies auf die prägnante Formel gebracht: „Professionalisierung soll soziale Arbeit verbessern, Qualitätsmanagement soll die Organisation sozialer Arbeit verbessern. Professionalisierung setzt *Autonomie*, d.h. Selbstkontrolle fachlich qualifizierter Personen (...), voraus. Qualitätsmanagement dagegen zielt auf die *Kontrollierbarkeit und Transparenz* organisatorischen Handelns nach innen und nach außen“ (2000, S. 27; Hervorh. i. Org.).

Stimmt man dieser Formulierung zu, dann bedeutet dies zugleich, dass die Einführung von Qualitätsmanagement-Systemen keineswegs automatisch zu einer Verbesserung der Qualität sozialpädagogischen Handelns führen muss. Im Gegenteil: Befürchtet wird, dass die Einführung von Qualitätsmanagement-Systemen mit einer Standardisierung und Taylorisierung professionellen Handelns verbunden sein könnte. Aus theoretisch-konzeptioneller Perspektive stellt sich daher die Frage, ob eine produktive Verknüpfung von beidem möglich ist, ob also die Formalisierung von Verfahrensabläufen durch Qualitätsmanagement-Systeme nicht zwangsläufig mit einer Verschlechterung der Handlungsspielräume professionellen Arbeitens, sondern vielmehr mit einer Ermöglichung professionellen Handelns einhergehen könnte. Beckmann u.a. (vgl. 2006, S. 7ff.) machen im Anschluss an Adler und Borys (vgl. 1996) den Vorschlag, nicht nur Grade, sondern auch Formen von Formalisierung innerhalb einer Organisation zu unterscheiden und in dieser Hinsicht sowohl „ermächtigende“ als auch „restringierende“ Formen der Formalisierung ins Auge zu fassen. Während restringierende Formen der Formalisierung hierarchische Formen der Kontrolle und Disziplinierung darstellen, die Handlungsspielräume einschränken, wirken ermächtigende Formen der Formalisierung in dem Sinne unterstützend für professionelles Handeln, als sie die Routinisierung „bewährter“ professioneller Problemlösungen unterstützen, ohne unnötige Standardisierungen zu erzwingen (vgl. für den Schulbereich auch Hoy/Sweetland 2001).

Damit erweisen sich ermächtigende Formen der Formalisierung als besonders geeignet, sozialpädagogisches Handeln organisatorisch abzustützen. Denn sowohl die sozialwissenschaftliche Dienstleistungstheorie (vgl. Gross/Badura 1977; Olk 1986 sowie Olk/Otto/Backhaus-Maul 2003) als auch die strukturtheoretische und interaktionistische Professionstheorie weisen auf die Nicht-Standardisierbarkeit sowie die Kommunikations- und Interaktionsintensität des sozialpädagogischen Handelns hin. So muss der Klient bzw. „Konsument“ der sozialpädagogischen Dienstleistung aktiv am Erbringungsprozess beteiligt sein, um so anspruchsvolle Ziele wie Verhaltensänderungen oder die Behebung sozialer oder psychischer Probleme erreichen zu können. Zudem verweist die Professionstheorie auf strukturelle Spannungen und Antinomien im Bereich des professionellen Handelns von Sozialpädagogen (vgl. Schütze 1996 sowie zusammenfassend Combe/Helsper 2002). Bei der berufsrollenförmigen Bearbeitung lebenspraktischer Konfliktsituationen von Klienten werden die sozialpädagogischen Fachkräfte mit der paradoxen Anforderung konfrontiert, sowohl den Anforderungen einer adressatenbezogenen biografischen Fallbearbeitung als auch den innerbetrieblichen Ablaufmustern der jeweiligen Organisation gerecht zu werden. Darüber hinaus sind gerade Sozial-

pädagogen unweigerlich in die Zuteilung knapper gesellschaftlicher Güter einbezogen, also – ob sie wollen oder nicht – in Kontroll-, Selektions-, Sanktions- und Ausgrenzungsmechanismen involviert (vgl. Combe/Helsper 2002, S. 32f.).

Die Orientierung am Klientenwohl und an einem Arbeitsbündnis und einer hiermit verbundenen Vertrauens- und Beratungskultur reibt sich fortwährend an den Kontrollstrukturen der Organisation. Gerade die herkömmlichen Formen der professionellen Selbstvergewisserung (wie z.B. Supervision, Fallkonferenzen, Teambesprechungen) dienen der fachlichen Bearbeitung solcher unabweisbarer Organisationsparadoxien professionellen pädagogischen Handelns.

Dieses professionstheoretische Modell sozialpädagogischer Fachlichkeit steht in scharfem Kontrast zu solchen Konzepten und Verfahren des Qualitätsmanagements, die auf die Durchsetzung formaler Standards, schematisierter Verhaltensabläufe und auf Kosten-Nutzen-Rechnungen beruhender Ressourcenzuweisungen abzielen. Dennoch ist unübersehbar, dass bei aller Skepsis und Ablehnung instrumentell-technischer Methodenlehren und standardisierter Checklisten die sozialpädagogischen Fachkräfte gerade wegen der Diffusität und strukturellen Widersprüchlichkeit des sozialpädagogischen Handelns immer wieder ein starkes Interesse an wie auch immer gearteten Handlungsmethoden und Verfahrensvorschlägen artikuliert. Insofern lassen sich insbesondere solche Qualitätskonzepte, die darauf abzielen, die Besonderheiten sozialpädagogischen Handelns stärker zu berücksichtigen und ein Qualitätsmanagement zu entwickeln, dass den Eigenheiten sozialpädagogischen Handelns besser entspricht, durchaus als Beiträge zur professionellen Selbstvergewisserung hinsichtlich der Ziele des Handelns, der angemessenen Vorgehensweisen und der Ergebnisüberprüfung verstehen (vgl. Merchel 2005, S. 39f.). Hierzu zählen solche Instrumente und Verfahren, die weniger dem Gedanken der „Qualitätssicherung“ verpflichtet sind, als vielmehr zu einer kritischen (Selbst-)Reflexion und einer Entwicklung innovativer Problemlösungen beitragen sollen. Dies gilt beispielsweise für Fallbesprechungen, Supervision, kollegiale Beratungen und Selbstevaluation (vgl. Spiegel 2004). Insbesondere die Selbstevaluation hat in den letzten Jahren an Bedeutung gewonnen. Ausgehend von einer Kritik an fehlenden reflexiven Konzepten und Verfahren zur Dokumentation und Auswertung von Ergebnissen und alltäglichen Arbeitsabläufen in der Sozialen Arbeit wurde eine Vielzahl unterschiedlicher Verfahren der Selbstevaluation entwickelt (vgl. Heiner 1999 und 2000; Beywl/Bestvater 1998; Spiegel 2004). Selbstevaluation kann dabei sowohl nach innen (z.B. Erfolgskontrolle, Aufklärung, Qualifizierung) als auch nach außen (z.B. Öffentlichkeitsarbeit, Legitimierung) wirken (vgl. Spiegel 2004). Grundsätzlich lassen sich zwei Selbstevaluationsverständnisse in der Sozialen Arbeit unterscheiden: Bei der Selbstevaluation im engeren Sinne, die in der sozialpädagogischen Disziplin favorisiert wird, soll über einen längeren, aber zeitlich befristeten Zeitraum ein Forschungsprojekt zum eigenen Handeln mit allen Konsequenzen und aufeinander aufbauenden Arbeitsschritten durchgeführt werden (Selbstevaluation als Forschungsprojekt). Selbstevaluation im weiteren Sinne wird dagegen von den einzelnen Fachkräften systematisch, d.h. kontinuierlich und integriert im Arbeitsalltag durchgeführt und bedarf daher handhabbarer und praxistauglicher Instrumente (Selbstevaluation als integraler Bestandteil der Arbeit).

Konflikte und Reibungsflächen zwischen Konzepten und Verfahren des Qualitätsmanagements bzw. der Fremd- und Selbstevaluation einerseits und dem Eigensinn sozialpädagogischem Handeln andererseits sind insbesondere dort zu erwarten, wo an quantitativ messbaren Ergebniszielen und Kostenrechnungen orientierte schematisierte Verfahrensvorschriften mit den professionellen Prinzipien einer fall- und kontextsensiblen Herangehensweise kollidieren (vgl. Langer 2007; vgl. auch das Fallbeispiel bei van der Laan 2000). Insofern müssen die möglichen deprofessionalisierenden Effekte von Instrumenten wie Produktbeschreibungen, Qualitätshandbücher oder Checklisten auf das professionelle Handeln im Auge behalten werden.

Auf der anderen Seite könnte die aktuelle Qualitätsdebatte aber auch produktive Anstöße für die Weiterentwicklung sozialpädagogischer Vorgehensweisen geben. Diese könnten in der klaren Aufforderung liegen, die Ziele des professionellen Handelns zu präzisieren, den relativen Erfolg bestimmter Verfahrensweisen empirisch zu überprüfen und die Ergebnisse des eigenen Handelns systematisch zu evaluieren (vgl. z.B. die Ansätze dazu bei Schmidt u.a. 2002 sowie zusammenfassend Otto 2007). Solche Vorgehensweisen der empirisch abgestützten Selbstvergewisserung hinsichtlich der Ziele und Vorgehensweisen professionellen Handelns würden sowohl die fachliche Kompetenz nach innen als auch die Legitimation des eigenen Handelns nach außen stärken. Denn die Ausweisung nachvollziehbarer Handlungsziele, das transparent Machen von „Gütekriterien“ professionellen Handelns und der Nachweis von Wirkungen stellen einen nicht zu unterschätzenden Nachweis des gesellschaftlichen Nutzens des eigenen Tuns nach außen dar.

Die bisherigen Ausführungen haben deutlich gemacht, dass komplexe Wechselwirkungen zwischen Qualitätsmanagement, innerorganisatorischer Einflussverteilung zwischen unterschiedlichen Gruppen und der professionellen Handlungsautonomie bestehen (vgl. z.B. Cook/Verma 2004). Beckmann u.a. (vgl. 2006) können nun nachweisen, dass sich diese Bedingungskonstellationen auch auf die Qualität der Interaktion zwischen Professionellen und Klienten und damit auf die Effekte sozialpädagogischen Handelns auswirken.

Mittels Faktoren- und Clusteranalysen konnten insgesamt drei Typen von Organisationen herauskristallisiert werden. Unterscheidbar sind demnach „klassische Profi-Organisationen“, „managerielle Maschinenbürokratien“ und „Qualitätsbewusste Profi-Organisationen“ (vgl. Beckmann u.a. 2006, S. 58ff.). Die klassischen Profi-Organisationen zeichnen sich durch einen vergleichsweise geringen Konsens über Qualitätsfragen, professionelle, ermächtigende Arbeitsbedingungen sowie ein offensichtlich als von außen aufgezwungen erlebtes Qualitätsmanagement-System aus. Managerielle Maschinenbürokratien hingegen verfügen zwar auch über einen geringen Konsens über Qualitätsfragen und tendenziell ein formales und von „oben“ aufgezwungenes Qualitätsmanagement-System, allerdings existieren bei ihnen stark deprofessionalisierende, restringierende Arbeitsbedingungen. Organisationsbindung und professionelle Wertbindung der Professionellen sind gering. Die Qualitätsbewussten Profi-Organisationen hingegen weisen – dies erscheint zunächst überraschend – die geringste Wahrscheinlichkeit eines formalen Qualitätsmanagements aus. Sie zeichnen sich stattdessen durch einen hohen

Konsens über Qualitätsfragen, eine starke Organisationsbindung und professionelle Wertbindung der Fachkräfte sowie vergleichsweise gute Arbeitsbedingungen aus (vgl. Beckmann u.a. 2006, S. 58ff). Als entscheidend für eine hohe Dienstleistungsqualität erweisen sich die Einflussmöglichkeiten der Klienten (Eltern und Kinder) auf die eigentliche professionelle Arbeit und die Gestaltungsspielräume der Fachkräfte in den Organisationen (vgl. Beckmann u.a. 2006, S. 62).

## 6. Fazit

Mit der verpflichtenden Einführung von Qualitätsentwicklungsverfahren für Einrichtungen der Kinder- und Jugendhilfe, die stationäre und teilstationäre Dienste nach dem SGB VIII anbieten, zum 1.1. 1999 ist eine starke Tendenz zur Einführung von Qualitätsmanagementverfahren zu beobachten. Im Gegensatz zum schulisch geprägten Bildungssystem verläuft dieser Prozess in der Kinder- und Jugendhilfe in einer durch Wettbewerb und Konkurrenz geprägten pluralen Jugendhilfeträgerlandschaft dezentral, indem einzelne Träger, Einrichtungen und Dienste jeweils spezifische Instrumente und Verfahren des Qualitätsmanagements adaptieren und in ihre Formalstruktur übernehmen. Dabei deuten die vorliegenden empirischen Untersuchungen darauf hin,

- dass die Einführung von umfassenden Qualitätsmanagement-Systemen bzw. einzelner Instrumente und Verfahren des Qualitätsmanagements vor allem der Verbesserung der „Marktposition“ der betreffenden Einrichtung bzw. des betreffenden Trägers dienen soll. Die Übernahme von Qualitätsmanagement in die Formalstruktur einer Organisation muss daher keineswegs mit einer tiefgreifenden Umstrukturierung der Aktivitätsstruktur der Organisation verbunden sein; symbolische Inszenierung und organisationale Realität sind zwei unterschiedliche Realitätsebenen.
- dass die Einführung von Qualitätsmanagement-Verfahren und -Systemen in der Regel mit einer Umverteilung der relativen Macht- und Einflusststrukturen unterschiedlicher Akteursgruppen einhergeht. Über die unterschiedlichen Organisationstypen und inhaltlichen Arbeitsbereiche hinweg scheint die Einführung von Qualitätsmanagement-Verfahren durchgängig eine Professionalisierung von Organisations- und Managementkompetenzen und eine Bedeutungsaufwertung des mittleren Managements zu Lasten der sozialpädagogischen Fachkräfte nach sich zu ziehen. Während die Führungs- und Leitungskräfte in den Einrichtungen und Diensten an Kontroll- und Einflusschancen hinzu gewinnen, verlieren die sozialpädagogischen Professionellen einen Teil ihrer Definitions- und Kontrollmacht.
- dass die Einführung von Qualitätsmanagement-Verfahren zu einer Formalisierung des sozialpädagogischen Handelns führt. Allerdings bedeutet Formalisierung nicht in jedem Fall Deprofessionalisierung und Taylorisierung. Formalisierung kann sowohl „ermächtigend“ als auch „restringierend“ wirken, je nach dem, welche Form formaler Regelungen implementiert wird. Allerdings scheint es in der Mehrzahl der Fälle zu einer bürokratischen Überformung professionellen Handelns zu kommen.

- dass die Einführung von Qualitätsmanagement-Verfahren trotz aller Rhetorik der Kundenorientierung keineswegs im Selbstlauf zu einer Stärkung der Partizipations- und Definitionsmacht von (potentiellen) Klienten führt. Während – wie skizziert – die Kontroll- und Definitionsmacht von administrativen Führungs- und Leitungskräften bei der Ergebnis- und Prozesssteuerung eher zunehmen, stellen die meisten Qualitätsmanagement-Verfahren für die Stärkung der Einflusschancen von Klienten allenfalls indirekte Verfahren wie Kundenbefragungen und Beschwerdemanagement bereit. Selbst in rechtlich stark durch Partizipationsnormen geprägten Verfahren wie etwa dem Hilfeplanverfahren nach § 36 SGB VIII scheinen die Partizipationschancen von Klienten unter dem Einfluss neuer Definitionschancen der Prozessmanager abzunehmen.

Vor dem Hintergrund solcher Tendenzen und der unbefriedigenden Forschungslage wären insbesondere die folgenden Forschungsfragen zu klären:

- In der prozessualen Perspektive bedarf es gründlicher qualitativer Fallstudien, in denen die Beweggründe, die Verlaufsformen und die Ergebnisse der Einführung jeweils spezifischer Qualitätsmanagement-Verfahren unter konkreten organisatorischen Bedingungen analysiert werden. In solchen Studien könnten die je spezifischen Interessenlagen und Relevanzsetzungen unterschiedlicher Akteursgruppen (Führungs- und Leitungskräfte, mittleres Management, sozialpädagogische Fachkräfte, Klienten etc.), die hieraus resultierenden Einfluss- und Bündnisstrategien sowie deren Auswirkungen auf die Resultate solcher Organisationsentwicklungsprozesse rekonstruiert werden.
- Aus der Querschnittsperspektive wäre über die vorliegende Erkenntnislage hinaus genauer zu klären, welche systematischen Zusammenhänge zwischen bestimmten Merkmalen einer Organisation (Größe, Struktur der Ressourcenabhängigkeit, Tätigkeitsfeld etc.), der Einführung bestimmter Qualitätsmanagement-Verfahren, dem Grad der Autonomie professionellen Handelns und den Partizipationschancen (potenzieller) Klientengruppen bestehen. Auf diese Weise ließe sich genauer herausarbeiten, unter welchen Bedingungen Deprofessionalisierungstendenzen bzw. „ermächtigende“ Formen der organisationalen Abstützung professioneller Handlungsformen zu erwarten sind.
- Ein erhebliches Forschungsdefizit besteht auch hinsichtlich der Auswirkungen von Qualitätsmanagement-Verfahren auf die Partizipationschancen (potenzieller) Klienten. Obwohl es die erklärte Absicht der Befürworter von Qualitätsmanagement-Systemen ist, die Nachfragemacht der „Kunden“ zu stärken, gibt es für die tatsächlichen Auswirkungen des Qualitätsmanagements auf die Einflusschancen unterschiedlicher Klientengruppen kaum empirische Untersuchungen. Auch in dieser Hinsicht ginge es darum, pauschalisierende, allenfalls normativ begründete Behauptungen (Stärkung von Konsumentensouveränität versus Einschränkung von Partizipationschancen) zu differenzieren und die jeweiligen Einflusskonstellationen zu klären.

**Literatur**

- Adler, P.S./Borys, B. (1996): Two Types of Bureaucracy: Enabling and Coercive. In: *Administrative Science Quarterly* 41, H. 1, S. 61–89.
- Ahire, S.L./Waller, M.A./Golhar, D.Y. (1996): Quality Management in TQM Versus non-TQM Firms: An Empirical Investigation. In: *International Journal of Quality & Reliability Management* 13, H. 8, S. 8–27.
- Beckmann, C./Otto, H.U./Richter, M./Schrödter, M. (Hrsg.) (2004): *Qualität in der Sozialen Arbeit. Zwischen Nutzerinteresse und Kostenkontrolle*. Wiesbaden: VS-Verlag.
- Beckmann, C./Otto, H.-U./Schaarschuch, A./Schrödter, M. (2006): *Qualität und Wirkung in der Sozialpädagogischen Familienhilfe. Vorläufige Ergebnisse des DFG-Projektes „Dienstleistungsqualität“*. Bielefeld/Wuppertal. Online unter [http://www.dlq-online.de/document/Ergebnis\\_bericht\\_DFG-Projekt\\_DLQ.pdf](http://www.dlq-online.de/document/Ergebnis_bericht_DFG-Projekt_DLQ.pdf), letzter Zugriff am 12.08.2007.
- Beckmann, C./Otto, H.-U./Schaarschuch, A./Schrödter, M. (2007): Qualitätsmanagement und Professionalisierung in der Sozialen Arbeit. Ergebnisse einer Studie zu organisationalen Bedingungen ermächtigender Formalisierung. In: *Zeitschrift für Sozialreform* 53, H. 3, S. 275–295.
- Bermann, E.M. (1995): Implementing TQM in State Welfare Agencies. In: *Administration in Social Work* 19, H. 1, S. 55–72.
- Beywl, W./Bestvater, H. (1998): Selbst-Evaluation in pädagogischen und sozialen Arbeitsfeldern, Ergänzung und Alternative zur Fremdevaluation, In: Bundesvereinigung Kulturelle Jugendbildung (Hrsg.): *Qualitätssicherung durch Evaluation. Konzepte, Methoden, Ergebnisse – Impulse für die kulturelle Kinder- und Jugendbildung*. Remscheid: Eigenverlag, S. 33–44.
- Bogumil, J./Grohs, S./Kuhlmann, S./Ohm, A.K. (2007): *Zehn Jahre Neues Steuerungsmodell. Eine Bilanz kommunaler Verwaltungsmodernisierung*. Berlin: Edition Sigma.
- Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) (2002): *QS Kompendium. Gesamtausgabe aller QS-Hefte von 1-36. Materialien zur Qualitätssicherung in der Kinder- und Jugendhilfe*. (CD-ROM)
- Bundesministerium für Jugend, Familie, Frauen und Gesundheit (BMJFFG) (1990): *Achter Jugendbericht – Bericht über Bestrebungen und Leistungen der Jugendhilfe*. Bonn.
- Combe, A./Helsper, W. (2002): Professionalität. In: Otto, H.-U./Rauschenbach, T./Vogel, P. (Hrsg.): *Erziehungswissenschaft: Professionalität und Kompetenz*. Opladen: Leske+Budrich, S. 29–47.
- Cook, L.S./Verma, R. (2004): Exploring the Linkages Between Quality System, Service Quality, and Performance Excellence: Service Providers' Perspectives. In: *Quality Management Journal* 9, H. 2, S. 15–35.
- Engel, M./Flösser, G./Gensink, G. (1998): Quality-Enhancement in the Service Society: Perspectives for Social Work. In: Flösser, G./Otto, H.-U. (Hrsg.): *Towards More Democracy in Social Services*. Berlin/New York: Walter de Gruyter, S. 357–370.
- Flösser, G./Oechler, M. (2004): Chancen und Risiken von Qualitätsmanagement in der Sozialen Arbeit. In: Beckmann, C./Otto, H.-U./Richter, M./Schrödter, M. (Hrsg.): *Qualität in der Sozialen Arbeit. Zwischen Nutzerinteresse und Kostenkontrolle*. Wiesbaden: VS Verlag, S. 175–184.
- Galiläer, L. (2005): *Pädagogische Qualität. Perspektiven der Qualitätsdiskurse über Schule, Soziale Arbeit und Erwachsenenbildung*. Weinheim/München: Juventa.
- Gerull, K.-P. (2000): *Hand- und Werkbuch Soziales Qualitätsmanagement. Konzepte und Erfahrungen*. Hannover: EREV.
- Gerull, K.-P. (2004): *Qualitätsmanagement in der Jugend- und Sozialhilfe. Literaturanalytische und empirische Studien*. Dissertation. Göttingen.
- Gottlieb, H.-D./Kaufhold, S./Thomsen, M. (2003): Rahmenverträge nach § 78f SGB VIII. In: Verein für Kommunalwissenschaften e.V./AFET e.V. (Hrsg.): *Die Vereinbarungen nach §§ 78a ff. SGB VIII. Bestandsaufnahme und Analyse der Leistungs-, Entgelt- und Qualitätsentwicklungsvereinbarungen sowie der Rahmenverträge*, Berlin, S. 17–37.

- Gross, P./Badura, B. (1977): Sozialpolitik und soziale Dienste: Entwurf einer Theorie personenbezogener Dienstleistungen. In: Ferber, C.v./Kaufmann, F.-X. (Hrsg.): *Soziologie und Sozialpolitik*. Sonderheft 19 der Kölner Zeitschrift für Soziologie und Sozialpsychologie, S. 381–386.
- Heiner, M. (1999): Qualitätsentwicklung durch Evaluation. In: Peterander, F./Speck, O. (Hrsg.): *Qualitätsmanagement in sozialen Einrichtungen*. München/Basel: Ernst Reinhardt, S. 63–88.
- Heiner, M. (2000): Interne Evaluation institutioneller Steuerung und pädagogischer Reflexion – Sieben Thesen. In: Müller-Kohlenberg, H./Münstermann, K. (Hrsg.): *Qualität von Humandienstleistungen*. Evaluation und Qualitätsmanagement in Sozialer Arbeit und Gesundheitswesen. Opladen: Leske+Budrich, S. 207–226.
- Hollenstein, E. (1993): Wirkungsanalysen in der ambulanten Erziehungshilfe. Eine Evaluationsstudie in der Erziehungsbeistandschaft. In: *Neue Praxis* 4, S. 346–360.
- Honig, M.-S./Joos, M./Schreiber, N. (2004): Was ist ein guter Kindergarten? Theoretische und empirische Analysen zum Qualitätsbegriff in der Pädagogik. Weinheim/München: Juventa.
- Hornstein, W. (2000): Qualität und Evaluation in der Sozialpädagogik. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): *Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule*. 41. Beiheft der Zeitschrift für Pädagogik. Weinheim/Basel: Beltz, S. 129–136.
- Hoy, W.K./Sweetland, S.R. (2001): Designing Better Schools: The Meaning and Measure of Enabling School Structures. In: *Educational Administration Quarterly* 37, S. 296–321.
- KGSt (Kommunale Gemeinschaftsstelle für Verwaltungsvereinfachung (Hrsg.) (1994): *Outputorientierte Steuerung in der Jugendhilfe*. 9/1994, Köln.
- KGSt (Kommunale Gemeinschaftsstelle für Verwaltungsvereinfachung (Hrsg.) (1995): *Qualitätsmanagement*. 6/1995, Köln.
- Köpp, C./Neumann, S. (2003): *Sozialpädagogische Qualität. Problembezogene Analysen zur Konzeptualisierung eines Modells*. Weinheim/München: Juventa.
- Kröger, R. (Hrsg.) (1999): *Leistung, Entgelt und Qualitätsentwicklung in der Jugendhilfe. Arbeitshilfen mit Musterbeispielen zur praktischen Umsetzung der §§ 78a –g SGB VIII*. Neuwied/Kriftel: Luchterhand.
- Kuper, H. (2002): Stichwort: Qualität im Bildungssystem. In: *Zeitschrift für Erziehungswissenschaft* 5, S. 533–551.
- Laan, G. v. d. (2000): Handlungsspielräume von Sozialarbeitern unter Marktbedingungen – Das Beispiel Niederlande. In: Otto, H.-U./Schurr, S. (Hrsg.): *Privatisierung und Wettbewerb in der Jugendhilfe: Marktorientierte Modernisierungsstrategien in internationaler Perspektive*. Neuwied/Kriftel: Luchterhand, S. 87–110.
- Langer, A. (2007): Dienstleistungsstrukturen in der Sozialen Arbeit zwischen Verwaltungsreform und Professionalisierung. In: *Zeitschrift für Sozialreform* 53, H. 3, S. 223–246.
- Lee, B.R./McMillen, J.C./Knudsen, K./Woods, C.M. (2007): Quality-Directed Activities and Barriers to Quality in Social Service Organizations. In: *Administration in Social Work* 31, H. 2, S. 67–85.
- Mamier, J./Pluto, I./van Santen, E./Seckinger, M./Zink, G. (2003): *Geschäftsstellen freier Träger in der Kinder- und Jugendhilfe. Empirische Analysen zu Funktionen und Handlungsmöglichkeiten auf kommunaler Ebene*. München. Online unter [http://www.dji.de/bibs/64\\_2110.pdf](http://www.dji.de/bibs/64_2110.pdf), letzter Zugriff am 12.08.2007.
- Merchel, J. (2000): Zwischen Hoffnung auf Qualifizierung und Abwehrreflex. Zum Umgang mit dem Qualitätsthema in der Jugendhilfe. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): *Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule*. Zeitschrift für Pädagogik, 41. Beiheft. Weinheim/Basel: Beltz, S. 161–183.
- Merchel, J. (2005): Was hat die Qualitätsdebatte in der Jugendhilfe gebracht? Versuch einer Zwischenbilanz. In: *Archiv für Wissenschaft und Praxis der Sozialen Arbeit* 36, H. 2, S. 38–59.
- Merchel, J. (2006): Qualitätsmanagement in der Sozialen Arbeit. Erfahrungen zur Verarbeitung und zur Umsetzung des Themas „Qualität“. In: *Pädagogischer Blick* 14, S. 195–208.
- Meyer, J.W./Rowan, B. (1977): Institutionalized Organizations: Formal Structure as Myth and Ceremony. In: *American Journal of Sociology* 83, S. 340–363.

- Meyer, J.W./Rowan, B. (1978): The Structure of Educational Organizations. In: Meyer, J.W. (Hrsg.): *Environments and Organizations*. San Francisco: Jossey-Bass, S. 78–109.
- Müller, B. (2000): Qualitätsmanagement und professionelle Autonomie. In: Institut für soziale Arbeit e.V. (Hrsg.): *Prädikat wertvoll – Qualität sozialer Arbeit. Jugendhilfe im Kontext sozialer Konflikte und individueller Krisen*. Materialien und Beiträge zum ISA-Kongress 27.-29.9.1999 in Dortmund. Münster: Votum, S. 26–40.
- Münder, J./Baltz, J./Jordan, E. (2006): *Frankfurter Kommentar zum SGB VIII: Kinder- und Jugendhilfe*. 5., vollst. überarb. Aufl., Gesetzesstand: 01.04.2006. Weinheim/München: Juventa.
- Münder, J./Tammen, B. (2003): *Die Vereinbarungen nach § 78a ff. SGB VIII. Eine Untersuchung von Leistungs-, Entgelt- und Qualitätsentwicklungsvereinbarungen*. Eine Studie im Auftrag des Bundesministeriums für Familie, Senioren, Frauen und Jugend. Berlin (Verein für Kommunalwissenschaften).
- Olk, T. (1986): *Abschied vom Experten. Sozialarbeit auf dem Weg zu einer alternativen Professionalität*. Weinheim/München: Juventa.
- Olk, T. (2001): *Träger der Sozialen Arbeit*. In: Otto, H.-U./Thiersch, H. (Hrsg.): *Handbuch Sozialarbeit/Sozialpädagogik*. 2., vollst. Überarb. Aufl. Neuwied/Kriftel: Luchterhand, S. 1910–1926.
- Olk, T./Otto, H.-U./Backhaus-Maul, H. (2003): *Soziale Arbeit als Dienstleistung – Zur analytischen und empirischen Leistungsfähigkeit eines theoretischen Konzepts*. In: Olk, T./Otto, H.-U. (Hrsg.): *Soziale Arbeit als Dienstleistung*. Grundlegung, Entwürfe und Modelle. München/Unterschleißheim: Luchterhand, S. IX–LXXII.
- Olk, T./Speck, K./Albert, E. (2004): *Abschlussbericht zur wissenschaftlichen Begleitung der Entwicklung von Zuwendungsverträgen in der öffentlich geförderten Familienbildung des Landes Sachsen-Anhalt*. Im Auftrag des Ministeriums für Gesundheit und Soziales sowie des Landesjugendamtes in Sachsen-Anhalt. Halle.
- Olk, T./Speck, K./Sattler, C. (2004): *Abschlussbericht zur wissenschaftlichen Begleitung der Qualitätsentwicklung in der öffentlich geförderten, außerschulischen Kinder- und Jugendbildung im Land Sachsen-Anhalt 2003*. Im Auftrag des Ministeriums für Gesundheit und Soziales sowie des Landesjugendamtes in Sachsen-Anhalt. Halle.
- Otto, H.-U., Arbeitsgemeinschaft für Kinder- und Jugendhilfe (Hrsg.) (2007): *Zum aktuellen Diskurs um Ergebnisse und Wirkungen im Feld der Sozialpädagogik und Sozialarbeit – Literaturvergleich nationaler und internationaler Diskussion. Expertise im Auftrag der Arbeitsgemeinschaft für Kinder- und Jugendhilfe – AGJ*. Berlin.
- Peterander, F./Speck, O. (Hrsg.) (1999): *Qualitätsmanagement in sozialen Einrichtungen*. München/Basel: Reinhardt.
- Peters, F./Ritter, K./Schollmeyer, K. (2001): *Forschungsprojekt Qualitätsentwicklung in der Jugendhilfe*. Endbericht. Online unter <http://www.fh-erfurt.de/so/projekte/qualitaet/ergebnisse>, letzter Zugriff am 08.05.2003.
- Rauschenbach, T./Sachße, C./Olk, T. (Hrsg.) (1996): *Von der Wertgemeinschaft zum Dienstleistungsunternehmen*. Jugend- und Wohlfahrtsverbände im Umbruch. 2. Aufl., Frankfurt a.M.: Suhrkamp.
- Schaefers, C. (2002): *Der soziologische Neo-Institutionalismus. Eine organisationstheoretische Analyse- und Forschungsperspektive auf schulische Organisationen*. In: *Zeitschrift für Pädagogik* 48, S. 835–855.
- Schmidt, M./Schneider, K./Hohm, E./Pickartz, A./Macsenaere, M./Petermann, F./Flosdorf, P./Hötzl, H./Knab, E. (2002): *Effekte erzieherischer Hilfen und ihre Hintergründe*. Stuttgart: Kohlhammer.
- Schütze, F. (1996): *Organisationszwänge und hoheitsstaatliche Rahmenbedingungen im Sozialwesen: Ihre Auswirkung auf die Paradoxien des professionellen Handelns*. In: Combe, A./Helsper, W. (Hrsg.): *Pädagogische Professionalität. Untersuchungen zum Typus pädagogischen Handelns*. Frankfurt a.M.: Suhrkamp, S. 183–275.
- Shaw, I. (1995): *The Quality of Mercy: The Management of Quality in the Personal Social Services*. In: Kirkpatrick, I./Lucio, M.M. (Hrsg.): *The Politics of Quality in the Public Sector. The Management of Change*. London/New York: Routledge, S. 128–150.

- Speck, K. (2006): Qualität und Evaluation in der Schulsozialarbeit. Konzepte, Rahmenbedingungen und Wirkungen. Wiesbaden: VS Verlag.
- Spiegel, H. v. (2000) (Hrsg.): Jugendarbeit mit Erfolg. Arbeitshilfen und Erfahrungsberichte zur Qualitätsentwicklung und Selbstevaluation. Ein Modellprojekt des Landesjugendamtes. Münster: Votum.
- Spiegel, H. v. (2004) (Hrsg.): Methodisches Handeln in der Sozialen Arbeit. München: Reinhardt.
- Volkmar, S. (1998): Qualität sozialer Einrichtungen. In: Brunner, E.J./Bauer, P./Volkmar, S. (Hrsg.): Soziale Einrichtungen bewerten. Theorie und Praxis der Qualitätssicherung. Freiburg: Lambertus, S. 54–68.
- Zbaracki, M.J. (1998): The Rhetoric and Reality of Total Quality Management. In: Administrative Science Quarterly 43, S. 602–636.

*Anschrift der Autoren:*

Prof. Dr. Thomas Olk, Martin-Luther-Universität Halle-Wittenberg, Philosophische Fakultät III – Erziehungswissenschaften, Institut für Pädagogik, Franckeplatz 1, Haus 6, 06099 Halle (Saale), Tel.: ++49 345 5523801, Fax.: ++49 345 5527062, E-Mail: thomas.olk@paedagogik.uni-halle.de.

Dr. Karsten Speck, Universität Potsdam, Humanwissenschaftliche Fakultät, Institut für Erziehungswissenschaft, Golm Campus, Haus 14, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Tel.: ++49 331 9772697, Fax.: ++49 331 9772067, E-Mail: speck@rz.uni-potsdam.de.

*Philipp Gonon*

## **Qualitätssicherung in der beruflichen Bildung**

*Eine Bilanz*

Seit den 1990er Jahren hat sich Qualität als vielgestaltiges Konzept auch in der beruflichen Bildung etabliert. Im Folgenden werden drei Ebenen unterschieden, in welchen Qualität und Qualitätssicherung sich als Begrifflichkeiten gefestigt haben: die Makro-, Meso- und Mikroebene. Es wird gerade in der Diffusität und im Facettenreichtum von Qualitätsprinzipien, -maßnahmen und -umsetzungen eine neue Form der Steuerung und Legitimation beruflicher Bildung sichtbar.

Die Omnipräsenz des Qualitätsbegriffs innerhalb des Bildungs- und Erziehungsdiskurses ist beachtlich. Bis zu Beginn der 90er Jahre war das Konzept einer Qualitätssicherung vorwiegend der industriellen Produktion vorbehalten. Nun ist das Sichern und Managen von Qualität Bestandteil des pädagogischen Glossars der Gegenwart geworden (Bakic 2006). Qualität ist zum „Modernisierungsimperativ“ in der Bildungsreform avanciert. In der aktuellen Qualitätsdiskussion überlagern sich hierbei unterschiedliche Handlungslogiken (Hartz 2004, S. 241).

Hervorgegangen aus einem primär betriebswirtschaftlich ausgerichteten Managementkonzept, hat sich die Sicherung der Qualität als Zielvorgabe und Steuerungsinstrument auch in weitere gesellschaftliche Bereiche verbreitet. Dieses „Inside out“ und „Going Public“ des ursprünglich internen betrieblichen Kontrollsystems kennzeichnet gemäß Michael Power die so benannte „Explosion“ von Audits und Akkreditierungen in der „Audit-Gesellschaft“. Im Management von Unsicherheit und Risiko wird im Sinne einer Erhöhung der Legitimität von Maßnahmen und Regeln, Beratungen und Strategien, Qualitätssicherung zur regulatorischen Ressource (Power 2007, S. 52). Nicht lediglich betriebliche und institutionelle Prozesse, auch die große Politik der Berufsbildung selbst wird als „Qualitätssicherung“ bestimmt. Darüber hinaus wird guter Unterricht und erfolgreiches oder Erfolg versprechendes Lehren und Lernen als Qualitätsentwicklung, so etwa im „Lernerorientierten Qualitätsmodell“ rekonstruiert (Zech 2005).

### **1. Thematische Schwerpunkte der Qualitätssicherung in der beruflichen Bildung**

Was Qualität in der beruflichen Bildung einschließt, wird zu Recht als vage und bestrittbar bezeichnet, es wird in der Regel jedoch ein „customer-driven approach“ als Gemeinsamkeit hervorgehoben (Muller/Funnell 1992, S. 257). Es ist augenfällig, dass die Qualitätssicherung sich in der beruflichen Bildung vorwiegend im Bereich der Berufsschulen, der beruflichen Weiterbildung und im Bezug auf Modellversuche thema-

tisch äußert, weniger jedoch die Qualität der betrieblichen Ausbildung. Im Blickpunkt ist immerhin auch das Zusammenspiel der unterschiedlichen Lernorte der Berufsbildung. Für viele Reformvorhaben wie z.B. für die „Ausbildung im Verbund“ wird dank eines solchen Konzepts eine Steigerung der Qualität und Effizienz erwartet (Precht/Scherbe 2007). Auch in Ländern, die sich nicht auf eine Tradition betrieblicher Ausbildung stützen, wird die „Qualität“ der Zusammenarbeit zwischen Schulen und der Vermittlung von betrieblichen „internships“, also betriebliche Praktika, untersucht (Bailey/Hughes/Barr 2000), bzw. auch die Einführung von dualen Ausbildungsformen unter dieser Perspektive thematisiert (Smits 2006).

Modellversuche in der beruflichen Bildung weisen sich durch Qualitätssicherungsmaßnahmen aus, so wie die niedersächsischen regionalen berufsbildenden Kompetenzzentren (ReKo), die aufgrund eines EFQM-Qualitätsmanagements Legitimität und Akzeptanz bei Geldgebern und Beteiligten beanspruchen (Rütters, Roggenbrodt, Künzel 2004, S. 187).

Gelegentlich wird auch die Berufsbildungsforschung selbst, die es im allgemeinen zu heben gelte, als Frage der Qualitätssicherung thematisiert. Dazu gehören auch die internationalen und von der EU geförderten Forschungsverbünde und -projekte, die Qualitätskriterien erstellen, die z.B. für die Förderung der Frauen in der Weiterbildung relevant sind (Foster 2000), oder das „eLearning“ und die Berufsbildung thematisieren.

Wie viele Autoren, so unterscheidet beispielsweise auch Nickolaus zur „Strukturierung des Problemraums“ eine „Input-, Prozess- und Outputqualität“ (2007, S. 161). Es gibt darüber hinaus Ansätze, die diesem „Bildungsproduktionsmodell“ weitere Dimensionen hinzufügen, nämlich eine Transferqualität, die pädagogisch erwünschte Ziele ausweist, und eine außerpädagogische Outcome-Qualität (z.B. Timmermann/Windschild 1996, S. 89). Wir können die Perspektiven der Qualität und Qualitätssicherung in folgender Matrix darstellen:

<b>Qualitätsdimension/ Qualitätsebenen</b>	<b>Inputqualität</b>	<b>Prozessqualität</b>	<b>Outputqualität</b>
Mikroebene	Ausstattung der Lernumgebung	Erfassung, Bewertung und Standardisierung des Lehrens und Lernens	Zertifizierter Lernerfolg, berufliche Karrieren
Mesoebene	Infrastruktur, Lehrplan und Lehrgänge, Professionalisierung der Lehrkräfte und Ausbilder	Leitbilder, Feedbackkultur, Qualitätssicherungsmaßnahmen	Jahresberichte, Absolventenstatistik, Public Relations
Makroebene	Regeln, Gesetze und Verordnungen	Nationale Standards und Setzungen von Politik und Verbänden	Nationale und internationale Vergleiche von Indikatoren, Benchmarks

Auch der Diskurs über Qualität und Qualitätssicherung in der beruflichen Bildung ist in diesen verschiedenen Feldern auffindbar. Er hat sich aus der ursprünglich betriebs- und innovationsbezogenen Perspektive als Kundenorientierung herausgelöst und schließt heute bildungspolitische Maßnahmen mit ein. Ebenso wurde die traditionelle Fokussierung auf Inputsteuerung auf eine Prozess- und Outputdimension ausgeweitet.

Allerdings gab es bereits historisch zu Zeiten der Entstehung der Berufsbildungssysteme Ende des 19. Jahrhunderts international angelegte Struktur- und Leistungsvergleiche, die die Wirkungen von Berufsbildungsreformen überprüften. Im Lichte der Globalisierung und Individualisierung wird Bildungsqualität heute umgekehrt kleineren Einheiten überantwortet und zur Managementaufgabe der Betriebe bzw. Bildungsinstitutionen transformiert (Gonon 2003, S. 15ff.). Die Verlagerung der Qualitätssicherung lässt sich daher auch als Abkehr eines Top-down-Ansatzes bestimmen, das heißt die Qualitätssicherung wird zur Aufgabe, an welcher alle Beteiligten partizipieren sollen (Gonon 2005).

Im Folgenden werden Qualität und Qualitätssicherung im Einzelnen wie auch in ihren Verknüpfungen auf der Meso-, Mikro- und Makroebene dargestellt.

## **2. Mesoebene: Qualität als organisationsorientiertes Managementkonzept**

Die klassische Domäne der Qualitätssicherung auch in der beruflichen Bildung ist diejenige, die sich auf Bildungseinrichtungen als Organisationen bezieht. Qualität sollte sich hierbei an den Aufgaben eines betrieblichen Controllings ausrichten (Berset 2001). In den 1990er Jahren zentrierte sich diese Frage auf die Einführung von Qualitätssicherungssystemen auf die Sekundarstufe II (Gonon et al. 1998). Die traditionelle Nähe zur Berufs- und Arbeitswelt legte zunächst, zu Beginn der 1990er Jahre, eine Übertragung des ISO-Konzepts auf Berufsschulen nahe. Aber auch andere Qualitätssicherungsverfahren, wie das EFQM (European Foundation for Quality Management) sollten durch Pilotprojekte oder Modellversuche erprobt und wissenschaftlich evaluiert werden. Dass hierbei nicht uneingeschränkt das gesamte Instrumentarium auf Schulen und die Lehrpersonen, deren Zentrum des beruflichen Wirkens im Unterrichten besteht, zu applizieren sei, wurde immer wieder betont (so Tenberg 2003). Daher richtete sich das Augenmerk auf spezifisch zu adaptierende Konzepte, Verfahren und Instrumente zum Aufbau eines Qualitätsmanagements, so in der Schweiz z.B. als „Qualität durch Evaluation und Entwicklung (Q2E)“ (Landwehr/Steiner 2003).

Bereits in den ersten Diskussionsbeiträgen wurde festgehalten, dass die Einführung von Qualitätsverfahren die Berufsschulen veränderte. Eine konkrete Zuordnung von Maßnahmen und Wirkungen sei aber ebenso schwer zu eruieren, wie die bekannte Fragestellung, wie sich die Bedingungen eines Pilotprojektes auf den „Normalbetrieb“ oder andere Schulen übertragen ließen (Hagmann/Ebner 2002, S. 242). Diese Umstellung auf Qualität schaffte gerade in Schulen neuartige „Konfliktzonen“ (Altrichter 2000).

EFQM als verordnetes, und seit 2005 verpflichtendes Qualitätsmanagementsystem an den Berufsschulen in Niedersachsen, berge, so die ersten Erfahrungen, viele Unschär-

fen, die durch konsensorientierte und pragmatische weitere Konkretisierungen zu erfolgreichen Umsetzungen führen könnten (Szewczyk/Alexander 2006, S. 231). Es ginge jedoch nicht nur um Bewertungsaspekte, sondern man sollte im Sinne eines „gestaltungsorientierten Qualitätsmanagement(s)“ insbesondere die Ausbildungsqualität als gestaltbar betrachten (Spöttl/Becker 2006, S. 59). Eine repräsentative Befragung, die allerdings schwerpunktmäßig Weiterbildungsinstitutionen erfasste, weist auf die starke thematische Präsenz der Qualitätssicherung hin, der aber eine deutlich weniger weit fortgeschrittene Umsetzung entspreche (Bötel/Krekel 2004, S. 40). Immerhin wird eine Verbesserung der Arbeitsabläufe und eine höhere Zufriedenheit der Mitarbeitenden diagnostiziert (Gnahs/Kuwan 2004, S. 44). Man kann also der Qualitätssicherung attestieren, dass sie vor allem in den Berufsschulen ein Thema geworden ist.

### **3. Mikroebene: Qualitätsentwicklung des Lehrens und Lernens und Standardisierung des Unterrichts**

Die innovative Funktion prozessorientierter Qualitätsmodelle ist darin zu sehen, den Blick auf die organisationsbezogenen Prozesse des Bildungsgeschehens zu legen. Gleichzeitig sei aber der Stellenwert des professionellen pädagogischen Handelns, vor allem die Lehr-Lern-Interaktion in den Hintergrund getreten und die Gefahr bestehe, diese Differenz auszublenden (Jütte 2004, S. 243). Gegenüber einer solchen Kolonisierung im Geiste des Postfordismus soll daher, so einige Pädagogen, die (traditionelle) Professionalität der Lehrenden mit ihren Entscheidungsspielräumen verteidigt werden (Hodkinson 1998). Erziehung und Unterricht als Managementproblem zu betrachten, wirkt insofern anrühlich, obwohl auch in der Schul- und Unterrichtsdiskussion Effizienz- und Wirkungsfragen an Bedeutung gewonnen haben (Gonon 2003a). Immerhin ist die Thematisierung des Lehrens und Lernens unter dem Begriff der Qualitätsentwicklung statt demjenigen der Effizienzsteigerung im pädagogischen Diskurs akzeptanzfähiger. Autoren von Handbüchern und Ratgebern rechtfertigen ihre Tipps und Vorschläge für erfolgreichen Unterricht und Selbstreflexion daher gerne mit Qualitätsbemühungen (vgl. Crittin 2003). Eine solche Perspektive der Qualitätsentwicklung prägt auch die argumentative Einbettung der Berufsschulunterrichtsforschung und des Lehrerhandelns.

Seeber und Squarra verstehen ihre durchgeführte standardisierte Schülerbefragung als Beitrag zur Analyse der Unterrichtsqualität und Schulentwicklung (2003, S. 15). Die „Qualität der Unterrichtsarbeit“ werde, so ein anderer Befund, durch die Evaluation der Lehrenden durch die Beurteilungen der Lernenden erhöht (Greimel 2002). Auch im Betrieb sei für die Qualität des informellen Lernens wesentlich, organisatorisch und technisch lernförderliche Arbeitsplätze zu gestalten, die auch eine partizipative Personalentwicklung im Unternehmen ermöglichen (Markert 2006, S. 55).

Im Zentrum der neueren Qualitätsdiskussion im Bereich des Lehrens und Lernens steht darüber hinaus die Frage der Standards. Ausgehend von den Ergebnissen der international vergleichenden PISA-Studie wird auch für die Berufsbildung eine solche gefordert und als „machbar“ betrachtet (Baethge et al. 2006). Qualitätsstandards seien

für den Unterricht zentral (Fach-Overhoff 2004, S. 255). Die Festlegung und Einführung „instruktorischer Standards“ sollen helfen, gerade auch in der Berufsbildung Qualitätsmanagement und -sicherung über organisationelle Ebenen hinauszuführen (Ebner 2006, S. 184ff.).

Unter dem Titel „Qualität der beruflichen Bildung“ sammeln Oser und Kern eine Reihe von Teilfragestellungen, um ausgehend von den „Akteuren“ der Lehr-Lernprozesse „die Handlungskompetenz von Berufsbildungsverantwortlichen in Schule und Betrieb“ zu beschreiben, zu erklären und zu verbessern (Oser/Kern 2006). Lehrerhandeln soll zum Gegenstand der Qualitätsentwicklung gemacht werden, hierbei soll via Standards der Output gesteuert werden (vgl. dazu kritisch: Beck 2006). Gerade die weit vielfältigere Einbindung der Berufsbildung in verschiedene Lernorte und Berufsfelder lässt jedoch die Entwicklung von Standards und darauf Bezug nehmender Kompetenzmodelle als schwierig erscheinen, wenn sie sich auf das Lehren und Lernen auswirken sollten (Hallmann 2006). Immerhin wird auch die Güte des Lehrerhandelns und des Unterrichts in den Rahmen einer umfassenderen Qualitätsperspektive eingebettet.

#### **4. Makroebene: Europäische Union und nationale Bildungspolitiken: Qualität als Steuerungsvorgabe**

Die „Förderung der Qualität“ ist seit Beginn der Jahrtausendwende eines der herausragenden berufsbildungspolitischen Ziele der Europäischen Gemeinschaft. Im Hinblick auf die globalen Herausforderungen setzte sich der Europäische Rat im März 2000 in Lissabon als umfassendes – und auf 10 Jahre anberaumtes – Ziel, die Qualität und Wirksamkeit der Systeme der allgemeinen und beruflichen Bildung zu erhöhen. Die EU-Lissabon-Strategie geht davon aus, dass für die weltweite Wettbewerbsfähigkeit der Beitrag der Erziehung zentral und entsprechende Reformanstöße auch weit herum zustimmungsfähig sind. Eine generelle Erhöhung der Beteiligung am Bildungssystem sei daher ebenso bedeutsam, wie ein Einbezug gut ausgebildeter und auf höchster Stufe qualifizierter Fachkräfte in die Arbeitswelt (Derose/Young 2005, S. 18f.).

Anvisiert ist ein erleichterter Zugang zur allgemeinen und beruflichen Bildung und eine Öffnung dieser Systeme, so dass die europäischen Bürger jeden Alters Zugang zu lebensbegleitenden Bildungsmaßnahmen haben, was auch eine Kompatibilität dieser Systeme bedinge. Als Instrument zur Herstellung einer „größeren Konvergenz im Bezug auf die wichtigsten Ziele der EU“ sei daher die „offene Koordinierungsmethode“ (Europäische Kommission 2000), die sich auf Indikatoren, Richtwerte sowie vergleichend-bewertende Verfahren der regelmäßigen Beobachtung und Evaluation stützt, anzuwenden.

Fortschritte im Bereich von fünf bestimmten „benchmark areas“, die europaweit u.a. sowohl eine Reduktion der „early school leavers“ bis 2010 auf eine Quote von 17% auf 10% und eine Anhebung der Weiterbildungsbeteiligung von 7,5% auf 12,5% anstreben, sind zwar erst ansatzweise erfolgt, dennoch bzw. gerade deshalb wird einer Verbesserung der Attraktivität und Qualität der Berufsbildung das Wort geredet (EU 2006).

Eine „Technical Working Group“ kümmert sich insofern um die Qualität in der Berufsbildung als sie bezüglich europäischer Kooperation beratend und konsensbildend agiert und damit auch einen Beitrag zur Schaffung und Qualitätssicherung des EQF, d.h. des Europäischen Qualifikationsrahmens leistet (European Commission 2004, S. 28). Ein „Work Programme“ des „European Network on Quality Assurance in VET“ soll dazu beitragen, dass auf freiwilliger Basis und in gemeinsamer Kooperation Daten erhoben werden können, die sowohl die Ausstattung wie auch die Lernergebnisse erfassen. Die Zielsetzung besteht darin, die Flexibilität („responsiveness“) der nationalen Berufsbildungssysteme zu erhöhen, die sich u.a. auch in der verbesserten Durchlässigkeit zum Höheren Bildungswesen zeige (European Commission 2006, S. 5). Die Qualitätsverbesserung wird an mehr Erfahrungsaustausch, Orientierung an „Benchmarks“ und Förderung von vertiefenden Analysen geknüpft. Im Rahmen der Lissabon-Strategie sei die Sicherung der Qualität zu unterstützen und zu verstärken, um die Arbeitslosigkeit zu bekämpfen und in der neuen wissensbasierten und globalisierten Wirtschaft bestehen zu können (Europäische Kommission 2000, S. 2). „Berufsbildungsqualität“ wird anhand dreier politischer Zielsetzungen bestimmt:

- 1) Beschäftigungsfähigkeit,
- 2) Kongruenz von Angebot und Nachfrage und
- 3) Verbesserung des Zugangs zur Berufsbildung.

Die Qualitätskontrolle, die zunächst input-orientiert erfolgte, soll künftig stärker auch Prozesse, Output und Kontexte berücksichtigen. Die hierbei einzubeziehenden harten und „weichen“ Qualitätsindikatoren richten sich nach diesen politischen Vorgaben (Seufried/Slickers/Lassnigg 2002).

Mit der erfolgten Schaffung eines „Common Quality Assurance Framework (CQAF) for VET in Europe“ soll sowohl ein Planungs- und Umsetzungsmodell, eine Methodologie für ein Assessment, ein Monitoringsystem und ein Messinstrument der Berufsbildungssysteme der Mitgliedstaaten ermöglicht werden, das zu erhöhter Effektivität, mehr Transparenz und gegenseitigem Vertrauen beitragen soll. Die Qualitätsbemühungen auf europäischer Ebene sollen darüber hinaus durch „Good Practice“-Beispiele angespornt werden (Bundesministerium für Bildung, Wissenschaft und Kultur 2006). Im markanten Gegensatz zu diesen Vorstößen auf europäischer Ebene zeigen sich die Berufs- und Wirtschaftspädagogik und teilweise auch einzelne Akteure der Berufsbildungspolitik betont skeptisch, denn die lernergebnisorientierte Perspektive wird als Gefahr und Unterminierung des dualen Berufsbildungssystems wahrgenommen (vgl. zur Übersicht Kraus 2007). Insgesamt fehle der europäischen Berufsbildungspolitik ein ausreichend „deliberatives“ Moment (Grollmann & Ruth 2006, S. 76). So sieht etwa Rauner die Angleichung an solche europäische Vorgaben kritisch. Die „Qualität und Effizienz der dualen Berufsbildung“ sei vielmehr dadurch gewährleistet, dass das bestehende duale Berufsbildungsmodell kostengünstiger und bezüglich Organisation und Durchführung flexibler zu gestalten sei (Rauner 2007).

Generell erfährt das Argument der Qualität auch auf nationaler Ebene eine stärkere Gewichtung. Das novellierte deutsche Berufsbildungsgesetz hebt die Bedeutung der

Qualitätssicherung hervor: es gelte die Qualitätsentwicklung der schulischen Berufsausbildung zu verstärken, gleichzeitig die Qualitätssicherung am Lernort Betrieb zu entwickeln und insgesamt die lernortübergreifende Qualität zu steigern. Die Bemühungen zur Verbesserung der Qualität der Berufsausbildung sollten auch zu einer Halbierung der Ausbildungsabbrecherquote führen (Catenhusen 2007, S. 6).

Im Rahmen nationaler „Qualitätssicherung“ werden für ein „Antizipationssystem“ vorhandene Daten bezüglich Angebot und Nachfrage für die Berufsbildung Österreichs – auf nationaler wie europäischer Ebene – integriert, um zu erreichende und auf Indikatoren beruhende Zielsetzungen umzusetzen (Lassnigg/Markowitsch 2005, S. 232). Auch auf der Ebene der Berufsschulen ist in Österreich der Einfluss des gesamteuropäischen Qualitätsprozesses sichtbar. Die im Jahre 2005 begonnene Implementierung des QUIBB (Qualitätsinitiative Berufsbildung) will nationale Bildungsstandards für die berufliche Bildung etablieren und berufsfeldbezogenen Kompetenzen gemäß dem Europäischen Qualitätsrahmen für den Unterricht definieren (Timischl 2006).

Neuere bildungspolitische Vorstöße auf nationaler und internationaler Ebene figurieren demgemäß unter dem Label der „Qualitätssicherung“.

## **5. Qualität und Qualitätssicherung aus einer Forschungsperspektive**

Die Fachliteratur zur Qualitätssicherung ist einerseits unermesslich, andererseits stark anwendungsbezogen. Die überwiegende Mehrzahl der Veröffentlichungen sind Anleitungen und Ratgeber mit Leitfäden und Checklisten (z.B. Schöni, Tomforde & Wicki 1997) oder aber Erfahrungsberichte, Dokumentationen und Stellungnahmen. Einzelne Beiträge sind kritisch-reflexiv, indem sie die begrifflichen Unschärfen der Qualität und Qualitätssicherung und deren Varianzen thematisieren (so Harvey/Green 2000). Einige bilanzierende Berichterstattungen sind die Folge von Begleitstudien (vgl. Diesner et al. 2004). Insgesamt fehlt jedoch bis anhin eine umfassende Forschung im Bereich der beruflichen Bildung, welche die Qualitätssicherung und Qualitätsentwicklung selbst evaluiert.

Eine qualitative Studie geht auf die Einführung von Qualitätssicherung und ISO, zu welcher auch die US-amerikanischen berufsbildenden Schulen angehalten seien, ein. Neben Vorteilen in der Teamentwicklung sieht Bevans-Gonzales, die vergleichend Fokusgruppen befragte, vor allem die Einführung als zeit- und papierintensiv an (2004, S. 177f.). Die Interpretation und Anwendung von Standards sei schwierig und zu wenig auf die Besonderheiten im Erziehungsbereich eingerichtet. Damit bestätigt sie ähnliche Aussagen, wie sie bereits für die europäische Berufsbildung van den Berge (1997) und für Deutschland Kuwan/Waschbüsch (1996) festhielten. Im Umfeld der Schulentwicklungsforschung, die zu standardisierten Befragungen anregte, ergab sich auch zum Leitungshandeln an 14 Berliner Berufsschulen eine Befragung, die zum Schluss kommt, dass alle Funktionsträger einer Schule im Sinne einer Erfolgsorientierung einzubinden seien (Seeber 2003, S. 177). Schneeberger/Petanovitsch/Nowak (2006) untersuchten unter dem Signum der Qualitätssicherung der österreichischen beruflichen Bildung mit

Fragebögen in 6 ausgewählten Lehrberufen die Einstellung der Berufsschulen und Betriebe hinsichtlich der Optimierung der Kooperation. Die Zufriedenheit mit den Leistungen und Ergebnissen der Berufsschule steige, bei kontinuierlicher Information und Aufbau persönlicher Kontakte (ebd., S. 10ff.). Als explorative Studie ist auch die Untersuchung von Richter zu bezeichnen, welche die Erwartungen der Unternehmen an das Qualitätsmanagement in der Berufsschule mittels postalischer Befragung eruierte (Richter 2006, S. 183ff.).

Eine international vergleichende Studie hebt hervor, dass in den untersuchten Ländern Australien, Vereinigtes Königreich und Deutschland, unterschiedliche Formen der Qualitätssicherung entstanden, die allerdings nach gleichen Prinzipien funktionieren. Es sei eine Konvergenz in dem Sinne feststellbar, dass Standards für Bildungsanbieter und Auditoren der Akkreditierung sich etablieren, die sowohl eine Außensicht wie auch „Self-Assessment“ einschließen (Misko 2006). Eine branchenspezifische Gegenüberstellung der Rolle der Qualitätssicherung im Zusammenhang mit der beruflichen Bildung in der Automobilindustrie in Korea und Deutschland stellt eine stärkere Einbindung der Qualitätssicherung mit der beruflichen Bildung und Weiterbildung für die deutschen Betriebe heraus, im Gegensatz zu einer eher an der Produktion orientierten Qualitätssicherung in Korea (Lee et al. 2006).

Die dominante Form der wenigen empirischen Studien geht in dem Sinne evaluativ vor, als formulierte Zielsetzungen des Qualitätsverfahrens geprüft werden. Insofern folgen sie dem „klassischen“ Anspruch einer kunden- und produktbezogenen Ergebnisqualität, die sich auf die Erreichung der gesetzten Ziele und die Kundenzufriedenheit konzentriert. Weniger im Blick ist hierbei eine wirkungsorientierte Evaluation. Diejenige Perspektive herrscht vor, die die organisationelle und professionelle Entwicklung der Bildungsinstitution und der Lehrenden, also die Mesoebene, im Blick hat.

## 6. Bilanz

Qualität als organisationelles Bildungsmanagement, als Entwicklungs- und Standardisierungsmodell des Lehrerhandelns und Unterrichtens, sowie als bildungspolitische Steuerungsvorlage auf nationaler und internationaler Ebene – dies sind die Facetten der Qualitätssicherung in der beruflichen Bildung. Auch gut ein Jahrzehnt nach der ersten Welle der Qualitätssicherung besteht zur „Qualität“ der Qualitätssicherung in der beruflichen Bildung wenig ausgewiesene Evidenz zu den günstigen Wirkungen und kaum empirische Überprüfung. Wahrscheinlich kann der Qualitätssicherung ein Verdienst zugeschrieben werden: sie hat eine anschlussfähige Sprache mit international konvertibler Währung geschaffen. Darüber hinaus wurde Qualität als Chiffre für Reformbereitschaft sowohl in den Mikrostrukturen des Lernens wie auch in der großen Bildungspolitik als eine verbindende Komponente eingeführt. Die beständige und explizit formulierte Sorge um „Qualität“, die nach fortlaufender Veränderung und Fortbildung ruft, nach Neugestaltung und Zielerreichungsvereinbarungen, ist das spezifische Organisationsmerkmal das auch Schulen und die Berufsbildung insgesamt erfasst. Eine Ergeb-

nisorientierung mit impliziter Standardisierungstendenz bildet sich als die „organisierende Idee“ (Power 2007, S. 51) heraus, die sowohl Mikro-, Meso- wie Makroebene der Bildung umfasst.

Diese Ausrichtung auf Qualität hat das Berufsbildungswesen zweifellos verändert, es quasi in beständige Selbstalarmierungsbereitschaft gestellt. Residuen selbstbestimmter und eigensinniger Initiativen wurden hierbei einem stromlinienförmigen und von Bürokratien und Bildungsapparaten gesteuerten Kontrollregime ausgesetzt, bzw. es wurden undefinierte Räume durch Maßnahmen der Qualitätsentwicklung explizit gemacht.

Qualität ist auch beschreibbar als Übergang von einer auf bürokratischer Disziplinierung ausgerichteten Bildungsorganisation hin zu einer Kontrollgesellschaft, in der wir – in den Begriffen Foucaults – je frei zirkulieren und uns selbst kontrollieren: alle Beteiligten setzen sich hierbei einem Qualitätstribunal aus (Simons 2002). Die Qualitätssicherung in der beruflichen Bildung lässt sich auch als Paradoxie auffassen, die darin besteht, Reputationsverluste und Risiken zu minimieren, indem man Qualität zur Legitimationstechnik transformiert, die gleichzeitig jedoch wieder neue Risiken erzeugt.

## Literatur

- Altrichter, H. (2000): Konfliktzonen beim Aufbau schulischer Qualitätssicherung. In: Zeitschrift für Pädagogik, 41, Beiheft, S. 93–110.
- Baethge, M./Achtenhagen, E./Arends, L./Babic, E./Baethge-Kinsky, V./Weber, S. (2006): Berufsbildungs-PISA. Machbarkeitsstudie. Stuttgart: Steiner.
- Bailey, Th./Hughes, K./Barr, T. (2000): Achieving Scale and Quality in School-to-Work Internships: Findings from Two Employer Surveys. In: Educational Evaluation and Policy Analysis, 22, pp. 41–64.
- Bakic, J. (2006): Qualitätsmanagement. In: Dzierzbicka, A./Schirlbauer, A. (Hrsg.): Pädagogisches Glossar der Gegenwart. Von Autonomie bis Wissensmanagement. Wien: E. Löcker, S. 218–227.
- Beck, K. (2006): Standards – ein Mittel zur Qualitätsentwicklung in der Lehrerbildung? In: Wirtschaft und Erziehung (WuE), 2, S. 44–54.
- Berset, J.-E. (2001): Introduction d'un Controlling éducationnel dans la formation professionnelle suisse. Fribourg: Editions universitaires.
- Bevans-Gonzales, Th. (2004): The Strengths and Weaknesses of ISO 9000 in Vocational Education. In: Journal of Vocational Education and Training (JVET), 56, S. 163-180.
- Bötel, C./Krekel, M. (2004): Trends und Strukturen bei Bildungsträgern. In: Balli, C./Krekel, E./Sauter, E. (Hrsg.): Qualitätsentwicklung in der Weiterbildung – Wo steht die Praxis? (Bundesinstitut für Berufsbildung – Berichte zur beruflichen Bildung, Heft 262). Bielefeld: Bertelsmann, S. 19–40.
- Bundesministerium für Bildung, Wissenschaft und Kultur (2006): Quality in Education and Training. Cases of Good Practice in Vocational Education and Training and Higher Education. Vienna: bmbwk.
- Catenhusen, W. (2007): Die Zukunftsfähigkeit des deutschen Bildungssystems sichern. Politische Stellungnahme im Namen des Netzwerk Bildung anlässlich der Präsentation der neuen Studie der Friedrich-Ebert-Stiftung „Berufsbildung im Umbruch – Signale eines überfälligen Aufbruchs von Prof. Dr. Martin Baethge“. Berlin: Netzwerk Bildung – Friedrich Ebert Stiftung, 26. Januar 2007.
- Crittin, J.-P. (2003): Erfolgreich unterrichten. Die Vorbereitung und Durchführung von lernwirksamen Schulungen. (4. überarbeitete Auflage). Bern: Haupt.

- Derose, S./Young, D.: Overview (2005): Quality and efficiency of education and training in Europe. In: European Commission (ed.). *European Economy – Quality and efficiency in education* (Special report No.3/2004). Brussels: Directorate General for Economic and Financial Affairs, S. 6–21.
- Diesner, I./Euler, D./Walzik, S./Wilbers, K. (2004): Abschlussbericht des Modellversuchsprogramms KOLIBRI (Kooperation der Lernorte in der beruflichen Bildung) 09/1999 bis 02/2004. In: *Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK): Abschlussbericht des Programmträgers zum BLK-Programm*. (Materialien zur Bildungsplanung und Forschungsförderung. Bonn: BLK Geschäftsstelle.
- Ebner, H. (2006): Standards als Instrumente des Qualitätsmanagements im Bildungsbereich. In: Eckert, M./Zöller A. (Hrsg.): *Der europäische Berufsbildungsraum – Beiträge der Berufsbildungsforschung*. 6. Forum der Arbeitsgemeinschaft Berufsbildungsnetz (AG BFN). Bielefeld: Bertelsmann, S. 179–188.
- European Commission (2004): ‚Copenhagen Process‘ Quality Assurance in VET. Technical Working Group Progress Report.
- European Commission (2006): *European Network on Quality Assurance in VET. Work Programm 2006–2007*.
- Europäische Kommission (2000): *Qualität der Berufsbildung*. R: /Ve-ETV/ETV html/Doc online/Themes-II/Qualite/Marques-Oliveira-DE.doc (27.11.2007).
- EU (2006): *The Helsinki Communiqué on Enhanced European Cooperation in Vocational Education and Training*. Communiqué of the European Ministers of Vocational Education and Training, the European Social Partners and the European Commission to review the priorities of the Copenhagen Process. Helsinki, 5. 12. 2006.
- Fach-Overhoff, M. (2004): Qualitätsstandards für die Unterrichtspraxis. In: Rützel, J./Bendig, B./Herzog, C./Sloane, P.F. (Hrsg.): *Berufsbildung in der globalen NetzWerkGesellschaft: Quantität – Qualität – Verantwortung*. Bielefeld, S. 237–260.
- Foster, H. (2000): Mit Qualitätskriterien zu mehr Chancengleichheit in der Weiterbildung: Das QUASAR-Projekt. In: Rützel, J. (Hrsg.): *Entwicklung und Umsetzung von Qualitätsstandards in der Berufsbildung*. Bertelsmann, S. 31–38.
- Gnahn, D./Kuan, H. (2004): Qualitätsentwicklung in der Weiterbildung – Effekte, Erfolgsbedingungen und Barrieren. In: Balli C./Krekel E./Sauter, E. (Hrsg.): *Qualitätsentwicklung in der Weiterbildung – Wo steht die Praxis?* (Bundesinstitut für Berufsbildung – Berichte zur beruflichen Bildung, Heft 262). Bielefeld: Bertelsmann, S. 41–59.
- Gonon, Ph./Hügli, E./Landwehr, N./Ricka, R./Steiner, P. (1998): *Qualitätssysteme auf dem Prüfstand*. Aarau: Sauerländer.
- Gonon, Ph. (2003): Qualität als „Qualitätssicherung“ – eine Herausforderung für die (Berufs-) Schule? In: Zöller, A./Gerds, P. (Hrsg.). *Qualität sichern und steigern. Personal- und Organisationsentwicklung als Herausforderung für berufliche Schulen*. Bielefeld: Bertelsmann, S. 11–24.
- Gonon, Ph. (2003a): *Erziehung als Managementproblem. Bildungsinstitutionen zwischen Charisma und Taylorismus*. In: Mangold, M./Oelkers, J.:(Hrsg.): *Demokratie, Bildung und Markt*. Bern: Peter Lang, S. 281–301.
- Gonon, Ph. (2005): *Partizipative Qualitätssicherung*. In: Rauner, F. (Hrsg.): *Handbuch Berufsbildungsforschung*. Bielefeld: Bertelsmann, S. 421–426.
- Grollmann, Ph. & Ruth, K. (2006): *Zertifizierung, Innovationen und Qualitätsentwicklung in der beruflichen Bildung in Europa*. In: Clement U./Le Mouillour I./Walter, M. (Hrsg.): *Standardisierung und Zertifizierung beruflicher Qualifikationen in Europa*. Bonn: Schriftenreihe BIBB, S. 61–78.
- Greimel, B. (2002): *Lehrerevaluation durch Beurteilungen der Lernenden – eine Analyse des Standes der Evaluationsforschung*. In: *Zeitschrift für Berufs- und Wirtschaftspädagogik (zbw)*, 98, S. 197ff.
- Hagmann, S./Ebner, H. G. (2002): *Qualitätskonzepte auf dem Weg in die beruflichen Schulen – erste Ergebnisse einer Evaluationsstudie*. In: Eckert, M./Horlebein, M./Lisop I. Reinisch, H./Tramm T. (Hrsg.): *Bilanzierungen. Schulentwicklung, Lehrerbildung und Wissenschaftsgeschichte im Feld der Wirtschafts- und Berufspädagogik*. Frankfurt: GABF-Verlag, S. 223–243.

- Hallmann, P. (2006): Nationale Bildungsstandards – Motor der Qualitätsentwicklung auch an beruflichen Schulen oder nur Lernziel-Retrospektive? In: *Wirtschaft und Erziehung (WuE)*, 1, 2006, S. 20–25.
- Harvey, L./Green, D. (2000): Qualität definieren. Fünf unterschiedliche Ansätze. In: *Zeitschrift für Pädagogik*, S. 17–40.
- Hartz, S. (2004): Qualität in der Weiterbildung: Die Perspektivengebundenheit von Qualitätsanforderungen am Beispiel der Differenz von Organisation und Profession. In: Fröhlich, W./Jütte, W. (Hrsg.): *Qualitätsentwicklung in der postgradualen Weiterbildung. Internationale Entwicklungen und Perspektiven*, Münster: Waxmann, S. 231–246.
- Hodkinson, Ph. (1998): Technicism, Teachers and Teaching Quality in Vocational Education and Training. In: *JVET (Journal of Vocational Education and Training)*, 50, S. 193–208.
- Jütte, W. (2004): Qualitätsmanagement und betriebliche Weiterbildung. In: Gonon, Ph./Stolz, S. (Hrsg.): *Betriebliche Weiterbildung – empirische Befunde, theoretische Perspektiven und aktuelle Herausforderungen*. Bern: hep, S. 229–246.
- Landwehr, N./Steiner, P. (2003): *Q2E – Qualität durch Evaluation und Entwicklung*. Bern: hep.
- Lassnigg, L./Markowitsch, J. (Hrsg.) (2005): *Qualität durch Vorausschau. Antizipationsmechanismen und Qualitätssicherung in der österreichischen Berufsbildung*. Innsbruck: Studien Verlag.
- Lee, Y./Lee, D./Härtel, M./Zinke, G. (2006): *Qualification Strategies and New Media for Quality Assurance in Manufacturing – the example of the automotive industry in Korea and Germany*. (Discussion Paper No. 79). Bonn: Bundesinstitut für Berufsbildung.
- Markert, W. (2006): Qualitätsbestimmung zwischen Organisationsentwicklung und lernendem Subjekt. In: W. Markert (Hrsg.): *Qualität des beruflichen Lernens in der Weiterbildung*. Baltmannsweiler: Schneider Verlag Hohengehren, S. 48–56.
- Misko, J. (2006): *Vocational education and training in Australia, the United Kingdom and Germany*. Adelaide: NCVER (National Centre for Vocational Education Research).
- Muller, D./Funnell, P. (1992): An Exploration of the Concept of Quality in Vocational Education and Training. In: *Educational and Training Technology International*, 29, p. 257–261.
- Kraus, K. (2007): Die „berufliche Ordnung“ im Spannungsfeld von nationaler Tradition und europäischer Integration. In: *Zeitschrift für Pädagogik*, 53, S. 382–398.
- Kuwan, H./Waschbüsch E. (1996): *Zertifizierung und Qualitätssicherung in der beruflichen Bildung*. Bielefeld: Bertelsmann.
- Nickolaus, R. (2007): Qualität in der Beruflichen Bildung. In: *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 103, S. 161–166.
- Oser, F./Kern, M. (Hrsg.) (2006): *Qualität der beruflichen Bildung – Eine Forschungsbaustelle*. Bern: hep
- Power, M. (2007): *Organized Uncertainty. Designing a World of Risk Management*. Oxford: University Press.
- Prechtel, C./Scherbe, N. (2007): Ausbildung im Verbund: Steigerung von Qualität und Effizienz. In: Severing, E./Loebe, H. (Hrsg.): *Effizienz in der Ausbildung. Strategien und Best-Practice-Beispiele*. Bielefeld: Bertelsmann, S. 69–78.
- Rauner, F. (2007): *Kosten, Nutzen und Qualität der beruflichen Ausbildung*. Bremen: Universität, Institut Technik & Bildung (ITB), Forschungsberichte 23/2007.
- Richter, A. (2006): *Qualitätsmanagement als Gegenstand der Berufsbildung*. Bielefeld: Bertelsmann.
- Rütters, K./Roggenbrodt, G./Künzel, J. (2004): Ziele und Aufgaben einer externen Evaluation für das berufsbildende Schulwesen. Überlegungen am Beispiel Niedersachsens. In: Rützel, J./Bendig, B./Herzog, C./Sloane, P.F. (Hrsg.): *Berufsbildung in der globalen NetzWerkGesellschaft: Quantität – Qualität – Verantwortung*. Bielefeld, S. 173–192.
- Schneeberger, A./Petanovitsch, A./Nowak, S. (2006): *Optimierung der Kooperation Berufsschule-Lehrbetrieb. Erhebungen und Analysen der Verbesserung der Qualität der Berufsbildung*. Wien: Bundesministerium für Bildung, Wissenschaft und Kultur.

- Schöni, W./Tomforde, E./Wicki, M. (1997): Leitfaden Bildungsqualität. Evaluation und Gestaltung der Bildungsarbeit in Betrieb und Büro. Zürich: Rüegger.
- Seeber, S. (2003): Schulmanagement zwischen strategischem Entscheiden und Routine. In: J. van Buer/O. Zlatkin-Toitschanskaia (Hrsg.): Berufliche Bildung auf dem Prüfstand. Entwicklung zwischen systemischer Steuerung, Transformation durch Modellversuche und unterrichtlicher Innovation. Frankfurt a. M.: Peter Lang, S. 157–179.
- Seeber, S./Squarra, D. (2003): Lehren und Lernen in beruflichen Schulen. Schülerurteile und Unterrichtsqualität. Frankfurt a. M.: Peter Lang.
- Seyfried, E./Slickers, C./Lassnigg, L. (2002): Indikatoren für eine europäische Strategie zur Förderung der Qualität der Berufsbildung (vorläufiger Bericht). Thessaloniki: Cedefop.
- Simons, M. (2002): Governmentality, Education and Quality Management. Towards a Critique of the Permanent Quality Tribunal. In: Zeitschrift für Erziehungswissenschaften, 5, S. 617–633.
- Smits, W. (2006): The Quality of Apprenticeship Training. In: Education Economics 14, pp. 329–344.
- Spöttl, G./Becker, M. (2006): Qualität in der beruflichen Bildung – Perspektiven für einen Handlungsrahmen. In: lernen & lehren, 82, S. 52–60.
- Szewczyk, M./Alexander, P. (2006): Qualitätsverbesserungen durch EFQM – ein Ansatz zur Stärkung der Selbsterhaltungsfähigkeit (Autopoiese) Berufsbildender Schulen. In: WuE (Wirtschaft und Erziehung), 8, S. 223–235.
- Tenberg, R. (2003): „Dienstleistung“ Unterricht? Unstimmigkeiten bei der Adaption betrieblicher Instrumente von Qualitätsmanagement an Schulen. In: Zöllner, A./Gerds, P. (Hrsg.). Qualität sichern und steigern. Personal- und Organisationsentwicklung als Herausforderung für berufliche Schulen. Bielefeld: Bertelsmann, S. 121–146.
- Timischl, W. (2006): Qualität als Schlüssel zum Erfolg in der Berufsbildung. In: wissenplus, 5, S. 8–9.
- Timmermann, D./Windschild, T. (1996): Implementierung von Qualitätsgrundsätzen und -maßstäben in der betrieblichen Bildung. Bielefeld: Bertelsmann, S. 79–90.
- Van den Berge, W. (1997): Application of ISO 9000 Standards to Education and Training: interpretation and guidelines in a European perspective. Thessaloniki: CEDEFOP.
- Zech, R. (2005): Lernerorientierte Qualitätstestierung in der Weiterbildung. LQW 2 – Das Handbuch. Hannover: ArtSet.

*Anschrift des Autors:*

Prof. Dr. Philipp Gonon, Universität Zürich, Institut für Gymnasial- und Berufspädagogik, Beckenhofstr. 35, CH 8006 Zürich, E-Mail: gonon@igb.uzh.ch

*Klaus Meisel*

## **Qualitätsmanagement und Qualitätsentwicklung in der Weiterbildung**

### **1. Der aktuelle Qualitätsdiskurs in der Weiterbildung**

Seit mittlerweile mehr als einem Jahrzehnt gibt es in der fachöffentlichen Auseinandersetzung der Weiterbildung einen kontinuierlichen Diskurs zu Fragen der Qualität, der Qualitätssicherung, der Qualitätsentwicklung und geeigneten Umsetzungsformen des Qualitätsmanagements (vgl. Gnahs 1996; Küchler/Meisel 1999; Heinold-Krug/Meisel 2002; Balli/Krekel/Sauter 2004; Hartz/Meisel 2006). Qualität ist nicht nur generell das Leitkonzept des allgemeinen Bildungsdiskurses (vgl. Terhart 2000, S. 809; Bühlow-Schramm 2006); die Debatte zur Qualität beherrschte in der Weiterbildung auch die inhaltlichen, organisationsbezogenen und ordnungspolitischen Auseinandersetzungen der letzten Jahre. Dabei zeigt sich, dass der Umgang mit dem Qualitätsbegriff nicht einheitlich ist. Unter Qualität wird häufig verstanden:

- das herausragende Praxisbeispiel (good practise),
- eine nachvollziehbare effiziente und kundenorientierte Gestaltung von Organisation und Prozessen,
- eine Weiterbildung, die in Inhalt und Methode dem Bedarf und den Teilnehmererwartungen gerecht wird.

Während Qualitätssicherung nur die Sicherung bereits vorhandener Qualität meint, zielt die Qualitätsentwicklung darüber hinaus auf einen Prozess der kontinuierlichen Reflexion der Bedingungen für Qualität und deren ständige Verbesserung. Dem Qualitätsmanagement wiederum liegt die Vorstellung zugrunde, dass Qualität systematisch durch ein bestimmtes Managementhandeln beeinflusst werden kann (vgl. Meisel 2001, S. 111f.).

Die Qualitätsdiskussion in der Weiterbildung wurde in den letzten Jahren auch auf unterschiedlichen Ebenen geführt (vgl. Kuper 2002, S. 539). Sie bezieht sich zum einen auf die Ebene des Weiterbildungssystems als Ganzes. Hier geht es um die Frage des gesellschaftlichen Bedarfs, der Systemausgestaltung und der Ressourcenbereitstellung sowie entsprechender ordnungspolitischen Rahmenbedingungen. Zum anderen bezieht sie sich auf die Organisationsebene, also in erster Linie auf die Organisationsstrukturen und -abläufe. Drittens gibt es einen Diskurs, der sich auf die Gestaltung der erwachsenpädagogischen Interaktion, also um die besondere qualitative Ausgestaltung des Lehrens und Lernens mit Erwachsenen konzentriert. Die hier vorgenommene grobe Restrukturierung des Qualitätsdiskurses in der Weiterbildung bedeutet nicht, dass die ge-

fürten Diskussionen sich eindeutig der einen oder anderen Ebene zuordnen ließen. Das Gegenteil ist der Fall: Die Schwerpunktsetzungen variieren zeit- und interessenbezogen. Die Auseinandersetzungen finden in den unterschiedlichen Akteursgruppen mit unterschiedlichen Akzentuierungen statt. So wird beispielsweise die Diskussion in der beruflichen Weiterbildung anders geführt als in der allgemeinen Weiterbildung. In der Familienbildung konzentrieren sich die Auseinandersetzungen auf andere Schwerpunkte (vgl. Schiersmann/Thiel/Pfizenmaier 2001) als etwa in der politischen Weiterbildung (vgl. Meisel 2005). Bei den Einrichtungsleitungen stehen primär Fragen der Institutionssicherung im Vordergrund. Für die praktisch-pädagogisch in der Weiterbildung Tätigen geht es eher um die professionelle Bewältigung des Paradigmenwechsels von der „Belehrungspädagogik“ zur „Ermöglichungsdidaktik“ (vgl. Arnold/Schüßler 2003). Die Weiterbildungsnachfrager haben wiederum noch einen anderen Blick auf die Qualitätsfragen. Je teurer für die Teilnehmenden die Angebote werden, desto mehr rücken Fragen nach der Qualität in den Mittelpunkt des Interesses (vgl. Meisel 2001, S. 110).

Grundsätzlich kann man eine inhaltliche, eine professionstheoretische, eine ökonomische und eine ordnungspolitische Orientierung in den Auseinandersetzungen unterscheiden (vgl. Hartz/Meisel 2006, S. 7f.):

- Inhaltlich geht es um Fragen des Selbstverständnisses, der Ziele und der Aufgaben der Weiterbildung im Kontext des lebenslangen Lernens.
- Professionstheoretisch steht das Problem der Sicherstellung und der Entwicklung der erwachsenenpädagogischen Qualität sowie eines professionsangemesseneren Handlungssystems von Qualitätsentwicklung im Vordergrund.
- Ökonomisch geht es um Marktanteile, Finanzierungsmöglichkeiten und Effizienz.
- Ordnungspolitisch werden Fragen der Transparenz des Verbraucherschutzes, der Anerkennung und der finanziellen Förderung behandelt.

In Anlehnung an Luhmann und Schorr interpretiert Kuper (2002, S. 536) das erstarrende Interesse an Qualität als eine semantische Begleitung tiefgreifender struktureller Veränderungen im Bildungssystem. Eine solche Einschätzung wird auch von den Ergebnissen einer Delphi-Studie zu Veränderungen im Bildungswesen bestätigt, wonach die Behandlung der Qualitätspolitik und des Qualitätsmanagements als ein Motor für die Veränderungen im Bildungswesen gesehen werden (vgl. Stock u.a. 1998, S. 76ff.). Bezogen auf die Weiterbildung deckt sich diese Interpretation mit einer früheren Feststellung von Faulstich (1991, S. 572), wonach eine Intensivierung der Qualitätsdiskussion in der Weiterbildung immer im Zusammenhang mit grundlegenden Gestaltungsfragen des Weiterbildungssystems zu sehen ist.

Der thematische Diskurs rund um die Qualitätssicherung, das Qualitätsmanagement und die Qualitätsentwicklung wurde also auf unterschiedlichen Ebenen geführt, aus unterschiedlichen Perspektiven aufgegriffen und dementsprechend inhaltlich anders akzentuiert. Dabei gibt es zahlreiche Wechselbezüge und Rückkoppelungseffekte zwischen den einzelnen Ebenen und den einzelnen Interessen. Darüber hinaus darf nicht übersehen werden, dass die weiterbildungsspezifischen Auseinandersetzungen zur

Qualitätsentwicklung in einem im Vergleich mit den anderen Säulen des Bildungssystems wenig professionalisierten Bildungssektor mit besonderen Spannungsfeldern geführt werden. Dabei handelt es sich zum einen um eine Verschiebung der Schnittstelle zwischen staatlicher Ordnungspolitik und Selbstregulierung des Marktes. Zwar agiert die Weiterbildung schon seit längerem in einem mehrfach geteilten Weiterbildungsmarkt (vgl. Friebel 1993), doch orientieren sich ihre Handlungsgrundsätze immer stärker an Marktprinzipien, weshalb marktorientierte Qualitätsdimensionen (z.B. Tests und Qualitätssiegel) an Bedeutung gewinnen. War es früher für Anerkennung und Finanzierung bedeutsam, dass eine Einrichtung eine staatliche Anerkennung hatte, so ist es heute entscheidender, ob sich eine Weiterbildungseinrichtung einer externen Zertifizierung unterzogen hat. Stand also früher dem Staat quasi ein Monopol der Anerkennung zu, ist dieses Recht nun tendenziell an private Zertifizierungsgesellschaften übergegangen. Nicht zuletzt deshalb muss auch auf interne Spannungsfelder in den Weiterbildungseinrichtungen verwiesen werden. Einerseits wird gerade von öffentlichen Zuwendungsgebern eine erhöhte Wirtschaftlichkeit in der Arbeit gefordert und die öffentliche Finanzierung erheblich reduziert (vgl. Kuhlenkamp 2003, S. 25ff.), andererseits erfordern Qualitätsentwicklungsprozesse in der Regel einen zusätzlichen Ressourcenbedarf (vgl. Schöll 1996, S. 92ff.). Die strategischen und konzeptionellen Auseinandersetzungen mit der Qualitätsentwicklung sind also eingebettet in einen allgemeineren Strukturwandel der Weiterbildung (vgl. Meisel 2007, S. 106ff.). Da die Diskussionen aus unterschiedlichen Perspektiven und mit differierenden Interessen geführt werden, erscheint die Qualitätsdebatte in der Weiterbildung häufig diffus und unübersichtlich.

## **2. Rückblick: Mehr Qualität durch Professionalisierung**

Die Bemühungen um eine Verbesserung der Weiterbildungsqualität sind älter als der aktuelle Qualitätsdiskurs. Sie waren früher nur nicht immer mit dem Qualitätsbegriff verbunden. Eher waren Qualitätsanstrengungen mit Professionalisierungsanstrengungen verknüpft (vgl. Giesecke 1997). Beispielsweise waren mit dem Ende der siebziger Jahre des letzten Jahrhunderts entwickelten Baukastensystemen im fremdsprachlichen und naturwissenschaftlichen Unterricht zur Förderung des systematischen Lernens Erwachsener immer auch gezielte Anstrengungen zur Evaluation verbunden (vgl. Gerl/Pehl 1983). Die Einrichtung der Diplomstudiengänge oder die Entwicklung von Selbststudienmaterialien für neben- und hauptberufliche pädagogische Mitarbeitende in der Weiterbildung waren Ausdruck zunehmender Bemühungen um die Qualifizierung Personals, also auch um die Qualität des professionellen Handelns. Professionalisierungsbemühungen einerseits und die Anstrengungen zur Qualitätsentwicklung andererseits waren in der Weiterbildung eng miteinander verbunden. Erhard Schlutz hat darauf verwiesen, dass – professionstheoretisch gesehen – Qualitätssicherung ein Ziel aller Professionalisierungsbemühungen gewesen sei und insofern Professionalität und Qualität in einem engen Verhältnis zueinander stehen. „Mehr Qualität durch Professionalisierung und Professionalität meint

1. mehr Programmqualität durch theoretische Reflexion, Bedarfsermittlung, Transparenz, Systematik, curriculare Bausteine mit Fach- und Wissenschaftsorientierung;
2. mehr Durchführungsqualität durch Teilnehmerberatung, Dozentenauswahl und -fortbildung, Begleitmaterialien, Kurshospitationen usw.;
3. mehr Ergebnisqualität durch operationalisierte Tests und Prüfungen, durch Befragungen der Beteiligten und Programmevaluation“ (Schlutz 1996, S. 28).

Auch für die Weiterbildungspolitik sind die Qualitätsfragen im Grunde genommen nichts gänzlich Neues. Auf öffentlich diskutierte Qualitätsmängel wurde je nach Zuständigkeit punktuell mit Gesetzen und Erlassen reagiert. So hatte eine öffentliche Kritik an mangelnder Fernunterrichtsqualität den Staat bereits in den siebziger Jahren mit dem Fernunterrichtsgesetz (1974) zu einem ordnungspolitischen Eingriff veranlasst. Auch die Bundesagentur für Arbeit hat mit umfangreichen Träger- und maßnahmebezogenen Anforderungskatalogen in Erlassen versucht die Qualität in der außerbetrieblichen beruflichen Weiterbildung zu sichern (vgl. Bundesanstalt für Arbeit 1999). Trotz zusätzlicher Qualitätsüberprüfungen ist es ihr aber nicht gelungen, erhebliche Qualitätsmängel in der beruflichen Weiterbildung in den neuen Ländern gleich nach der deutschen Einigung zu verhindern (vgl. Sauter 1992). Vor dem Hintergrund dieser Situation hat das Bundesministerium für Wirtschaft und Arbeit im Jahr 2002 im Rahmen des Sozialgesetzbuches eine entsprechende Gesetzesänderung ausgearbeitet, nach der Einrichtungen der beruflichen Weiterbildung eine Qualitätszertifizierung nachweisen müssen (§§84ff. SGB III). Die nun gültige Anerkennungs- und Zulassungsverordnung – Weiterbildung (AZWV) vom 16. Juni 2004 legt das Zulassungsverfahren von Trägern und Maßnahmen der beruflichen Bildung nach dem Dritten Sozialgesetzbuch fest. Demnach müssen sowohl Träger als auch Maßnahmen, die von der Bundesagentur für Arbeit gefördert werden, einer Prüfung durch so genannte fachkundige Stellen unterzogen werden. Die Anforderungen beziehen sich zum einen auf die Trägerebene. Sie müssen ein wirksames Qualitätsmanagementsystem und spezifische Anforderungen (z.B. kundenorientiertes Leitbild, Zusammenarbeit mit der Wirtschaft, Anforderungen an die Leiter und Mitarbeiter, Arbeitsmarktorientierung der Bildungsarbeit) nachweisen. Zum anderen beziehen sich die Anforderungen auf die Maßnahmeebene. Gemäß §9 Abs. 1 Punkt 2 AZWV müssen sich Maßnahmen an definierten Kriterien wie Einbindung der Maßnahme in arbeitsmarktrelevante, regionale Entwicklungen, Abschlussorientierung etc. messen lassen.

Qualitätssicherungsstrategien fanden auch schon immer Berücksichtigung in den Weiterbildungsgesetzen der Länder (vgl. Melms 2002). Eine Gemeinsamkeit der Ländergesetze ist es, den staatlichen Anspruch nach Qualität in der Weiterbildung zum Schutz der Lernenden zu bekräftigen. Die staatliche Anerkennung und der Zugang zur finanziellen Förderung setzt in der Regel die Einhaltung professioneller Mindeststandards voraus. Hierzu gehören u.a. offen zugängliche Programme, Offenlegung der Arbeitsinhalte und der Finanzierungsstrukturen, wirtschaftliche Leistungsfähigkeit, Programmbreite und notwendige sachliche Voraussetzungen, eine einschlägig qualifizierte hauptberufliche Leitung sowie qualifiziertes nebenberufliches Personal etc. (vgl. Gnahs

1999, S. 33). Weitergehende staatliche Interventionen lehnen die meisten Länder ausdrücklich ab. In einigen Gesetzen gilt der Teilnehmerschutz als Anerkennungsgrundsatz (Bremen, Mecklenburg-Vorpommern, Niedersachsen, Schleswig-Holstein und Thüringen). In Niedersachsen werden die Einrichtungen verpflichtet, ihre Arbeit extern evaluieren zu lassen, die Ergebnisse zu dokumentieren und sich spätestens nach vier Jahren erneut einer externen Evaluation zu stellen. Ein Teilnehmerschutzgesetz war in Hamburg in der Diskussion, wurde jedoch durch eine freiwillige Selbstkontrolle ersetzt. Diese wird vom „Verein Weiterbildung Hamburg e.V.“, in dem fast alle Hamburger Weiterbildungseinrichtungen organisiert sind, durchgeführt (vgl. Krüger 2002). Die weit reichendste Regelung existiert im Stadtstaat Bremen. Die in vielen Ländern alternativ diskutierte Frage, „staatliche Vorgaben versus Deregulierung“ (Seevers 2002, S. 22) hat Bremen im Sinn einer „mittleren Systematisierung“ beantwortet. Darunter wird verstanden, dass der Stadtstaat im Interesse der Teilnehmenden und der Angebotsqualität den Einrichtungen zwar verpflichtende Vorgaben macht, dass aber eine hohe Eigenverantwortlichkeit für die Form der Qualitätsentwicklung bei den Einrichtungen erhalten bleibt. Alle Einrichtungen werden neben einem Nachweis rechtlicher, wirtschaftlicher und sachlicher Voraussetzungen dazu verpflichtet, ein entwicklungsorientiertes Qualitätsmanagementsystem schriftlich zu dokumentieren und selbiges von externen Gutachtern beurteilen zu lassen. Mit Fragen der Qualitätsentwicklung haben sich die unterschiedlichen Weiterbildungsakteure schon immer beschäftigt. Das Neue an dem Qualitätsdiskurs in der Weiterbildung des letzten Jahrzehnts ist, dass die Qualitätsentwicklung im systemischen Zusammenhang mit deren organisationalen Bedingungen als Ganzes in den Mittelpunkt des Interesses rückt.

### **3. Qualitätsaktivitäten: Vielfalt oder Zerfaserung?**

Ein Anfang des neuen Jahrtausends von der Bundesregierung und den Ländern eingesetzter Arbeitsstab Forum Bildung (2001) empfiehlt die Entwicklung eines Qualitätstests, das für alle Weiterbildungseinrichtungen offen ist. Es soll eine Qualitätstransparenz für die Interessenten an Weiterbildung und für die Öffentlichkeit gewährleisten. Die Konzertierte Aktion Weiterbildung, ein Politikberatungsorgan des Bundesministeriums für Bildung und Forschung, in dem die wichtigsten überregionalen Träger der allgemeinen und beruflichen Weiterbildung vertreten sind, forderte im Anschluss an eine bundesweite Expertentagung die Entwicklung eines übergreifenden Referenzsystems, das der Unterschiedlichkeit der Einrichtungen gerecht werden soll (vgl. KAW 2002, S. 148). Bezogen auf die allgemeine Weiterbildung entwickelte die Kultusministerkonferenz der Länder Vorstellungen zu einem von Bund, Ländern, Wissenschaft und Praxis getragenen trägerübergreifenden Testierungssystem (vgl. Krug 2002, S. 41). Einige Akteure verweisen darauf, dass die Qualitätsförderung in der Weiterbildung einen erweiterten Blick braucht. Die ausschließliche Betrachtung der Träger- und Maßnahmenebene genüge nicht. Entsprechend reklamiert die KAW weitergehende Strukturelemente für eine systemorientierte Qualitätspolitik. Dazu zählen:

- „Entwicklung eines Referenzrahmens im Sinne eines Leitbildes unter Berücksichtigung der ... [relevanten] Kriterien für die Qualitätsentwicklung,
- Verknüpfung des Qualitätsmanagements mit Konzepten der Organisations- und Personalentwicklung,
- Stärkung der Position der Verbraucher, z.B. durch Bildungstests oder durch Integration vorhandener Checklisten,
- Erhöhung der Transparenz des Angebots durch Weiterentwicklung z.B. bereits bestehender regionaler und überregionaler Datenbanken im Internet, Bestandsaufnahmen, (...)
- Förderung der Professionalität des Personals in der Weiterbildung,
- Unterstützung von träger- und einrichtungsübergreifenden Initiativen wie beispielsweise Vergleichsringen und Entwicklungsgruppen,
- Ausbau internationaler Kontakte zur Sicherung des Anschlusses an Entwicklungen von Qualitätsmanagementsystemen,
- Förderung der Qualitätsforschung in der Weiterbildung, insbesondere der Wirkungsforschung“ (KAW 2002, S. 148).

Die weiterbildungspolitischen Empfehlungen wurden in den letzten Jahren von zahlreichen Initiativen, Projekten und Modellversuchen (vgl. Meisel 2001; Hartz/Meisel 2006) aufgegriffen.

Bildungsbereichsspezifische Initiativen: Das Bundesinstitut für Berufsbildung hat eine Reihe von Modellversuchen zur Qualitätsentwicklung der betrieblichen Weiterbildung initiiert und wissenschaftlich begleitet. Als Ergebnis wurden praxisorientierte Handreichungen entwickelt (vgl. Mohr u.a. 1998). Primär für die beruflichen Weiterbildungseinrichtungen haben die deutschen Wirtschaftsverbände die Firma CERTQUA gegründet. Sie bietet Beratung und Zertifizierung von Qualitätsmanagementprozessen nach DIN/EN/ISO 9000ff. an (vgl. Wuppertaler Kreis e.V./CERTQUA 2002). Für den Bereich der öffentlichen Weiterbildung hat eine Projektgruppe am Deutschen Institut für Erwachsenenbildung (DIE) eine auf die Weiterbildungsbranche abgestimmte Version des Modells der European Foundation for Quality Management (EFQM) erarbeitet (vgl. Heinold-Krug/Griep/Klenk 2001). Im Fremdsprachenbereich existieren auf europäischer Ebene unterschiedlich organisierte Qualitätssicherungssysteme (vgl. von der Handt 1999). Zudem können Initiativen in einzelnen Programmbereichen der allgemeinen Weiterbildung beobachtet werden (vgl. u.a. zur kulturellen Bildung Stang 1999, zur wertheorientierten Bildung Hohmann 1999, zur Gesundheitsbildung Venth 1999). Auch für die politische Weiterbildung wurden eigenständige konzeptionelle Überlegungen angestellt (vgl. Meisel 2005). Auf der Grundlage eines vom Bundesministerium für Jugend und Familie geförderten Modellversuchs hat die Universität Heidelberg Strukturierungshilfen für ein Qualitätsmanagement in der Familienbildung publiziert (vgl. Schiersmann/Thiel/Pfizenmaier 2001). Der Bundesverband der Träger beruflicher Bildung (BBB) hat als Konsequenz auf die Anforderungen der Bundesagentur für Arbeit ein „Bildungs-Qualitäts-Management“ (BQM) für die berufliche Weiterbildung entwickelt (vgl. Bundesverband der Träger beruflicher Bildung 2005).

Träger- und länderübergreifende Initiativen und Projekte: Das Deutsche Institut für Erwachsenenbildung (DIE) entwickelte in einem Bund-Länder-Projekt „Qualitätssicherung“ übergreifende Prinzipien eines Qualitätsmanagements für die allgemeine Erwachsenenbildung (vgl. Kuchler/Meisel 1999). Seitens des Bundesministeriums für Bildung und Forschung (BMBF) wurde ein ebenfalls vom DIE realisiertes Projekt zur Entwicklung und Erprobung eines „Fortbildungskonzepts zum/zur Qualitätsentwickler/in für Weiterbildungsorganisationen“ gefördert (vgl. Mathes 2002). Das BMBF förderte darüber hinaus ein Projekt „Qualitätsentwicklung in Weiterbildungsverbänden“. Angesprochen waren solche Träger, die in dezentralen und teilweise nicht-betriebsförmigen Organisationsstrukturen arbeiten, wie Teile der evangelischen Erwachsenenbildung oder Weiterbildungsinitiativen des Arbeitskreises Arbeit und Leben. In diesem Zusammenhang wurden Eckpunkte für ein Rahmenmodell zur Qualitätsentwicklung im Verbund von Bildungseinrichtungen (QVB) entwickelt (vgl. Liebold/Seiverth 2005; DEAE/AuL 2004). Ziel eines Bund-Länder-Verbundprojekts „Qualitätstestierung in der Weiterbildung“ (vgl. Zech 2005, Hartz/Herr/Veltjens 2006) war es, ein Modell Lernerorientierte Qualitätstestierung (LQW) bundesweit in Einrichtungen der allgemeinen und beruflichen Weiterbildung einzuführen. Als eine weitere übergreifend angesiedelte Initiative kann auch das vom Bundesministerium für Bildung und Forschung (BMBF) geförderte Projekt zu vergleichenden Bildungstests bei der Stiftung Warentest eingeordnet werden (vgl. Töpfer 2004), das zukünftig dauerhaft etabliert werden soll.

Träger- und landesbezogene Aktivitäten: Bei der Entwicklung eines „Fragenkatalogs zur Selbstevaluierung von Weiterbildungseinrichtungen“ des Landesverbandes der Volkshochschulen Niedersachsen (1996) handelt es sich ebenso um eine trägerbezogene Initiative wie bei der Implementation des EFQM-Konzepts im Zuständigkeitsbereich des bayerischen Volkshochschulverbandes (vgl. Lang 2002) oder bei der Einführung einer EFQM-Adaption in den Volkshochschulen in Baden-Württemberg. Zu erwähnen ist hier auch ein Projekt, das vom Land Rheinland-Pfalz gefördert wurde und die Qualitätsentwicklung in der Evangelischen Arbeitsgemeinschaft für Erwachsenenbildung im Bundesland unterstützte. Zu den in erster Linie landesbezogenen Aktivitäten zählen ebenfalls das Gütesiegel des Vereins Weiterbildung Hamburg e.V. (vgl. Krüger 2002) sowie das Qualitätsentwicklungssystem (QES) (vgl. Wiesner/Knoll u.a. 2004), das unter Mitwirkung der Universitäten Leipzig und Dresden für Weiterbildungseinrichtungen in Sachsen erarbeitet wurde. Bereits 1996 wurde in Sachsen ein Qualitätsprüfungssystem für solche Einrichtungen entwickelt, die Finanzmittel aus dem EU Sozialfonds Mittel erhalten wollten (vgl. Gnahn 1996).

Die hier nur exemplarisch genannten Aktivitäten belegen, dass in den letzten Jahren auf den unterschiedlichen Ebenen und in beinahe allen Bereichen des Weiterbildungssystems zahlreiche Akteure bei der Verwirklichung von Qualitätssicherungs- und Qualitätsentwicklungsmaßnahmen eingebunden waren und sind. Die Vielfalt der Aktivitäten kann als ein Indiz dafür gelten, dass sich an der Qualitätspolitik Weichenstellungen für die zukünftige Gestaltung des Weiterbildungssystems festmachen (vgl. Meisel 2003). Gezielt gefördert wird die Nachfrageseite (z.B. mit dem Projekt der Stiftung Warentest oder der Vergabe von Bildungsgutscheinen in der beruflichen Weiterbildung). Die poli-

tischen Entscheidungsträger fördern zudem solche Projekte, die Einrichtungen in die Lage versetzen, sich mit einer eigenverantworteten Qualitätsstrategie auf dem Weiterbildungsmarkt zu bewegen. Durchgesetzt hat sich der Trend zur externen Zertifizierung. Von staatlicher Seite aus werden für die unterschiedlichen Weiterbildungsbereiche unterschiedliche Verbindlichkeiten definiert: Für die bundesverantwortete berufliche Weiterbildung im Rahmen der Bundesagentur für Arbeit ist der Nachweis eines angemessenen Qualitätsmanagements zur Pflicht gemacht worden. In der von den Ländern verantworteten allgemeinen Weiterbildung werden wiederum unterschiedliche Strategien verfolgt.

Diese unterschiedlichen Strategien auf den verschiedenen ordnungspolitischen Ebenen führten zu durchaus beeindruckenden und vielfältigen Qualitätsanstrengungen im Weiterbildungssystem. Die Breite und die Vielfalt enthalten eine enorme und nicht zu unterschätzende Produktivität. Die Kehrseite der Vielfalt ist jedoch eine fehlende Transparenz – ein für individuelle und institutionelle Nachfragende sowie die öffentlichen Förderungsinstanzen gleichermaßen wichtiger Aspekt. Vor dem Hintergrund der ordnungspolitisch unterschiedlichen Zuständigkeiten erscheint es unrealistisch, dass sich in naher Zukunft alle Weiterbildungsakteure auf ein gemeinsames Qualitätsmanagementkonzept festlegen. „Um so wichtiger ist die Herausforderung, sich auf Dimensionen zu einigen, die eine Vergleichbarkeit der Konzepte ermöglichen und deren Angemessenheit für den Bereich der Weiterbildung überprüfbar machen“ (Schiersmann 2002, S. 25). Vorgeschlagen wurde deshalb ein Strukturmodell für die Analyse und Einordnung von Qualitätskonzepten (vgl. Meisel 2002). Das DIE hat sich dieser Aufgabe angenommen und entsprechende Übersichten erstellt (Veltjens 2006).

Aktuellerweise zeichnet sich ab, dass in den praktischen Entwicklungen mehr die Gemeinsamkeiten als die Differenzen zwischen den unterschiedlichen Ansätzen in den Vordergrund gestellt werden. In der Weiterbildungslandschaft haben sich, sieht man einmal von regionalen Besonderheiten ab, im Grunde genommen drei überregionale relevante Qualitätsmanagementmodelle etabliert.

Erstens das Konzept nach der DIN-EN-ISO 9000 ff, das im Weiterbildungsbereich in erster Linie von der Zertifizierungsfirma CERTQUA (vgl. Orru 2001) vermarktet wird. Das ausgeprägt prozessorientierte Verfahren lässt einrichtungsbezogene Ausdeutungen durchaus zu und orientiert sich primär an dem Ziel der externen Zertifizierung. Das Modell basiert auf einem Regelkreis von den Kundenerwartungen bis zur Herstellung von Kundenzufriedenheit, die als zu erreichendes Ziel in diesem Konzept eine zentrale Rolle spielt. Das Modell setzt sich aus fünf Qualitätsmanagement-Elementen zusammen: Verantwortung der Leitung, Management und Ressourcen, Produktrealisierung, Messen, Analysieren, Verbessern und kontinuierliche Verbesserung des Qualitätsmanagements. Schwächen des ISO-Ansatzes werden in der Fachwelt in erster Linie darin gesehen, dass eine Bezugnahme des Konzepts auf die Lehr-Lern-Interaktion äußerst problematisch ist. Das Qualitätsmanagement richtet sich eindeutig auf die Organisation aus. Dem relativ hohen Dokumentations- und Kostenaufwand steht zumindest in der beruflichen Weiterbildung ein positiver Erkennungseffekt seitens der Betriebe, die ebenfalls eine Zertifizierung nach der ISO-Norm, erworben haben, gegenüber (vgl. Hartz/Meisel 2006, S. 66ff.).

Zweitens ist das EFQM-Modell (European Foundation for Quality Management) in unterschiedlichen Varianten verbreitet, das von einer Selbstbewertung von Befähigerkriterien (Leitung, Mitarbeitende, Profil und Strategie, Ressourcen und Kooperationen, Prozesse) sowie Ergebniskriterien (mitarbeiterbezogene, kundenbezogene, gesellschaftsbezogene Ergebnisse und andere wichtige betriebliche Ergebnissen) ausgeht. Der Versuch, eine branchenspezifische Version (vgl. Heinold-Krug/Griep/Klenk 2001) breit im Feld zu implementieren, ist zwar gescheitert. Dennoch wird dieses Qualitätsmanagementmodell besonders in der allgemeinen Weiterbildung häufig angewendet. Das Modell setzt eine intensive Phase der Selbstevaluation und eine hohe betriebliche Kompetenz in Auswahl und Bearbeitung von Verbesserungsprojekten voraus. Zwischenzeitlich existiert auch die Möglichkeit einer externen Zertifizierung. Die Offenheit des Konzepts lässt die Berücksichtigung von einrichtungsbezogenen Besonderheiten zu, was aber die Vergleichbarkeit von Einrichtungen untereinander einschränkt. Das Modell ist stark beteiligungsorientiert und entspricht damit der in den Bildungseinrichtungen häufig vorfindbaren Organisationskultur. Es folgt im Grunde nach der Struktur und inneren Logik von Organisationen im Allgemeinen, aber eben nicht der pädagogischen Qualität im besonderen (Hartz/Meisel 2006, S. 78ff.).

Das dritte überregional relevante Modell ist das Konzept der „Lernerorientierten Qualitätstestierung“ (Ehse/Heinen-Tenrich/Zech 2001). Ausgegangen wird von der Anforderung an die Weiterbildungsorganisation für sich selbst „gelungenes Lernen“ zu definieren. In elf Qualitätsbereichen (Leitbild, Bedarfserschließung, Schlüsselprozesse, Lehr- und Lernprozesse, Evaluation von Bildungsprozessen, Infrastruktur, Führung, Personal, Controlling, Kundenkommunikation, und strategische Entwicklungsziele) werden jeweils Mindestanforderungen beschrieben. In einem Selbstreport muss die Einrichtung nachweisen, dass sie diese erfüllen. Nach einem Audit werden in einem abschließenden Workshop die weiterführenden Entwicklungsziele herausgearbeitet. Dieses Modell unterscheidet sich insofern von den anderen Ansätzen, als dass es die spezifischen Ziele und Bedingungen von Bildungseinrichtungen ausdrücklich berücksichtigt. Zwar ermöglicht es einrichtungsindividuelle Umgangsvarianten; da aber Mindestvoraussetzungen definiert sind, ist zumindest tendenziell eine Vergleichbarkeit der Einrichtungen gewährleistet. Das Modell kombiniert Selbst- und Fremdevaluation und entlässt die Einrichtungen nicht aus der Verantwortung, die Bezüge zwischen organisationaler und pädagogischer Qualität zu bedenken. Wissenschaftliche Begleituntersuchungen (vgl. Bosche/Veltjens 2005; Hartz/Schrader/Berzbach 2005; Hartz 2006) legen gleichwohl die Vermutung nahe, dass sich das Modell in der Anwendungspraxis ebenfalls eher auf die organisatorische Seite der Weiterbildung ausrichtet.

Bei allen Unterschieden im Detail lassen sich bei den unterschiedlichen Modellen auch Gemeinsamkeiten feststellen. Selbst- und Fremdevaluation werden in unterschiedlicher Reichweite kombiniert. Im Mittelpunkt der Qualitätsanstrengungen steht die Optimierung der organisationalen Bedingungen des Lehrens und Lernens. Die Umsetzung des Qualitätsmanagements wird prozessorientiert angegangen und konzentriert sich auf Schlüsselsituationen des Weiterbildungsmanagements (vgl. Hartz/Meisel 2006 S. 89ff.).

#### **4. Perspektiven: Professionalitätsentwicklung – die konsequente Fortsetzung der Qualitätsdebatte in der Weiterbildung**

Mit den hier unvollständig dargestellten, aber dennoch beeindruckenden Anstrengungen weist das Weiterbildungssystem, anders als bei früheren Modernisierungsschüben, die Fähigkeit auf, für eine strukturelle Verankerung von der Qualitätsentwicklung in der Profession zu sorgen. Die Breite und Vielfalt der Initiativen beinhaltet eine eminente Produktivität, wie aus zahlreichen Erfahrungsberichten erkennbar ist. Auf den unterschiedlichsten Ebenen sind die verschiedenen Akteure aktiv in die Umsetzungsarbeiten der Qualitätsentwicklung eingebunden, was deren Akzeptanz sicherlich erhöht. Angesichts der Heterogenität der Weiterbildungslandschaft, die sich in unterschiedlichen politischen Zuständigkeiten, finanziellen Bedingungen, institutionellen Verfasstheiten und inhaltlichen Ausrichtungen darstellt, wäre es unrealistisch, ein allgemeingültiges Qualitätsmanagementverständnis zu erwarten. Die Produktivität der Vielfalt hat jedoch auch Medaillenkrückseite: Die Zerfaserung der Qualitätsanstrengungen im Weiterbildungssystem widerspricht zumindest teilweise den in der Qualitätsdebatte stets vorgetragenen Zielen wie z.B. dem der Qualitätstransparenz für die „Kunden“. Aufgrund der unterschiedlichen Zuständigkeiten für die Weiterbildung im föderalen Bermudadreieck zwischen Bund und Ländern und den einzelnen Ressorts werden von den Einrichtungen z.T. Mehrfachzertifizierungen, die an Ressourcenverschwendung grenzen, verlangt.

Auch wenn einzelne Qualitätsmanagementmodelle mit ihrem spezifischen Zugang versuchen, die Organisationsqualität wie auch die pädagogische Qualität gleichermaßen zu fördern, ist bei einem Blick in die Weiterbildungspraxis nicht zu übersehen, dass sich die meisten Anstrengungen auf die organisationale Seite der Qualitätsentwicklung konzentrieren (vgl. Meisel/Dollhausen 2006, S. 60f.). Dies mag verständlich sein, weil die Organisationsqualität in der Weiterbildungsbranche in der Vergangenheit vernachlässigt wurde und ein Nachholbedarf festzustellen war. Unter den Bedingungen des Strukturwandels der Weiterbildung und deren Marktbedingungen ist es deshalb eher konsequent, ein systematisches organisationsbezogenes Qualitätsmanagement einzuführen und nachzuweisen. Organisationales Handeln und professionelles pädagogisches Handeln folgen aber einer anderen Handlungslogik (vgl. Hartz/Meisel 2006, S. 118ff.). Bei der Qualitätsentwicklung der Organisation geht es in der Regel um Standardisierung von Prozessen, Festlegung von verbindlichen Arbeitsschritten. Individuell ausgeprägte Handlungsformen sollen reduziert werden; die Organisation soll unabhängig von den Arbeitsstilen der beteiligten Personen „reibungslos“ funktionieren. Bei professionellem erwachsenenpädagogischem Handeln geht es demgegenüber eher um fallbezogenes, gerade nicht um standardisiertes Anwenden von theoretischem Wissen. Der Situationsbezug steht im Vordergrund. Professionelles pädagogisches Handeln durchbricht geradezu routiniertes Handeln. Es erfordert eine hohe Flexibilität und Kompetenzen, auch mit Unsicherheiten umzugehen. Hier ist authentisches Handeln der beteiligten Subjekte gefragt. In der Lehr-Lern-Interaktion greift die Logik des pädagogischen Handelns. Diese lässt sich gerade nicht in standardisiertem Verfahren pressen. Das Gegenteil ist der Fall. Professionelles Handeln in der Lehr-Lernsituation zeichnet sich durch „situative Kom-

petenz“ aus. Qualität bemisst sich über das diesbezügliche „Vermögen“ des Lehrenden. An dieser Herausforderung müssen weitergehende Qualitätsentwicklungsstrategien ansetzen. Diesen kann man sich über eine systematische Kompetenzentwicklung auf Seiten der Professionellen annähern. Damit verbunden wäre die Förderung der Professionalität der Weiterbildner. So gesehen, ist die Debatte um Qualitätsmanagement auch als Chance begreifbar, die in der Weiterbildung steckengebliebene Professionalisierung wieder aufzugreifen und diese weiter voranzutreiben (Meisel 2002; Kraft 2006). Die zunehmende Ausdifferenzierung unterschiedlicher Angebotsformen in der Weiterbildung, die Notwendigkeit zur professionellen Unterstützung von selbstgesteuerten Lernformen durch die Weiterbildungsorganisationen, die Anforderungen, die sich aus dem Nachweis von informell erworbenen Kompetenzen der Lernenden ergeben, die Integration medial unterstützter Lernorganisationsformen, der zunehmende Bedarf an Information und Beratung in der Weiterbildung – all diese Trends in der Weiterbildung verlangen nach gezielter Weiterentwicklung der professionellen Kompetenzen des erwachsenenpädagogischen Personals. Das pädagogische Personal stellt in erster Linie die Qualität des Lehr- und Lernprozesses mit den Lernenden her. Insofern besteht die Notwendigkeit die Seite des professionellen pädagogischen Handelns in der Qualitätsdebatte wieder stärker zu gewichten. Für organisierte Weiterbildung ist die Organisationsqualität eine wesentliche Voraussetzung für die pädagogische Qualität. Letztere entsteht aber nicht automatisch, sondern bedarf besonderer Anstrengungen.

Der konzentrierte Blick auf die Organisations- und pädagogische Qualität darf darüber hinaus nicht die Perspektive der Systemqualität verstellen. Nachdenklich muss stimmen, dass in den Jahren der angestregten Bemühungen um Qualitätsentwicklung die Weiterbildungsbeteiligung im internationalen Vergleich gesunken ist. Während die Akteure in der Weiterbildungspraxis sich den schwierigen Implementierungsprozessen des Qualitätsmanagements offensiv ausgesetzt haben, müssen die weiterbildungspolitischen Bemühungen um die Systemqualität im Kontext der europäischen Strategien zum Lebenslangen Lernen noch dringend gestärkt werden. Unterstützend könnte wirken, wenn in Sachen Qualitätsmanagement und Qualitätsentwicklung ein inhaltlicher Austausch zwischen den Bildungssektoren gestärkt wird. Auf lange Sicht könnte eine neue, übergreifende Expertenkultur entstehen, die der Qualität des ganzen Bildungssystems zugute käme.

## **Literaturverzeichnis**

- Arnold, R./Schüßler, I. (Hrsg.) (2003): Ermöglichungsdidaktik in der Erwachsenenbildung. Baltmannsweiler: Schneider-Verlag Hohengehren. S. 142–155.
- Balli, C./Krekel, E. M./Sauter, E. (2004): Qualitätsentwicklung in der Weiterbildung – zum Stand der Anwendung von Qualitätssicherungs- und Qualitätsmanagementsystemen bei Weiterbildungsanbietern. Bonn: BiBB.
- Bundesregierung (2004): Anerkennungs- und Zulassungsverordnung – Weiterbildung – AZWV: Verordnung über das Verfahren zur Anerkennung von fachkundigen Stellen sowie zur Zulassung von Trägern und Maßnahmen der beruflichen Weiterbildung nach dem Dritten Buch Sozialgesetz-

- buch, Teil I Nr. 28. Bonn. S. 1100–1104. URL: [www.bgblportal.de/BGBl/bgbl1f/bgbl104s1100.pdf](http://www.bgblportal.de/BGBl/bgbl1f/bgbl104s1100.pdf). (23.2.06).
- Bosche, B./Veltjens, B. (2005): LQW: Ein Qualitätsmanagementsystem in der Praxis. In: Die Österreichische Volkshochschule, Expressum-Verlag, H. 215, S. 15–21.
- Bülow-Schramm, M. (2006): Qualitätsmanagement in Bildungseinrichtungen. Münster: Waxmann-Verlag.
- Bundesanstalt für Arbeit (1999): Anforderungskatalog an Bildungsträger und Maßnahmen der beruflichen Weiterbildung, BA-I-FW-11/99. Nürnberg: Bundesanstalt für Arbeit.
- Bundesverband der Träger beruflicher Bildung e.V. (BBB) (Hrsg.) (2005): Bildungs-Qualitäts-Managementssystem. Kompendium. BQM – Der Qualitätsstandard für Bildungsträger. 2. Aufl., Hamburg: Bundesverband der Träger beruflicher Bildung e.V.
- Deutsche Evangelische Arbeitsgemeinschaft für Erwachsenenbildung (DEAE)/Bundesarbeitskreis Arbeit und Leben (AuL) (2004): Qualitätsentwicklung im Verbund von Bildungseinrichtungen – QVB. Das Rahmenmodell. Frankfurt a. M., Düsseldorf: DEAE/AUL.
- Ehres, C./Heinen-Tenrich, J./Zech, R. (2001): Das lernerorientierte Qualitätsmodell für Weiterbildungsorganisationen. Hannover: Expressum-Verlag.
- Faulstich, P. (1991): Qualitätskriterien für die Erwachsenenbildung als ein Fokus der Berufsbildungsforschung. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, H. 7, Steiner-Verlag, S. 572–581.
- Forum Bildung (2001): Qualitätsentwicklung und Qualitätssicherung im internationalen Wettbewerb. Vorläufige Empfehlungen und Expertenbericht. Bonn: Forum Bildung.
- Friebel, H. (1993): Der gesplittete Weiterbildungsmarkt und der Lebenszusammenhang der Teilnehmer/innen. In: Friebel, H. u.a. (Hrsg.): Weiterbildungsmarkt und Lebenszusammenhang. Bad Heilbrunn: Klinkhardt.
- Gerl, H./Pehl, K. (1983): Evaluation in der Erwachsenenbildung. Bad Heilbrunn: Klinkhardt.
- Gieseke, W. (1997): Die Qualitätsdiskussion aus erwachsenenpädagogischer Sicht. Was bedeutet Qualität in der Erwachsenenpädagogik? In: Arnold, R. (Hrsg.): Qualitätssicherung in der Erwachsenenbildung. Leske und Budrich: Opladen, S. 29–47.
- Gnahn, D. (1996): Handbuch zur Qualität in der Weiterbildung – Stand, Perspektiven, Praxis. Frankfurt a.M.: Max-Träger-Stiftung/GEW.
- Gnahn, D. (1999): Rahmenbedingungen der Qualitätsentwicklung: Schnittmengen und Unterschiede der gesetzlichen Regelungen in den einzelnen Bundesländern. In: Küchler, F. v./Meisel, K. (Hrsg.): Qualitätssicherung in der Weiterbildung – Auf dem Weg zu Qualitätsmaßstäben. Frankfurt a. M.: Bertelsmann-Verlag, S. 31–45.
- Hartz, S./Schrader, J./Berzbach, F. (2005): Systemevaluation im Rahmen des Projektes „Qualitätstestierung in der Weiterbildung“. Ergebnisbroschüre. Tübingen: Wissenschaftsgemeinschaft Gottfried-Wilhelm-Leibniz.
- Hartz, S. (2006): Lernerorientierte Qualitätsentwicklung – Ergebnisse der Systemevaluation. In: Hartz, S./Herr, M./Veltjens, B.: Qualitätstestierung in der Weiterbildung. Bielefeld: Bertelsmann-Verlag.
- Hartz, S./Herr, M./Veltjens, B. (2006): Qualitätstestierung in der Weiterbildung. Bonn: DIE.
- Hartz, S./Meisel, K. (2006): Qualitätsmanagement. Studentext Erwachsenenbildung. Bielefeld: Bertelsmann-Verlag.
- Heinold-Krug, E./Meisel, K. (Hrsg.) (2002): Qualität entwickeln – Weiterbildung gestalten. Bielefeld: Bertelsmann-Verlag, S. 23–37.
- Heinold-Krug, E./Griep, M./Klenk, W. (2001): EFQM. Version Erwachsenenbildung/Weiterbildung. Frankfurt a. M.: DIE.
- Hohmann, R. (1999): Qualitätssicherung in der wert- und normbezogenen Bildung. In: von Küchler, F./Meisel, K. (Hrsg.): Qualitätssicherung in der Weiterbildung – Auf dem Weg zu Qualitätsmaßstäben. Frankfurt a. M.: Deutsches Institut für Erwachsenenbildung, S. 92–101.
- KAW – Konzertierte Aktion Weiterbildung (Hrsg.) (2002): Stellungnahme des ständigen Ausschusses der Konzertierte Aktion Weiterbildung zur Qualitätsentwicklung in der Weiterbildung. Bonn: KAW.

- Kraft, S. (2006): Umbrüche in der Weiterbildung – dramatische Konsequenzen für das Weiterbildungspersonal. In: DIE-Fakten. URL: [www.die-bonn.de/esprid/dokumente/doc-2006/Kraft06\\_01.pdf](http://www.die-bonn.de/esprid/dokumente/doc-2006/Kraft06_01.pdf) (6.5.06):
- Krug, P. (2002): Ansätze und Perspektiven der Länder zur Weiterbildungsqualität. In: Heinold-Krug, E./Meisel, K. (Hrsg.): *Qualität entwickeln – Weiterbildung gestalten*. Bielefeld: Bertelsmann-Verlag, S. 38–49.
- Krüger, T. (2002): Es geht auch ohne Weiterbildungsgesetz – das Modell Weiterbildung Hamburg e.V. In: Melms, B. (Hrsg.): *Relevanz rechtlicher Regelungen für die Qualitätssicherung der Weiterbildung auf Ebene der Länder in der Bundesrepublik Deutschland*. Bielefeld: Bertelsmann-Verlag, S. 390–396.
- Küchler, F. v./Meisel, K. (Hrsg.) (1999): *Qualitätssicherung in der Weiterbildung – Auf dem Weg zu Qualitätsmaßstäben*. Frankfurt a.M., Bielefeld: Bertelsmann-Verlag.
- Kultusministerkonferenz (1996): Beschluss der 146. Amtschefkonferenz vom 25. und 26. April 1996. Bonn.
- Kuhlenkamp, D. (2003): Weiterbildung als Teil öffentlicher Wohlfahrt? In: Wittpoth, J. (Hrsg.): *Erwachsenenbildung und Demokratie. Dokumentation der Jahrestagung der Deutschen Gesellschaft für Erziehungswissenschaft*. Beiheft REPORT, H. 1, Bielefeld: Bertelsmann-Verlag, S. 25–44.
- Kuper, H. (2002): Stichwort: Qualität im Bildungssystem. In: *Zeitschrift für Erziehungswissenschaft*, H. 2, Opladen: Barbara Budrich-Verlag, S. 533–551.
- Landesverband der Volkshochschulen Niedersachsens (Hrsg.) (1996): *Qualitätssicherung in der Volkshochschule – Fragenkatalog zur Selbstevaluation*. Hannover: Niedersächsischer Volkshochschulverband.
- Lang, W.F. (2002): Qualitätsmanagement als Verbandsstrategie. In: Heinold-Krug, E./Meisel, K. (Hrsg.): *Qualität entwickeln – Weiterbildung gestalten*. Bielefeld: Bertelsmann-Verlag, S. 61–68.
- Liebold, C./Seiverth, A. (2005): Qualitätsentwicklung und Zertifizierung – Aufgaben für die Evangelische Erwachsenenbildung. In: *Forum Erwachsenenbildung*, H. 1, S. 28–32.
- Luhmann, N. (2002): *Das Erziehungssystem der Gesellschaft*. Frankfurt a. M.: Suhrkamp.
- Mathes, E. (2002): Qualität in den eigenen Reihen entwickeln. In: *DIE Zeitschrift für Erwachsenenbildung*, H. 3, Frankfurt a.M.: DIE, S. 28–29.
- Meisel, K. (2001): Managementprobleme in öffentlichen Erwachsenenbildungseinrichtungen – Anforderungen an intermediäre Forschungs- und Entwicklungsarbeiten für ausgewählte Managementaufgaben. Baltmannsweiler: Schneider-Verlag Hohengehren.
- Meisel, K. (2002): Qualitätsentwicklung im Aufbruch. In: Heinold-Krug, E./Meisel, K. (Hrsg.): *Qualität entwickeln – Weiterbildung gestalten. Handlungsfelder der Qualitätsentwicklung*. Bielefeld: Bertelsmann-Verlag, S. 9–19.
- Meisel, K. (2003): Zur politischen Qualität der Qualitätspolitik. In: *REPORT – Literatur- und Forschungsreport Weiterbildung*, Bielefeld: Bertelsmann-Verlag, H. 1, S. 61–67.
- Meisel, K. (2005): Qualitätsentwicklung in der Politischen Bildung. In: *Redaktionen Politische Bildung & kursiv – Journal für politische Bildung, Bildungsstandards – Evaluation in der politischen Bildung*. Schwalbach/Ts.: Bildung & Kursiv-Verlag 2005, S. 121–133.
- Meisel, K./Dollhausen, K. (2006): Erwachsenenpädagogische Qualität als Ziel des Qualitätsmanagements. In: Markert, W. (Hrsg.): *Qualität des beruflichen Lernens in der Weiterbildung*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Meisel, K. (2007): Öffentliche Weiterbildungsorganisationen im Strukturwandel. In: *Hessische Blätter für Volksbildung*, Bielefeld: Bertelsmann-Verlag, H. 2, S. 106–116.
- Melms, B. (2002): *Relevanz rechtlicher Regelungen für die Qualitätssicherung der Weiterbildung auf Ebene der Länder in der Bundesrepublik Deutschland*. Bielefeld: Bertelsmann-Verlag.
- Mohr, B. (Mitarb.) (1998): *Qualitätssicherung in der Weiterbildung*. In: *Leitfaden für Bildungsträger im Marktsegment „öffentlich geförderte Weiterbildung“*. Bielefeld: Bertelsmann-Verlag.

- Orru, A. (2001): Qualitätsmanagement und Zertifizierung nach ISO 9000 in der beruflichen Bildung. In: Zeitschrift für Bildungsverwaltung, H. 1–2, Deutsche Gesellschaft für Bildungsverwaltung e.V., S. 65–68.
- Sauter, E. (1992): Ansätze zur Qualitätssicherung in der beruflichen Weiterbildung. Maßstäbe und Maßnahmen. In: WSI-Mitteilungen, H. 6, S. 380–387.
- Schiersmann, C./Thiel, H.-U./Pfitzmaier, E. (2001): Organisationsbezogenes Qualitätsmanagement. EFQM-orientierte Analyse und Qualitätsentwicklungs-Projekte am Beispiel der Familienbildung. Opladen: Leske und Budrich.
- Schiersmann, C. (2002): Zweierlei Herausforderungen. In: DIE Zeitschrift für Erwachsenenbildung, H. 3, Frankfurt a.M.: DIE, S. 25–26.
- Schlutz, E. (1996): Zur Qualitätssicherung als Professionsaufgabe. In: Meisel, K. (Hrsg.): Qualität in der Weiterbildung. Frankfurt a.M.: DIE Materialien für Erwachsenenbildung, Bd. 3, S. 27–33.
- Schöll, I. (1996): Qualitätssicherung – Erfahrungen aus der Praxis. In: Meisel, K. (Hrsg.): Qualität in der Weiterbildung. Frankfurt a.M.: DIE Materialien für Erwachsenenbildung, Bd. 3, S. 93–96
- SeEVERS, M. (2002): Staatliche Vorgaben vs. Deregulierung. In: Melms, B. (Hrsg.): Relevanz rechtlicher Regelungen für die Qualitätssicherung der Weiterbildung auf Ebene der Länder in der Bundesrepublik Deutschland. Bielefeld: Bertelsmann-Verlag, S. 365–370.
- Stang, R. (1999): Qualität Kulturreller Bildung. In: KÜCHLER, F. v./Meisel, K. (Hrsg.): Qualitätssicherung in der Weiterbildung – Auf dem Weg zu Qualitätsmaßstäben. Frankfurt am Main: DIE, S. 46–55.
- Stock, J./ Wolff, H./Mohr, H. (1998): Delphi-Befragung 1996/1998: Potentiale und Dimensionen der Wissensgesellschaft – Auswirkungen auf Bildungsprozesse und Bildungsstrukturen. Basel: Prognos AG München.
- Terhart, E. (2000): Qualität und Qualitätssicherung im Schulsystem? Hintergründe – Konzepte – Probleme. In: Zeitschrift für Pädagogik, H. 6, Weinheim: Beltz Verlag, S. 809–829.
- Töpfer, A. (2004): Aufgabe und Struktur von Bildungstests. In: Balli, C./Krekel, E./Sauter, E. (Hrsg.): Qualitätsentwicklung in der Weiterbildung – Wo steht die Praxis? Bielefeld: Bertelsmann, S. 89–94.
- Veltjens, B. (2006): Qualitätsmodelle im Überblick. : URL: [http://www.die-bonn.de/esprid/dokumentation/doc-2006/veltjens06\\_01.pdf](http://www.die-bonn.de/esprid/dokumentation/doc-2006/veltjens06_01.pdf) (1.8.2006).
- Venth, A. (1999): Gesundheitsbildung: Qualität versteht sich nicht von selbst. In: KÜCHLER, F. v./Meisel, K. (Hrsg.): Qualitätssicherung in der Weiterbildung. Frankfurt am Main: DIE, S. 69–76.
- von der Handt, G. (1999): Qualitätssicherung und Qualitätsentwicklung im Sprachunterricht. In: KÜCHLER, F. v./Meisel, K. (Hrsg.): Qualitätssicherung in der Weiterbildung – Auf dem Weg zu Qualitätsmaßstäben. Frankfurt am Main, Bielefeld: DIE, S. 56–68.
- Wiesner, G./Knoll, J. u.a. (2004): Das Qualitätsentwicklungssystem QES plus. Modellbeschreibung mit Hinweisen zur Implementierung und Bewertung. Leipzig: Sächsischer Volkshochschul Verband.
- Wuppertaler Kreis e.V./CERTQUA (2002): Qualitätsmanagement und Zertifizierung in der Weiterbildung nach dem internationalen Standard ISO 9000:2000. Neuwied: Ziel-Verlag.
- Zech, R. (2005): Lernerorientierte Qualitätstestierung in der Weiterbildung. LQW II. 3. Hannover: Expressum-Verlag.

*Anschrift des Autors:*

Prof. Dr. Klaus Meisel (Managementdirektor), Münchner Volkshochschule GmbH, Kellerstraße 6, 81667 München

*Lutz von Rosenstiel*

## **Qualitätssicherung in der betrieblichen Weiterbildung**

Qualität ist in der aristotelischen Philosophie eine spezifische Kategorie, die der Quantität gegenübergestellt wird und die die Besonderheit und Beschaffenheit eines Dings kennzeichnet. Der Qualitätsbegriff hat nun in jüngerer Zeit hohe Beachtung gefunden. Möglicherweise liegt dies daran, dass auf dem Markt der Güter und Dienstleistungen aufgrund einer beständig wachsenden internationalen Konkurrenz, aufgrund staatlicher Forderungen und Kontrollen, aufgrund eines Wandels von Wertorientierungen etc. die Qualität zu einem kauf- oder nutzungsentscheidenden Merkmal wird. Zunehmend hat sich dann der Qualitätsbegriff auf verschiedenartige Prozesse ausgeweitet, um deren qualitative Optimierung man sich bemüht. So verschreibt man sich in vielen Unternehmen einem so genannten „Total quality management (TQM)“ (Reiß/v. Rosenstiel/Lanz 1997), verpflichtet sich einer „Qualitätsoffensive“ oder fasst Mitarbeiter und Mitarbeiterinnen in „Qualitätszirkeln“ (Bungard/Wiendieck 1986) zusammen, um mit Hilfe derartiger Maßnahmen Prozesse, Produkte oder Dienstleistungen zu verbessern. Dabei sind freilich jeweils (unternehmens-) politisch die Kriterien zu benennen und zu gewichten, an denen die Qualität konkret zu bestimmen ist.

In eher pragmatischer Weise lässt sich die Qualität als Übereinstimmung der Leistungen mit den Ansprüchen von Bezugsgruppen bestimmen. Bei der Weiterbildung wird man hier in erster Linie an die Sicht der Teilnehmer einer Weiterbildungsmaßnahme und an jene der Auftraggeber denken, je nach Ziel und Perspektive aber evtl. auch an die Ansprüche von Kunden, Vorgesetzten, Kollegen oder Mitarbeitern derer, die geschult wurden. Dabei kann gezeigt werden, dass Qualität kein eindimensionales Konstrukt ist, sondern sich in unterschiedliche Teilqualitäten ausdifferenzieren lässt. Bei der Weiterbildung könnten sich derartige Teilqualitäten z.B. auf den vermittelten Stoff, die Didaktik, die bereitgestellten Unterlagen und Lernhilfen, den Transfer des Gelernten in die Praxis, die Anwendungsmöglichkeiten oder die Kosten-Nutzen-Relation beziehen. All dies könnte die Grundlage von konkreter Informationsgewinnung sein, mit deren Hilfe die Qualitätssicherung, die Qualitätskontrolle und die Qualitätsmehrerung der Weiterbildung erfolgen können. All dies ließe sich schließlich zu einem integrierten Qualitätsmanagement der Weiterbildung zusammenführen (Reiß 2000).

Auch für Bildungseinrichtungen gibt es ähnlich wie für andere Institutionen spezifische Qualitätsstandards. Beispiele sind ISO 2000–2004, EQA (European Award oder EFQM (European Foundation of Quality Management) (Büssing 2007) nach denen sich Unternehmen – auch der Fort- und Weiterbildung – zertifizieren lassen können. Auch Bildungsgänge bedürfen vermehrt der Akkreditierung – z.B. MBA-Studiengänge für Führungskräfte, nach FIBAA – um damit ihre Qualität im Sinne der zuvor festgeleg-

ten Standards nachzuweisen. Einzelne Kursangebote werden vielfach durch darauf spezialisierte Institutionen, z.B. solchen für Warentests bis hin zu Management-Zeitschriften oder Wirtschaftsteilen großer Tageszeitungen auf ihre Qualität untersucht. Darauf soll hier nicht eingegangen werden, sondern aus organisationspsychologischer Sicht gefragt werden, was man unter der Qualität betrieblicher Weiterbildung verstehen könnte und wie sie zu sichern ist.

## **1. Betriebliche Weiterbildung als Bestandteil der Personalentwicklung**

Die betriebliche Personalentwicklung (Conradi 1983; Neuberger 1994; Sonntag 2006) ist weit zu verstehen und wird in höchst unterschiedlicher Weise definiert. Als eine relativ typische Begriffsbestimmung sei die von Rüter (1988, S. 3) zitiert, der die Personalentwicklung als die „systematisch vorbereitete, durchgeführte und kontrollierte Förderung der Anlagen und Fähigkeiten des Mitarbeiters in Abstimmung mit seinen Erwartungen und mit besonderer Berücksichtigung der Veränderungen der Arbeitsplätze und Tätigkeiten ...“ versteht. Die betriebliche Weiterbildung, die breit verstanden werden soll, ist dann jeweils nur ein Teil der Personalentwicklung. Sie baut auf einem Basisberuf auf und führt zu einer Spezialisierung im Basisberuf oder soll Kenntnisse im Basisberuf aktualisieren und auf den neuesten Stand bringen – hier spricht man häufig auch von Fortbildung – oder das Wissen in einer durch Weiterbildung erreichten Spezialisierung vertiefen (Bieling 1980).

Sieht man also ein zentrales Ziel der Personalentwicklung darin, dass Mitarbeiter und Mitarbeiterinnen den aktuellen oder gar künftigen beruflichen Anforderungen gewachsen sein sollen, so geschieht das vielfach nicht allein durch Weiterbildung (Staudt/Kriegesmann 1999). Sicherlich sollte man die institutionalisiert und zielorientiert ablaufende betriebliche Weiterbildung, die ein explizites Lernen fordert, nicht unterschätzen, doch spielten vermutlich das „Lernen im Prozess der Arbeit“ sowie das „Lernen im sozialen Umfeld“ (Erpenbeck/Sauer 2000) eine noch größere Rolle. Menschen lernen durch herausfordernde, sie vor neue Anforderungen stellende Arbeit im Zuge des Job Enlargement, des Job Enrichment oder des Job Rotation wie z.B. durch zeitweise Entsendung ins Ausland, sie lernen durch Team- und Projektarbeit von Anderen, aber auch durch informelle Kommunikation innerhalb des Betriebes oder über die Betriebsgrenzen hinweg sowie durch ehrenamtliche Tätigkeiten in Gemeinden oder Bürgerinitiativen, aber auch in der Familie. Der Qualifizierung kann aber auch ein selbstverantwortliches und selbstorganisiertes Lernen (Greif/Kurtz 1996), durch „Action Learning“ (Hauser 2008) oder die Nutzung des Intra- oder Internets oder anderer Medien dienen. Bei diesen letztgenannten Aktivitäten erfolgt das Lernen vielfach implizit und begünstigt die so genannte Kompetenzentwicklung (Heyse/Erpenbeck 2004), wobei die Kompetenzen hier als Dispositionen zum selbstorganisierten Handeln verstanden werden sollen (Erpenbeck/v. Rosenstiel 2003).

## **2. Betriebliche Weiterbildung im Netz der Mikropolitik**

Betriebe werden häufig als zweckorientierte Organisationen gesehen, innerhalb derer auf rationale Weise die Mittel, zu denen auch die Mitarbeiter zählen, eingesetzt werden, um die Zwecke zu realisieren, wobei alle Personen, alle Ressourcen, alle Aktivitäten dem eigentlichen Ziel des Unternehmens, der Gewinnmaximierung, untergeordnet sind. Ein derartiges Bild ist eine erschreckende Utopie. Die Realität sieht anders aus. Viele Organisationsmitglieder verfolgen ihre eigenen Ziele; sie versuchen dabei diese zum Teil mit nicht rationalen Mitteln, durch die Instrumentierung Anderer zu erreichen, sie schmeicheln, drohen, lassen sich auf Tauschgeschäfte ein und betreiben in diesem Sinne „Mikropolitik“ (Neuberger 2006b). Das hat zur Folge, dass auch das empirisch registrierbare Führungsverhalten der Logik rationaler Konzepte, den sog. Managementmodellen vielfach widerspricht (Mintzberg 1973; Neuberger 2006a).

Auch die betriebliche Weiterbildung ist in ein derartiges Kräftefeld verwoben. Dabei erscheint es ratsam zwischen manifesten und latenten Zielen zu differenzieren. Manifest wird wohl fast stets darauf verwiesen, dass es darum geht, die Teilnehmer der Maßnahme für die heutigen oder gar die kommenden Aufgaben zu qualifizieren. Latent aber können ganz andere Zielsetzungen ausschlaggebend sein, so etwa die demonstrative Belohnung oder Bestrafung der von der Bildungsmaßnahme Betroffenen, der Versuch des Leiters der Personalentwicklung sein eigenes Ansehen durch ein von ihm konzipiertes Seminar, das „gut ankommt“, zu verbessern, die Nutzung der Maßnahmen als „Parkplatz“ für Führungskräfte, die nach einem mehrjährigen Auslandseinsatz zurückkehren und für die man zunächst keine adäquaten Positionen findet, die Wahl des Plenums zur Präsentation bestimmter Projekte, Ideen oder Personen, die Verformung der Weiterbildung zu einem heiteren Event mit Unterhaltungs- und Erholungscharakter, etc. Es gibt eine Vielzahl derartiger latenter Zielsetzungen. Man wird aber kaum die Qualität der Maßnahme danach bewerten wollen, ob sie den angedeuteten latenten Zielen gerecht wird, so dass hier eine Beschränkung auf jene Ziele erfolgen soll, die manifest angestrebt werden: Die Qualifizierung der Teilnehmer der Maßnahme im Sinne der definierten Lernziele.

## **3. Spannungen zwischen wissenschaftlichen Anforderungen und praktischen Notwendigkeiten**

Nur wenige Felder der Personal- und Organisationspsychologie übertragen so unmittelbar grundlagenorientierte wissenschaftliche Erkenntnis in die Praxis, wie die Personalentwicklung (Holling/Liepmann 2007; Sonntag 2006). So lassen sich vielfältige Befunde der empirischen Pädagogik, der Lern- oder der Sozialpsychologie in die Gestaltung einer Bildungsmaßnahme übersetzen; Methoden qualitativer oder quantitativer Art der empirischen Sozialforschung erscheinen geeignet, etwa bei der Anlage eines Designs, die Wirkung der Maßnahme zu evaluieren. Es ließen sich in diesem Sinne aus der Wissenschaft vielfältige Kriterien ableiten, an denen Qualitätsindikatoren der

Weiterbildung im Sinne ihrer Wissenschaftlichkeit abgeleitet werden könnten. Dem stehen allerdings betriebliche Realitäten gegenüber; vielfach ist es das Kostenargument, das pädagogisch geforderte Maßnahmen wie z.B. die Bearbeitung eines individuellen Problems mit anschließendem persönlichem Coaching, verbietet; bestimmte Formen der an sich anzustrebenden Evaluation der Maßnahme, die beim einzelnen Teilnehmer erhebt, welche Inhalte er warum besonders gut und welche er nicht umsetzte, werden am Widerstand des Betriebs- oder Personalrats oder am betrieblichen Datenschutz scheitern; der Einsatz bestimmter, für die Thematik besonders gut erscheinender Trainer wird daran scheitern, dass der bisherige einen besonders „guten Draht“ zu einem Vorstandsmitglied hat etc. Die Planung, Durchführung und Evaluierung der betrieblichen Weiterbildung wird also immer ein Kompromiss zwischen dem wissenschaftlich Wünschbaren und dem praktisch Möglichen sein. Qualitätskriterien müssen das berücksichtigen.

#### **4. Grundsatzforderungen**

Relativ häufig trifft man in der Praxis auf eine Bestimmung der Qualität von Produkten oder Dienstleistungen über Zufriedenheitsindikatoren der Kunden oder Nutzer, wobei diese Zufriedenheitswerte wieder nach verschiedenen inhaltlichen Bereichen differenziert werden (z.B. beim Auto in die Zufriedenheit mit der technischen Zuverlässigkeit, dem Service, der Qualitätsanmutung, etc.). Selbstverständlich können derartige Maße Hinweise auf eine spezifische Facette der Qualität geben, doch sind sie allein kaum ausreichend. Ähnlich wie die Zufriedenheit eines Patienten mit dem Arzt oder die Zufriedenheit des Klienten mit einem Berater die durch Expertenurteil gemessene Leistung des entsprechenden Arztes oder Beraters nur unvollkommen widerspiegelt, ähnlich wie die vielfach erhobene Arbeitszufriedenheit die Qualität der Arbeit – z.B. gemessen an Humanisierungskriterien (Ulich 2007) – kaum angemessen abbildet, z.B. deshalb, weil hohe Zufriedenheitswerte gelegentlich die Konsequenz eines geringen Anspruchsniveaus oder der Mangels an Kenntnis von Alternativen sind, so kann auch die Zufriedenheit der Teilnehmer an einer Weiterbildungsmaßnahme den Blick auf Schwächen dieser Maßnahme geradezu verstellen. So kann z.B. ein rhetorisch brillanter Trainer, der über eine Vielzahl amüsanter Anekdoten und emotional bewegender Erfahrungen verfügt, die Teilnehmer geradezu begeistern, obwohl der Beitrag zum zuvor definierten Lernziel überaus bescheiden ist, oder es wird ein Outdoortraining, das durch intensive Teamerlebnisse und überraschende Übungen den Teilnehmern lebenslang in Erinnerung bleibt, möglicherweise keinerlei Transfer in die konkrete Arbeitssituation mit sich bringen, wenn darüber nicht explizit gesprochen und Transfermöglichkeiten aufgezeigt wurden. Gelegentlich werden sogar für die Personalentwicklung verantwortliche Seminarmanager und Trainer ihres Ansehens oder der Folgeaufträge wegen ausschließlich die Zufriedenheit der Kunden oder Teilnehmer zu optimieren suchen, um den häufig steinigen und mit geringer subjektiver Zufriedenheit verbundenen Weg der Übertragung in die nüchterne Alltagspraxis zu umgehen. Kurz: Die Zufrieden-

heit des Auftraggebers oder der Teilnehmer greift, wenn nur sie erfasst wird, als Qualitätsindikator zu kurz. Ergänzende Indikatoren sollten bedacht und höher gewichtet werden.

#### *4.1 Relevante und realistische Zielbestimmungen*

Ziele betrieblicher Weiterbildungsmaßnahmen werden gelegentlich in einer höchst oberflächlichen Weise bestimmt. So schreibt etwa die Personalabteilung potentielle Teilnehmer an und bittet sie, aus einem umfangreichen Katalog von Themen das anzukreuzen, was sie besonders interessiert. Ob die gewählten Inhalte dafür geeignet sind bestehende Schwächen der Antwortenden zu kompensieren oder ob sie relevant für die Anforderungen sind, wird nicht weiter geprüft. Noch problematischer wird es, wenn modische Themen oder „populäre“ Referenten oder Trainer ausgewählt werden, weil man weiß, dass dies in anderen Unternehmen „gut ankam“.

Die Lernziele sollten sich aus einer fundierten Bildungsbedarfsanalyse ergeben (Landy 1985; v. Rosenstiel/Molt/Rüttinger 2005). Dabei ist zum einen an die Arbeits- oder Aufgabenanalyse (Frieling/Sonntag 1999; Ulich 2006, 2007; Hacker 2005) zu denken, aber auch an den Versuch einer Bestimmung, wie sich die Aufgaben und damit auch die Anforderungen in Zukunft weiterentwickeln werden. Man erkennt dabei z.B. schwer auszuführende Verhaltenselemente und kann entsprechend die Lernenden zur erfolgreichen Bewältigung dieser Schwierigkeiten qualifizieren. Weiterhin gilt es, eine Analyse des sozialen Potenzials, also eine Ermittlung erfolgsrelevanter Merkmale der Mitarbeiter und Mitarbeiterinnen vorzunehmen, ihre Stärken und Schwächen in Bezug auf ihre Anforderungen und dabei insbesondere das Potenzial zu analysieren, also dabei die feststellbaren Lernmöglichkeiten und –bereitschaften (Schuler/Görlich 2006) zu erfassen. Schließlich gilt es eine Organisationsanalyse (Bissing 2007) vorzunehmen, weil sich ein Lernbedarf ja auch aus den Strukturen und Prozessen der Organisation, insbesondere aber auch aus deren Strategie ergibt (Sonntag/Stegmaier/Schaper 2006). Der Erfolg einer Unternehmensstrategie setzt eine damit abgestimmte strategische Kompetenzentwicklung voraus (v. Rosenstiel/Pieler/Glas 2004).

Aus der Aufgaben-, Personen- und Organisationsanalyse ergibt sich, welche Lernziele notwendig oder wünschenswert sind, ob es – im Sinne des Konzepts „Nutzen x Wahrscheinlichkeit“ – auch realistisch erscheint, diese Ziele in der konkreten Situation anzustreben. Erlauben es die Kultur des Unternehmens, die Belastungen der für die Bildungsmaßnahme vorgesehenen Personen, die Einstellungen ihrer jeweiligen Vorgesetzten oder die bereitstellbaren materiellen oder immateriellen Ressourcen überhaupt, die an sich wünschenswert erscheinenden Ziele anzustreben, oder würde dies nur Frustration und Misserfolg nach sich ziehen? In diesem Punkt gilt es einen Kompromiss zwischen den wissenschaftlich begründeten Lernzielen und der von der Praxis her bestimmten Umsetzungsmöglichkeit zu finden.

#### 4.2 Wissenschaftliche Fundierung der Maßnahme

Es wurde bereits angesprochen: Nicht alles, was die Teilnehmer eine betrieblichen Weiterbildungsmaßnahme begeistert oder doch zufrieden stellt bedingt einen anforderungsadäquaten Zugewinn an Qualifikation. Es ist also nicht nur bei der Zielformulierung an die Relevanz zu denken, sondern auch mit Rückgriff auf wissenschaftliche Untersuchungen zu prüfen, mit welchen Lernmethoden und bei welchen Lernzielen man bei welcher Zielgruppe in einem spezifischen betrieblichen Umfeld am ehesten zum Erfolg gelangen kann. Generalisiert man trotz der soeben vorgenommenen einschränkenden Differenzierungen ein wenig, so ist – damit es zum Lernen und zum Lerntransfer kommt – mit Mandl und Reinmann-Rothmeier (1995) zu fordern:

- Authentizität, d.h. die Lernumgebung sollte die reale Situation widerspiegeln
- Situiertheit, d.h. der Lernende wird in Situationen versetzt, die ihm den Anwendungskontext anschaulich vor Augen führen
- Multiple Kontexte, d.h. es wird vermieden, dass das Wissen nur auf eine Situation bezogen wird, sondern dass es auf verschiedene Kontexte angewandt wird
- Multiple Perspektiven, d.h. die Inhalte und Probleme sollten aus unterschiedlicher Sicht reflektiert werden, sowie
- sozialer Kontext, d.h. das neue Wissen sollte zumindest zum Teil gemeinsam und in kooperativer Weise erworben werden.

#### 4.3 Akzeptanz und Akzeptabilität des Vorgehens

Schon vor geraumer Zeit hat man innerhalb der psychologischen Diagnostik neben die rein funktionalen psychometrischen Gütekriterien der Auswahlverfahren wie Objektivität, Reliabilität und Validität die Ökonomie und – was in diesem Kontext zentral ist – die „soziale Validität“, also die Akzeptanz und Akzeptabilität der Maßnahmen, die z.B. Merkmale wie Transparenz und Fairness umschließen, gestellt. Dies gilt zunehmend auch für die Weiterbildung. Es gibt – insbesondere wenn die Lernziele der so genannten Persönlichkeitsentwicklung den Aufbau von Selbstsicherheit im Umgang mit Stresssituationen beinhalten – gelegentlich Vorgehensweisen, die gegen Sitte und Anstand, gegen Kriterien der Würde des Menschen oder sogar gegen das Gesetz verstoßen. Bei der Gestaltung des Vorgehens ist also darauf zu achten, dass es im Sinne der Akzeptanz die Zustimmung derer findet, die geschult, oder gefördert werden sollen, wozu Partizipation und Transparenz wesentlich beitragen können. Es gilt aber auch die Akzeptabilität zu beachten, d.h. die Übereinstimmung des Vorgehens mit informellen Normen, mit vorherrschenden Auffassungen von Sitte und Anstand, aber auch mit gesetzten Rechtsbestimmungen.

## **5. Was sollte konkret gefordert werden?**

Die soeben beschriebenen Grundsatzforderungen liegen zum Teil auf einem abstrakten Niveau. Es soll nun etwas gezielter – fast im Sinne einer Checkliste – gefragt werden, was konkret zur Qualitätssicherung der betrieblichen Weiterbildung unternommen werden kann. Dabei orientieren wir uns an den Phasen, die für die Entwicklung derartiger Maßnahmen gelten oder doch gelten sollten (Pawlowsky/Bäumer 1996).

### *5.1 Erhebung des Bildungsbedarfs*

Hier ist zu fordern, dass Lernziele und -methoden nicht aufgrund modischer Tendenzen, guter Erfahrung damit in anderen Betrieben oder persönlichen Vorlieben der Weiterbildungsverantwortlichen oder der Trainer etc. vorgenommen werden, sondern differenziert nach betrieblichen Zielgruppen auf Grund fundierter Aufgaben-, Personen- und Organisationsanalysen, wie sie soeben bereits angesprochen wurden.

### *5.2 Ermittlung des Lernpotenzials*

Nicht immer können die an sich erstrebenswerten Lernziele bei der vorgesehenen betrieblichen Zielgruppe auch umgesetzt werden. Die entsprechende Lernfähigkeit und -bereitschaft ist gelegentlich nicht vorhanden. Entsprechend sind hier fundierte Potenzialbeurteilungen (Schuler/Görlich 2006) vorzunehmen. Diese Analysen sind schwierig, weil sie letztlich nicht einen gegebenen und messbaren Ist-Zustand erfassen, sondern unsichere Vermutungen über die Zukunft zu Grunde haben. Man wird ja – um das Problem am Beispiel aufzuzeigen – von einer neunjährigen Geigerin, die in ein Konservatorium aufgenommen werden soll, kann erwarten, dass sie bereits wie eine gefeierte Violinistin spielt, aber man sucht zu beurteilen, ob sie bei entsprechender Schulung in zehn Jahren so weit sein könnte. Gerade die Ermittlung des Potenzials gewinnt in jüngerer Zeit angesichts des demografischen Wandels besondere Bedeutung. Es wird zunehmend notwendig werden, auch ältere Mitarbeiter für neue Aufgaben zu qualifizieren, wobei es gilt, die noch gegebenen Potenziale valide abzuschätzen.

### *5.3 Bestimmung der Lernziele*

Bei der betrieblichen Weiterbildung sollte für die Lehrenden und Lernenden deutlich werden, was letztlich erreicht werden soll. Es gilt also Lernziele zu definieren (Mager 1965). Hier unterscheidet sich die betriebliche Weiterbildung von der Kompetenzentwicklung, bei der in der Regel ergebnisoffen vorgegangen wird. Bei der Bestimmung der Lernziele ist nun darauf zu achten, dass sie sich schlüssig aus dem zuvor erhobenen Bildungsbedarf ergeben, transparent für die Lernenden sind und deren Akzeptanz haben,

und so formuliert werden, dass der Grad ihrer Realisierung im Zuge der Evaluation empirisch überprüft werden kann. Schließlich sollte man bei den Lernzielformulierungen auch darauf achten, dass – wie bereits ausgeführt – ihre Realisierung wahrscheinlich ist, dass also weder von den Seiten der Lernenden noch vom Lernumfeld oder vom Betrieb insgesamt keine allzu großen Hindernisse oder Barrieren der Umsetzung entgegenstehen.

#### 5.4 *Planung der Intervention*

Betriebliche Weiterbildungsmaßnahmen sollten sorgfältig geplant werden. Eindeutige Empfehlungen können kaum ausgesprochen werden, da die Art des Vorgehens von der Kultur des Betriebes, den Lerninhalten und –zielen, dem Teilnehmerkreis, der Verfügbarkeit kompetenter Lehrender und den sonstigen bereitstehenden materieller und immaterieller Ressourcen, etc. abhängt. Vor diesem Hintergrund sollten spezifisch bedacht werden

- die äußeren Rahmenbedingungen: im Betrieb oder extern (mit Verpflichtung der Lernenden auch die jeweiligen Abende miteinander zu verbringen und am Schulungsort zu übernachten)?
- Falls extern: In einem großen Seminarhotel mit Möglichkeiten der Freizeitgestaltung, der Sportausübung, des Fitnesstrainings, der Begegnung mit Personen aus anderen Unternehmen, etc. oder in einer kleineren Pension, die durch die Veranstaltung voll ausgebucht ist und in der keinerlei störende Ablenkungen zu erwarten sind?
- Wie soll mit möglichen von den Teilnehmern selbst ausgehenden Störungen umgegangen werden? Wie also sind die Regeln hinsichtlich der Handynutzung, der Bearbeitung von E-Mails und Internet oder des Rauchens zu formulieren?
- Der Zeitgestaltung: Soll die Bildungsmaßnahme als ein Block (bei Nutzung eines externen Schulungsorts meist unumgänglich) oder über die Zeit verteilt, z.B. pro Woche je ein Tag (was bei innerbetrieblichen Schulungen empfehlenswert sein kann) durchgeführt werden?
- Ist eine follow-up-Veranstaltung zur Abschätzung erster Transfereffekte vorgesehen?
- Wie soll es um die Pausenregelung stehen? Mehrere kurze oder wenige längere Pausen? Kurze oder lange Mittagspause? Selbstbestimmte Zeiten für Lektüre und Erfahrungsaustausch? Zeit für sportliche Angebote am Morgen und/oder am Abend? Weiterführung des inhaltlichen Programms oder ein Blick über den „Tellerrand“ oder aber selbstbestimmte Zeit für den Erfahrungsaustausch an den Abenden?
- Teilnehmerkreis: Wie groß sollte die Lerngruppe sein (meist gilt: je kleiner, desto besser)? Nur für Teilnehmer des Betriebs oder gemeinsam mit solchen aus anderen Betrieben oder gar anderen Branchen, um Anregung und Erfahrungsaustausch zu sichern? Falls nur mit Angehörigen des Betriebs: Homogene oder heterogene Gruppen, was die hierarchische Ebene oder den Organisationsbereich betrifft, oder be-

wusst gemischt? Oder gar geschlossene „Organisationsfamilien“, um damit den Transfer wahrscheinlicher zu machen?

- Auswahl der Lehrenden: Nur ein oder mehrere Trainer? Interne oder externe (wobei sich ein Zweierteam mit einem internen und einem externen häufig bewährt, da dem Externen meist höhere Expertenmacht zugeschrieben wird und der interne leichter Transfermöglichkeiten aufzeigen kann)? Soll ein „Machtpromoter“ die Veranstaltung eröffnen, um auf diese Weise symbolisch deren Bedeutung zu unterstreichen? Sollen vorwiegend Wissenschaftler oder Praktiker (oder eine Mischung) für die Durchführung verantwortlich sein? Sollen die Trainer auch nach der Veranstaltung als Coaches oder Supervisoren beim Transfer zur Verfügung stehen?
- Lehrmethoden: Soll vorwiegend mit inhalts- oder prozessorientierten Methoden oder einer Mischung aus beidem vorgegangen werden? Sind bestimmte Inhalte im Selbststudium zu bearbeiten? Ist dafür e-learning oder doch blended learning vorgesehen? Sind bestimmte Zeiten für Outdoortraining, für Sportprogramme oder für Exkursionen angesetzt? In wie weit werden technische Hilfen wie Videofeedback, Groupware oder „Programm-Analyser“ eingesetzt? Sind Feedbackrunden zur formativen Evaluierung oder eine summative Evaluierung am Ende vorgesehen? Gibt es an jedem Abend bzw. am Folgemorgen einen zusammenfassenden Rückblick und einen vorausschauenden Blick auf das Kommende? Erstellen die Teilnehmer selbst jeweils Zusammenfassungen („lessons learned“)?
- Lernmaterialien: Wird den Teilnehmern vor oder während des Seminars Lernmaterial zur Verfügung gestellt? In schriftlicher oder elektronischer Form? Liegen ihnen Folienkopien bereits während der Schulung vor oder werden sie erst nachträglich bereitgestellt? Stehen standardisierte Fälle für Rollenspiele oder Analysen zur Verfügung oder werden diese von den Teilnehmern selbst aus ihrer Praxis erarbeitet?
- Umgang mit Störungen: Gibt es Regeln für Störungen oder gar Krisensituationen? Was tun, wenn es zu einem „Seminarkoller“ kommt? Was tun, wenn ein Trainer, Referent oder Moderator von der Gruppe abgelehnt wird? Wie geht man damit um, wenn mehrere Teilnehmer wegen eines Problems im Unternehmen plötzlich aus dem Seminar abgerufen werden? Gibt es einen Notfallplan, wenn bestimmte, für den Verlauf des Seminars wichtige Trainer ausfallen oder Unterlagen nicht oder fehlerhaft bereitgestellt wurden?

Die Liste ließe sich fortsetzen oder ist je nach Situation anders zu akzentuieren. Auf jeden Fall gilt es im Einzelnen zu planen, wie die Maßnahme konkret gestaltet sein sollte. Andererseits sollte ausreichend Flexibilität bestehen, um angemessen auf Unerwartetes reagieren zu können.

### *5.5 Durchführung der Maßnahmen*

Bei der konkreten Durchführung der Maßnahmen – ob dies nun durch den Betrieb selbst oder durch von diesem beauftragte Institutionen oder Personen geschieht – ist

darauf zu achten, dass im Vorfeld geprüft wird, ob die Vorbereitungen angemessen erfolgten und der Planung entsprechen. Außerdem sollte für den Fall von Problemen ein Ansprechpartner des Betriebs, zumindest aber eine „Hotline“, jederzeit zur Verfügung stehen.

### 5.6 Sicherung des Transfers

Ein zentrales Problem vieler Weiterbildungsveranstaltungen besteht darin, dass das Gelernte nicht in die Praxis umgesetzt wird. Zum Teil ist dies darauf zurückzuführen, dass das Transferproblem den Verantwortlichen gar nicht bewusst ist (Neudecker 1987) oder von ihnen angenommen wird, dass sich Transfer schon „von selbst“ einstellen werde und nicht explizit zu planen sei. Tatsächlich ist jedoch eine gezielte Transfersicherung unumgänglich, die zumindest darin bestehen sollte, dass bereits während der Veranstaltung ausdrücklich über Nutzungs- und Anwendungsmöglichkeiten diskutiert und auf Möglichkeiten des Transfers hingewiesen wird. Dabei sollte einem bewusst sein, dass das Zielverhalten nach der Maßnahme – wie es für alles Verhalten gilt – abhängig ist vom

- persönlichen Wünschen und Wollen (Motivation und Volition)
- individuellen Können (Wissen, Erfahrung, Fertigkeiten, Fähigkeiten)
- sozialem Dürfen und Sollen (Kultur des Betriebs, formelle und informelle Normen, Regelungen, Präferenzen des Führenden und der Kollegen)
- situativer Ermöglichung (fördernde und hemmende Einflussgrößen in der betrieblichen Situation).

Es liegt also nicht am Wollen, Wünschen und Können des Teilnehmers der Bildungsmaßnahme allein, ob das Gelernte umgesetzt wird, sondern auch an situativen Bedingungen, die bereits bei der Lernzielformulierung bedacht werden müssen. Überspitzt formuliert sollte also jede Personalentwicklungsmaßnahme von einer entsprechenden Organisationsentwicklung begleitet sein und hier entsprechend systemisch gearbeitet werden. Dies wird auch in einem viel beachteten Transfermodell von Baldwin und Ford (1988) deutlich, in dem als Voraussetzungen des Transfers Charakteristika der Lernenden, das Design des Trainings sowie das Arbeitsumfeld in detaillierter Weise benannt werden. Vielfältige Hinweise auf transferfördernde Maßnahmen findet man bei Lemke (1959), Weinert (2004) sowie Bergmann und Sonntag (2006). Es würde zu weit führen, diese hier im Einzelnen zu referieren.

### 5.7 Evaluation der Maßnahme

Es gehört zu den sozialwissenschaftlichen Standards bei Interventionen auch zu überprüfen, ob sie ihre Ziele erreichen. Eigentlich sollte dies bei kostenintensiven Investitionen, wie sie betriebliche Weiterbildungsmaßnahmen darstellen, auch selbstverständlich sein, doch ist dies oft nicht der Fall und geschieht bestenfalls in einer unzureichenden

Weise. So ist zuvor zu entscheiden, ob man auf eine formative Evaluierung, also eine begleitende Überprüfung, um die Maßnahme selbst im Verlauf zu optimieren, oder eine summative Evaluierung, die post-hoc die Effekte prüft, setzen möchte. Bei der formativen Evaluierung lässt sich beispielsweise untersuchen, ob die soeben unter 5.1 – 5.6 genannten Forderungen jeweils in angemessener Weise realisiert wurden. Es lässt sich aber ebenfalls in einer eher qualitativen Weise mit den Teilnehmern der Maßnahme in Zwischenfeedbacks erheben, was verbesserungsbedürftig erscheint und dann unmittelbar umgesetzt werden kann.

Bei der summativen Evaluierung sollte das, was schließlich überprüft wird, dem entsprechen, was in der Lernzielformulierung festgelegt wurde. Als Kriterien eignen sich hier in Anlehnung an Kirkpatrick (1987) folgende:

- Subjektive Wertung (z.B. Hat Ihnen die Maßnahme gefallen? Glauben Sie, dass Sie das Gelernte in die Praxis übertragen können?) Derartige Erhebungen sind die in der Praxis häufig.
- Wissen bzw. Wissenszuwachs (z.B. Kenntnis neuer Bestimmungen des Arbeits- bzw. Betriebsverfassungsrechts nach der entsprechenden Informationsveranstaltung).
- Verhalten in der Lernsituation (z.B. Differenz des beobachteten Verhaltens im Kritikgespräch – z.B. beim Rollenspiel – vor oder nach einschlägigen Trainingsmodulen).
- Verhalten am Arbeitsplatz (z.B. Beurteilung des Verhaltens der Trainierten durch Vorgesetzte, Kollegen, Mitarbeiter oder Kunden).
- Resultate (z.B. Umsatzsteigerung, Zahl/Qualität der Verbesserungsvorschläge, Fehlzeiten, Fluktuationsentwicklung).

Selbstverständlich ist für die Überprüfung der Effekte und deren Zurechnung auf Weiterbildungsmaßnahmen ein adäquates Design erforderlich. Dies kann häufig in der Praxis aus ökonomischen oder mikropolitischen Gründen nicht realisiert werden (Fisch/Fiala 1984; Neuberger 1994). So dominiert in der Weiterbildungspraxis die Erhebung subjektiver Indikatoren nach der Maßnahme durch anonyme schriftliche Befragung der Teilnehmer („Hat es Ihnen gefallen?“). Es wären jedoch möglicherweise zumindest – obwohl damit noch keineswegs alle einflussnehmenden Drittvariablen ausgeschlossen werden können – eine Vorher-Nachher-Messung oder aber ein Kontrollgruppendesign erforderlich (Thierau-Brunner/Wottawa/Stangel-Meseke 2006; v. Rosenstiel 2003, 2007). In der Folge ist selbstverständlich zu fordern, dass die Ergebnisse der Evaluation wiederum der Optimierung der laufenden Maßnahme (formative Evaluierung) oder künftiger Maßnahmen (summative Evaluierung) zugute kommen

## **6. Vernetzung von Wissenschaft und Praxis zu beiderseitigem Nutzen**

Es wurde in diesem Beitrag mehrfach auf Spannungen zwischen der Wissenschaft und der Praxis hingewiesen; es wäre auch unrealistisch, diese leugnen zu wollen. Legitimer Weise bemüht sich die Wissenschaft – insbesondere die empirisch orientierte – um ein-

deutige Begriffe, saubere Operationalisierungen, fehlerfreie Messungen und adäquate Untersuchungsdesigns und entwickelt, da diese rigorosen Forderungen im Feld selten umsetzbar sind, eine Neigung zum Rückzug ins Labor. Aus vielerlei (unternehmens-) politischen, ethischen, (arbeits- und betriebsverfassungs-) rechtlichen oder ökonomischen Gründen kann der Betrieb den rigorosen Forderungen des Wissenschaftlers meist nicht entsprechen. Dennoch erschien es lohnend, sich stärker als bisher aufeinander zu bewegen. Das gilt im Interesse der Praxis und der Wissenschaft. Letztlich wird dies in dem Kurt Lewin zugeschriebenen Satz: „Nichts ist praktischer als eine gute Theorie“ angesprochen. Irle (1975, S. 506) betont, dass die Praxis der Theorie bedarf, denn „Theorien werden in Techniken, hier in Verhaltens- und Sozialtechniken transformiert“, wobei die Techniken als Hypothesen zu interpretieren sind, deren Überprüfung in der Praxis erfolgt. Die Anwendung werde somit zu einem „Experiment in der sozialen Natur“. Die dabei gewonnenen Informationen – seien sie nun quantitativer oder qualitativer Art – können damit wesentlich dazu beitragen, sowohl die Theorie als auch die Sozialtechnologien weiter zu entwickeln.

## 7. Literatur

- Baldwin, T.T./Ford, J.K. (1988): Transfer of training: A review and directions for future research. In: *Personnel Psychology* 41, S. 63–105.
- Bergmann, B./Sonntag, K. (2006): Transfer: Die Umsetzung und Generalisierung erworbener Kompetenzen in den Arbeitsalltag. In: Sonntag, K. (Hrsg.): *Personalentwicklung in Organisationen*. Göttingen: Hogrefe, S. 355–388.
- Bieling, G. (1980): Bildungsprobleme der Diplompsychologen aus der Sicht der Praxis. In: Stephan, E. (Hrsg.): *Aus- und Weiterbildung in Psychologie*. Weinheim: Beltz, S. 256–265.
- Büssing, A. (2007): Organisationsdiagnose. In: Schuler, H. (Hrsg.): *Lehrbuch Organisationspsychologie*. Bern: Huber, S. 557–590.
- Bungard, W./Wiendiek, G. (Hrsg.) (1986): *Erfolgreich im Qualitätswettbewerb. – Quality circles als Instrumente zeitgemäßer Betriebsführung*. Landsburg: Moderne Industrie.
- Conradi, W. (1983): *Personalentwicklung*. Stuttgart: Enke.
- Erpenbeck, J./Rosenstiel, L. von (Hrsg.) (2007): *Handbuch der Kompetenzmessung*. Stuttgart: Schäffer-Poeschel.
- Erpenbeck, J./Sauer, J. (2000): Das Forschungs- und Entwicklungsprogramm „Lernkultur Kompetenzentwicklung“. In: *Arbeitsgemeinschaft Qualifikations-Entwicklungs-Management (Hrsg.): Kompetenzentwicklung 2000*. Münster: Waxmann, S. 289–337.
- Fisch, R./Fiala, S. (1984): Wie erfolgreich ist Führungstraining? In: *Die Betriebswirtschaft* 44, S. 193–203.
- Frieling, E./Sonntag, K. (1999): *Lehrbuch Arbeitspsychologie*. Bern: Huber.
- Greif, S./Kurtz, H. (Hrsg.) (1996): *Handbuch Selbstorganisiertes Lernen*. Göttingen: *Angewandte Psychologie*.
- Hacker, W. (2005): *Allgemeine Arbeitspsychologie. Psychische Regulation von Wissens-, Denk- und körperlicher Arbeit*. Bern: Huber.
- Hauser, B. (2008): *Action Learning im Management Development*. München/Mering: Hampp.
- Heyse, V./Erpenbeck, J. (2004): *Kompetenztraining*. Stuttgart: Schäffer-Poeschel.
- Holling, H./Liepmann, D. (2007): *Personalentwicklung*. In: Schuler, H. (Hrsg.): *Lehrbuch Organisationspsychologie*. Bern: Huber, S. 345–383.
- Irle, M. (1975): *Lehrbuch der Sozialpsychologie*. Göttingen: Hogrefe.

- Kirkpatrick, D. (1987): Evaluation of training. In: Craig, R. L. (Hrsg.): Training and development handbook: A guide to human resource development. New York: McGraw-Hill, S. 301–319.
- Landy, F. (1985): Psychology of work behavior. Chicago: Dorsey.
- Lemke, S. (1995): Transfermanagement. Göttingen: Hogrefe.
- Mager, R. F. (1965): Lernziele und Unterricht. Weinheim: Beltz.
- Mandl, H./Reinmann-Rothmeier, G. (1995): Unterrichten und Lernumgebung gestalten. In: Forschungsbericht. München: LMU München, Inst. für Päd. Psychologie, Bd. 60.
- Mintzberg, H. (1973): The nature of managerial work. New York: Wiley and Sons.
- Neuberger, O. (1994): Personalentwicklung. Stuttgart: Enke.
- Neuberger, O. (2006a): Führen und führen lassen. Stuttgart: Lucius.
- Neuberger, O. (2006b): Mikropolitik und Moral. Stuttgart: Lucius.
- Neudecker, M. (1987): Die innerbetriebliche Führungskräftebildung. Berlin: Duncker & Humblot
- Pawlowsky, P./Bäumer, J. (1996): Betriebliche Weiterbildung. München: Beck.
- Reiß, M. (2000): Qualitätsmanagement. In: Corsten, H. (Hrsg.) (2000): Lexikon der Betriebswirtschaftslehre. München: Oldenburg, S. 821–825.
- Reiß, M./Rosenstiel, L. von/Lanz, A. (Hrsg.) (1997): Change Management. Stuttgart: Schäffer-Poeschel.
- Rosenstiel, L. von (2003): Betriebliche Personalentwicklung – „ein blinder Fleck“ für die Evaluation. Zeitschrift für Evaluation, 1, S. 151–174.
- Rosenstiel, L. von (2007): Grundlagen der Organisationspsychologie. Stuttgart: Schäffer-Poeschel.
- Rosenstiel, L. von/Molt, W./Rüttinger, B. (2005): Organisationspsychologie. Stuttgart, Berlin, Köln: Kohlhammer.
- Rosenstiel, L. von/Pieler, D./Glas, P. (Hrsg.) (2004): Strategisches Kompetenzmanagement. Wiesbaden: Gabler.
- Rüter, H.D. (1988): Personalentwicklung bei der Landesbank Rheinland-Pfalz. In: Wirtschaft und Gesellschaft im Beruf Sonderheft Oktober, S. 35–41.
- Schuler, H./Görllich, Y. (2006): Ermittlung erfolgsrelevanter Merkmale von Mitarbeitern durch Leistungs- und Potenzialbeurteilung. In: Sonntag, K. (Hrsg.): Personalentwicklung in Organisationen. Göttingen: Hogrefe, S. 235–269.
- Sonntag, K.H. (Hrsg.) (2006): Personalentwicklung in Organisationen. Göttingen: Hogrefe.
- Sonntag, K.H./Stagmaier, R./Schaper, N. (2006): Ermittlung organisationaler Merkmale: Organisationsdiagnose und Lernkultur. In: Sonntag, K.H. (Hrsg.): Personalentwicklung in Organisationen. Göttingen: Hogrefe, S. 179–205.
- Staudt, E./Kriegesmann, B. (1999): Weiterbildung: Ein Mythos zerbricht. In: Arbeitsgemeinschaft Qualifikations-Entwicklungs-Management (Hrsg.): Kompetenzentwicklung 99. Münster: Waxmann, S. 17–60.
- Thierau-Brunner, H./Wottawa, H./Stangel-Meseke, M. (2006): Evaluation von Personalentwicklungsmaßnahmen. In: Sonntag, K. (Hrsg.): Personalentwicklung in Organisationen. Göttingen: Hogrefe, S. 329–354.
- Ulich, E. (2006): Lern- und Entwicklungspotentiale in der Arbeit - Beiträge der Arbeits- und Organisationspsychologie. In: Sonntag, K. (Hrsg.): Personalentwicklung in Organisationen. Göttingen: Hogrefe, S. 138–176.
- Ulich, E. (2007): Gestaltung von Arbeitstätigkeiten. In: Schuler, H. (Hrsg.): Lehrbuch Organisationspsychologie. Bern: Huber, S. 221–251.
- Weinert, A.B. (2004): Organisations- und Personalpsychologie. Weinheim: Beltz.

*Adresse des Autors:*

Prof. Dr. Lutz von Rosenstiel, Ludwig-Maximilians-Universität München, Leopoldstr. 13, 80802 München

Andrä Wolter/Christian Kerst

# Akkreditierung als Verfahren der Qualitätssicherung von Studiengängen in Deutschland

*Eine Policy-orientierte Analyse*

## 1. Akkreditierung als Element der Hochschulreform

Seit der zweiten Hälfte der 1990er Jahre befindet sich das deutsche Hochschulsystem in einer tiefgreifenden Phase der Umgestaltung, die nahezu jedes seiner Systemelemente einbezieht – von der Struktur des Hochschulsystems über die Hochschulzulassung, die Organisation des Studiums und die Abschlüsse bis hin zur Hochschulfinanzierung und zum Steuerungs- und Managementregime der Hochschulen. Die Etablierung von Institutionen und Verfahren der Qualitätssicherung, die dem deutschen Hochschulsystem im Bereich der Lehre und des Studiums (anders als in der Forschung) bis in die 1990er Jahre hinein weitgehend fremd waren, sind ein Teil dieses Prozesses. Der weit ausgreifende Charakter der sich gegenwärtig vollziehenden Veränderungen lässt es in einer historischen und international vergleichenden Perspektive durchaus als gerechtfertigt erscheinen, hier von einer fundamentalen Transformation des deutschen Hochschulsystems zu sprechen (zum Begriff der Transformation vgl. Jarausch 1983; Wittrock 1993; Wolter 1999; Kogan u.a. 2006).

Der wichtigste Grund für diesen Umbruch ist ein kumulativ über Jahrzehnte aufgebauter Reformdruck, der teilweise durch Nicht-Reform oder auf halbem Wege stecken gebliebene Partialreformen, teilweise durch eine sich auswachsende staatliche Überregulierung bei tendenzieller Unterfinanzierung der Hochschulen, nicht zuletzt aber durch europäische Anpassungsprozesse hervorgerufen wurde. In der Studienreform, einer schon seit den 1960er Jahren virulente Aufgabe, kam es zwar in den 1970er und 1980er Jahren zu zahlreichen Empfehlungen und Modellversuchen, auch zu neuen Rahmenprüfungsordnungen, jedoch selten zu den erforderlichen grundsätzlichen Veränderungen in der Struktur der Studiengänge und in der Organisation des Studiums. Die Studienreform wurde von den deutschen Hochschulen lange Zeit eher erlitten als aktiv vorangetrieben oder gestaltet. Qualitätssicherung fasste über die traditionellen Formen der Input-Steuerung hinaus ebenfalls nur zögerlich Fuß.

Es kommt hinzu, dass neben intentional herbeigeführten, politisch angestrebten Reformen die Entwicklungsdynamik des deutschen Hochschulsystems durch eine Reihe latenter, nur teilweise direkt geplanter Veränderungen bestimmt wird. In erster Linie gilt dieses für die seit den 1960er Jahren eingetretene massive Expansion der Studiennachfrage und das starke Wachstum des gesamten Hochschulsystems, das um die Jahrtausendwende einen neuen Höhepunkt erreichte (vgl. Konsortium Bildungsberichterstat-

tung 2006, S. 105 ff.). Die Hochschulexpansion hat ja nicht nur zu einer enormen quantitativen Erweiterung, sondern zu einem Funktions- und Identitätswandel der Hochschulen<sup>1</sup> geführt, der mit einem Verlust an Einheitlichkeit durch mehr Vielfalt einherging. Das enorme Wachstum des Systems – 1960 gab es in Deutschland 75 Hochschulen (ohne Pädagogische Hochschulen) mit 220.000 Studierenden, 2003 waren es 366 Hochschulen mit knapp über 2 Millionen Studierenden – hat in Deutschland den Reformdruck in Richtung stärkerer institutioneller Differenzierung und curriculärer Diversifizierung der Programme wesentlich verstärkt, dem die tatsächliche Entwicklung in den letzten vier Jahrzehnten allerdings nur zögerlich gefolgt ist. „Expansion und Differenzierung sind komplementäre Prozesse, die den Entwicklungspfad moderner Bildungssysteme bestimmen“ (Windolf 1990, S. 109). Mit einer stärkeren Differenzierung werden jedoch neue Anforderungen an Vergleichbarkeit und Anerkennung von Institutionen oder Programmen aufgeworfen (vgl. Teichler 2005a). Akkreditierungsverfahren werden deshalb häufig als eine Antwort auf diese neue Vielfalt gesehen.

Trotz oder gerade wegen der vielfältigen Reformen mit ihren zum Teil noch unbekanntem Nebenfolgen sind die zukünftigen Konturen des deutschen Hochschulsystems – nicht zuletzt unter dem Vorzeichen seiner Re-Föderalisierung – bislang nur in Umrissen zu erkennen (vgl. Wolter 2004). Die neuen Einrichtungen und Verfahren zur Akkreditierung von Studiengängen sind ohne Zweifel ein „gewollter“ Teil dieses Wandels. Zusammen mit anderen Verfahren (z.B. der Lehrevaluation) gehört Akkreditierung – zum einen – in den übergreifenden Rahmen neuer Modelle, Institutionen und Prozeduren der Qualitätssicherung innerhalb des deutschen Hochschulsystems. Zum anderen wurde ihre Einführung insofern durch den Bologna-Prozess beschleunigt, als sich Akkreditierung in Deutschland primär auf die neuen Studiengangstypen kapriziert. Akkreditierung wird etwa seit Mitte der 1990er Jahre zu einem zentralen hochschulpolitischen Thema. Seit 1998 hat sie schrittweise festere institutionelle Gestalt angenommen (vgl. dazu Abschnitt 3).

Akkreditierung – so lautet die überzeugende Definition, die Michael Daxner vorgeschlagen hat – „(...) ist der begründete, widerrufbare und auf nachvollziehbaren Kriterien aufbauende Vertrauensvorschuss (Kredit), den die Gesellschaft bzw. der Staat den Institutionen gibt“ (Daxner 1999, S. 47).

Im internationalen Vergleich fällt immer wieder auf, dass Deutschland bei Innovationen im Hochschulsystem häufig eine „verspätete Nation“ ist. Gerade am Beispiel von Qualitätsbewertung und -sicherung lassen sich exemplarisch einige der historisch gewachsenen Besonderheiten des deutschen Hochschulsystems aufzeigen (vgl. Teichler 2002a).

- Ihre starke Forschungsorientierung galt in der Humboldtschen Traditionslinie der Einheit von Forschung und Lehre weithin als eine international anerkannte Stärke der deutschen Universitäten, welche gleichsam als Nebenprodukt eine hohe Qualität

1 In der viel zitierten Formulierung von Martin Trow aus dem Jahr 1973: „from elite to mass higher education“ (Trow 1973).

der akademischen Lehre verbürgte. Forschungsbasierung der Lehre wurde daher gleichsam als funktionales Äquivalent anderer Verfahren der Qualitätssicherung gesehen. Tatsächlich haben sich jedoch mit der zunehmenden Spezialisierung der Forschung, unter dem Druck der „Massenuniversität“ und mit der „Curricularisierung“ der Studiengänge Forschung und Lehre immer weiter voneinander entfernt. Unter den Bedingungen eines funktionsdifferenzierten Hochschulsystems sichert hohe Forschungsqualität allein keine entsprechend Lehrqualität mehr, weil sich der alltägliche Lehrbetrieb weitgehend von einem unmittelbaren Forschungsbezug entkoppelt hat.

- Eng damit verbunden war die Annahme, dass sich die deutschen Hochschulen – innerhalb ihrer beiden Sektoren, den Universitäten und den Fachhochschulen – durch eine weitgehende Einheitlichkeit der Studiengänge und Standards auszeichnen und sich daher auch in ihrer Qualität nur wenig voneinander unterscheiden. Anders als in stark stratifizierten Hochschulsystemen waren Rangabstufungen zwischen Hochschulen in Deutschland mehr von informeller als von substanzieller Bedeutung. Tatsächlich zeigen sich jedoch erhebliche Unterschiede nicht nur in der Reputation, sondern auch in der Qualität der Studienangebote (und in der Praxis der Lehre) zwischen den wie auch innerhalb der einzelnen Hochschulen. Allerdings waren solche Unterschiede wenig transparent, und sie hatten keine strukturelle Bedeutung. Einheitlichkeit als Norm hat sich im Zeichen von Expansion und Differenzierung überlebt.
- Zu den internationalen Auffälligkeiten des deutschen Hochschulsystems gehört auch die dominante Rolle des Staates, der nicht zuletzt durch rechtliche Normierung und durch seine Verfahren der Input-Steuerung für eine Mindestqualität und hohe Gleichförmigkeit sorgen sollte. Auch hier hat sich eine gewisse Ernüchterung eingestellt. Die staatliche Regulierung ist nicht nur oft in Überregulierung umgeschlagen, die Innovationen eher behinderte als beförderte. Eine noch so machtvolle staatliche Steuerung hat im Übrigen nicht verhindern können, dass die Qualität des Studiums keineswegs über jeden Zweifel erhaben ist. Institutionelle Selbststeuerungsfähigkeit der Hochschulen ist in Deutschland tendenziell unterentwickelt. Die „Kritik am traditionellen System der staatlichen Genehmigung“ mit umständlichen und ermüdenden Prozeduren ist ein starkes Motiv für die Einführung von Akkreditierungsmechanismen in Deutschland gewesen (vgl. Kehm 2007, S. 80).

## **2. Kontexte der Akkreditierung: Qualität, Steuerung, Europäisierung und Differenzierung**

In der Einführung und Institutionalisierung studiengangsbezogener Akkreditierung in Deutschland manifestierten sich mehrere Entwicklungsstränge aus der hochschulpolitischen Reformdebatte der letzten 15 Jahre. An erster Stelle ist hier gewiss die Funktion von Akkreditierung im Kontext von Qualitätssicherung zu nennen. Zweitens bedeutet

Akkreditierung einen veränderten Steuerungsmodus, von dem eine – vorsichtige – Deregulierung ausgeht, indem bislang von staatlichen Behörden wahrgenommene Funktionen auf neue intermediäre Einrichtungen übergehen. Drittens hängt die Etablierung von Akkreditierungsverfahren eng mit dem Bologna-Prozess, überhaupt mit der wachsenden Internationalisierung der Hochschulen zusammen. Und viertens schließlich steht Akkreditierung in einem Zusammenhang mit der stärkeren institutionellen Differenzierung des Hochschulsystems.

*Studiengangsbezogene Qualitätssicherung:* Martin Trow hat das Aufkommen des Qualitätsthemas als Anzeichen dafür interpretiert, dass eine Gesellschaft ihr Vertrauen verliert, Hochschulen würden ohne intervenierende Maßnahmen gleichsam automatisch „Produktqualität“ garantieren (vgl. Trow 1996), auch wenn umgekehrt das Nicht-Vorhandensein spezieller Verfahren von Qualitätssicherung selbstverständlich nicht als Zeichen eines besonderen Vertrauens interpretiert werden kann. Im Blick auf die spezifisch deutsche Entwicklung lässt sich hinzufügen, dass auch das Vertrauen in die Wirksamkeit der herkömmlichen Instrumente staatlicher Input-Steuerung abgenommen hat. Seit den späten 1980er Jahren setzte sich das Qualitätsthema in Deutschland in verschiedenen Facetten mehr und mehr in Hochschulpolitik, Hochschulforschung und Hochschuldidaktik durch – zunächst in der Variante „Kriterien und Praxis guter Lehre“.

Studiengänge und ihre Qualität standen zunächst ganz am Rande, wenn sie überhaupt schon vorkamen. In den 1990er Jahren lässt sich dann eine zunehmende Etablierung von Verfahren der Lehrevaluation beobachten, sowohl an einzelnen Hochschulen wie auch in hochschulübergreifenden Verbänden und Einrichtungen. Dabei löste sich der Horizont langsam von der Evaluation einzelner Lehrveranstaltungen und der Praxis der Lehre und erweiterte sich in Richtung Studium und Studiengangsevaluation (vgl. Reissert 1994; Carstensen/Reissert 1995). Mit dem Bologna-Prozess kamen der Studiengang und dessen Struktur endgültig in den Brennpunkt des Interesses, wobei in Deutschland durchaus schon weit vor der legendären Bologna-Konferenz (1999) eine lebhaftige Debatte über eine Studienstrukturreform und die Einführung eines konsekutiven Studienmodells geführt wurde.

Der Paradigmenwechsel „vom Lehren zum Studieren“ hat ohne Zweifel die Institutionalisierung von Akkreditierungsverfahren befördert. Kennzeichnend für Deutschland ist jedoch, dass in den letzten 10 bis 15 Jahren sowohl auf Hochschul- wie auf Länderebene unterschiedliche, häufig parallele Verfahren und Einrichtungen der Qualitätssicherung etabliert wurden, bedingt nicht zuletzt durch eigenständige Wege der Länder (vgl. Hartwig 2003). Dadurch ist ein System „multipler Qualitätsbewertung“ entstanden, das gelegentlich schon Züge „eines superkomplexen Systems“ zeigt (vgl. Teichler 2003, S. 213 ff.), verbunden mit der Entstehung einer „neuen Qualitätsbürokratie“ (vgl. Nickel 2006). So können Qualitätsbewertung, Evaluation und Qualitätssicherung nicht nur auf unterschiedlichen institutionellen Ebenen stattfinden – von der einzelnen Lehrveranstaltung über einen Studiengang, eine Fakultät bzw. einen Fachbereich bis hin zu ganzen Hochschulen – sie stehen auch in sehr unterschiedlichen Handlungskontexten, die von Berufungsentscheidungen über Studienangebots- und Programmplanung bis hin zur Hochschulfinanzierung und Mittelverteilung reichen.

Zu dieser Vielfalt zählen neben der Akkreditierung von Studiengängen die herkömmlichen Verfahren von Lehrevaluation und studentischer Lehrveranstaltungs-kritik, studiengang- bzw. hochschulübergreifende Evaluationsnetzwerke, medienwirksam aufbereitete Rankingverfahren oder Lehrberichte aller Art. Im Kontext der neuen Steuerungsverfahren geht es um indikator- bzw. kennzahlenorientierte Mittelzuweisung, die immer auch eine evaluative Komponente enthält, um die Evaluation von Zielvereinbarungen oder Hochschulverträgen und neuerdings – als Konsequenz leistungsbezogener Besoldung – um individuelles „assessment“ von Hochschullehrerleistungen. Darüber hinaus werden von Zeit zu Zeit einzelne Hochschulen oder sogar die ganze Hochschullandschaft eines Landes einer Qualitätsbewertung unterzogen. Schließlich bestehen auch die schon seit langem vorhandenen Formen anlassbezogener Qualitätsbewertung (z.B. bei Forschungsanträgen, Berufungen oder Publikationen) fort (vgl. Teichler 2003).

*Akkreditierung als Element des neuen Steuerungsmodells:* In vielen OECD-Ländern sind in den letzten 10 bis 15 Jahren Reformen durchgeführt worden, mit denen neue Einrichtungen und Verfahren der Steuerung von Hochschulen auf der System- oder der institutionellen Ebene etabliert wurden (vgl. OECD 2003). Diese Reformen, häufig unter dem Begriff „governance“ diskutiert und analysiert (vgl. Braun/Merrien 1999; Kehm/Lanzendorf 2005, 2007), verfolgen in der Regel das Ziel, den Umfang der staatlichen Autorität und Kontrolle zu verringern und die institutionelle Autonomie und Verantwortlichkeit der Hochschulen und die Bedeutung wettbewerbsförmiger Mechanismen zu vergrößern. In der vergleichenden Hochschulforschung wird seit Burton Clarks viel zitiertem „triangle of coordination“ (vgl. Clark 1983, S. 143) gerade Deutschland immer wieder als Musterbeispiel für ein Hochschulsystem zitiert, in welchem staatlicher Autorität und akademischer Oligarchie ein sehr starker, Markt und Universitätsmanagement (ein Akteur, der in Clark's Dreieck nicht explizit erwähnt wird) dagegen ein schwacher Einfluss zukommt. Vor diesem Hintergrund besteht die mit der Einführung des neuen Steuerungsmodells in Deutschland verbundene Zielsetzung genau darin, in diesem Kräftefeld, eher ein Parallelogramm als ein Dreieck, den Einfluss der beiden bislang einflussreichsten Akteure – Staat und „academic community“ – zu verringern und demgegenüber den der anderen beiden Akteure, Management und Markt, zu verstärken (vgl. Wolter 2007).

In den späten 1960er und in den 1970er Jahren erfuhr der staatliche Einfluss im Hochschulsystem in der Folge des starken Wachstums, der neuen Hochschulgesetzgebung und der staatlichen Hochschulreformen einen erheblichen Zuwachs – in Form zunehmender Verrechtlichung (durch die Hochschulgesetze und Rechtsprechung), Bürokratisierung (durch administrative Detailsteuerung) oder anderer Regulierung (z.B. durch kameralistische Haushaltsführung). Die staatliche Kapazität stieß mehr und mehr an ihre Grenzen, ein zunehmend komplexeres Hochschulsystem und dessen sich verschärfende Entwicklungsprobleme überhaupt noch „beherrschen“ zu können. Insofern bildet Deregulierung oft einen eher reaktiven Handlungstypus, der primär der Erfahrung eigener begrenzter Handlungsfähigkeit entspringt.

Das sogenannte neue Steuerungsmodell sieht nun eine Reihe neuer Verfahren vor, welche direktive staatliche Steuerung reduzieren und die Selbststeuerungsmöglichkeiten

der Hochschulen vergrößern sollen. Input-orientierte Steuerung (z.B. durch die o.g. Form der Studiengangsgenehmigung oder die Praxis einer weitgehend leistungsunabhängigen, inkrementalistisch-diskretionären Mittelallokation) sollen von stärker ergebnis- und leistungsorientierten Steuerungsverfahren abgelöst werden; an die Stelle administrativer Detailsteuerung soll zielorientierte, strategische Steuerung treten. Die verschiedenen Verfahren, die im Rahmen des neuen Steuerungsmodells praktiziert werden, lassen sich zu zwei Hauptgruppen zusammenfassen; die Terminologie ist in Deutschland nicht ganz einheitlich: (1) verschiedene Formen des Kontraktmanagements (z.B. Hochschulverträge, Leistungs- und Zielvereinbarungen) und (2) die Modelle und Instrumente einer stärker leistungsorientierten, indikatorgestützten Mittelverteilung, verbunden mit einer höheren Flexibilität der Haushaltsführung (vgl. Leszczensky 2004; Orr 2007; Orr/Jäger/Schwarzenberger 2007; Wolter 2007; Ziegele 2002, 2006).<sup>2</sup> Damit gehen erhebliche Verschiebungen in der Autoritätsstruktur und im Kräftefeld der Hochschulen einher, insbesondere zwischen „academic community“ und institutionellem Management, wovon in erster Linie die neue Klasse der „executives“ in der Hochschule (vgl. Wolter 1999, S. 40ff.), der neuen managementorientierten Hochschulprofessionen, insbesondere auf der zentralen Leitungsebene, profitiert, die zum wichtigsten Träger des neuen Steuerungsmodells geworden ist.

Qualitätsentwicklung findet damit unter neuen Rahmenbedingungen statt. Sie wird zunächst stärker als je zuvor eine Gestaltungsaufgabe der Hochschulen, ohne allein auf die herkömmliche staatliche Input-Steuerung angewiesen zu sein. Diese Form der Deregulierung – vom Staat zu den Hochschulen – geht dabei in vielen Staaten gleichsam nach dem Gegenstromprinzip mit der Einführung neuer Instanzen und Verfahren der Kontrolle und Koordination zwischen Hochschule und Staat einher. Die maßgebliche Idee dabei besteht darin, eine stärkere Rechenschafts- und Transparenzpflicht der Hochschulen hinsichtlich der Qualität ihrer Leistungen, hier in der Lehre und Ausbildung, gegenüber Staat und Gesellschaft zu verankern. Mehr Autonomie und größere Rechenschaftspflicht werden dabei als zwei Seiten derselben Medaille angesehen. In den 1990er Jahren hatte sich hierfür das Konzept „accountability“ eingebürgert, ein mehrdeutiger Begriff, der im Kern die ex-post-Legitimation gesellschaftlicher Ressourcen durch Leistungstransparenz, Nachweis der Zielerreichung und effektiver und effizienter Leistungserbringung meint (vgl. Hüfner 1991). Ergebnis- und wirkungsorientierte Steuerung verlangt jedenfalls nach neuen Formen der Koordination und der Rechenschaftslegung. Dabei kann es sich um beteiligungsorientierte Formen der Koordination (z.B. durch Hochschulräte) oder um neue Formen des Berichtswesens (z.B. Hochschulberichte, Kosten-/Leistungsrechnung, Leistungs- und Ausstattungsvergleiche) handeln.

In Deutschland wie in vielen anderen Staaten bildet die Einführung nationaler oder dezentraler Mechanismen der Qualitätssicherung, der Evaluation und Akkreditierung einen festen Teil dieses Prozesses. Bei der Akkreditierung sind historisch zumeist solche

2 Modelle und Verfahren der externen Steuerung – auf der Ebene zwischen Land und Hochschule – finden ihre Fortsetzung in analogen Praktiken der hochschulinternen Steuerung, worauf hier aber nicht weiter eingegangen werden soll.

Staaten vorangegangen, in denen eine eher heterogene Hochschullandschaft und eine gering ausgeprägte öffentliche Verantwortung für die Hochschulen bestehen. Während Akkreditierung in der „bunten“ Hochschullandschaft der USA ein schon seit Jahrzehnten gebräuchliches, eher staatsfernes Verfahren ist (vgl. Myers u.a. 1998), haben inzwischen auch solche OECD-Länder, in denen der Staat traditionell eine starke Rolle als Steuerungsinstanz im Hochschulsystem ausübte, Evaluation und Akkreditierung als Teil einer stärkeren institutionellen Selbststeuerung der Hochschulen eingeführt oder verstärkt.

*Bologna-Prozess und Internationalisierung:* Der Bologna-Prozess hat die Einführung von Institutionen und Verfahren der Qualitätssicherung in Deutschland zusätzlich zu den gleichsam autochthonen Anstößen unterstützt. Von vornherein umfasste das ambitionierte Projekt, einen gemeinsamen Europäischen Hochschulraum zu schaffen, über die Struktur der Studiengänge und der Abschlüsse hinaus auch die Qualitätssicherung als zentrales Aktionsfeld (vgl. Teichler 2005b; Wolter 2006). Schon die ursprüngliche Bologna-Deklaration (1999), die am Ausgang dieses Prozesses steht, postulierte das Ziel, die europäische Zusammenarbeit bei der Qualitätssicherung zu fördern. Die nachfolgenden Bologna-Konferenzen – von Prag (2001) bis London (2007) – griffen dieses Ziel immer wieder auf. So wurde in Prag erstmals, wenn auch noch in recht allgemeiner Form, die Notwendigkeit herausgestellt, Szenarien für die wechselseitige Akzeptanz von Evaluations- und Akkreditierungsmechanismen zu entwickeln. Das Communiqué der Berlin-Konferenz (2003) forderte explizit, dass nationale Qualitätssicherungssysteme neben Evaluation immer auch Akkreditierung oder Zertifizierung als wesentliche Elemente umfassen müssen. Im Communiqué der Bergen-Konferenz (2005) wird nachdrücklich die Bedeutung nationaler Agenturen für diesen Prozess hervorgehoben.

In Bergen wie in der bislang letzten Nachfolgekonferenz in London (2007) wurden die vom European Network for Quality Assurance in Higher Education (ENQUA) entwickelten „Standards and Guidelines for Quality Assurance in the European Higher Education Area“ (deutsche Fassung vgl. Hochschulrektorenkonferenz 2006) explizit zu einem zentralen Teil des Bologna-Prozesses erhoben. In Europa – wie im übrigen auch in der außereuropäischen Hochschulwelt – hat sich in den letzten Jahrzehnten eine Vielzahl unterschiedlicher Modelle der Qualitätssicherung durch Evaluation und Akkreditierung ausgebildet, einschließlich einer Reihe nationaler, zum Teil dezentraler Agenturen, die für die Qualitätssicherung zuständig sind (vgl. Schwarz/ Westerheijden 2004a). Die Empfehlungen der ENQUA zur internen und externen Qualitätssicherung sowie für europäische Standards für externe Qualitätsagenturen spiegeln nicht nur das Bemühen um eine stärkere europäische Koordinierung wider, basierend auf gemeinsamen Standards und Kriterien, als Teil des europäischen Harmonisierungs- und Mobilitätsprojekts. Darüber hinaus zeigen diese Empfehlungen, dass sich neben der Programm- und Studiengangorientierung die Frage der Anerkennung externer Qualitätssicherungsagenturen und die nach den hochschuleigenen Qualitätsmanagementkonzepten mehr und mehr in den Vordergrund schiebt.

In Deutschland wurde die Studienreform als Aktionsfeld des Bologna-Prozesses mit der ebenfalls durch Bologna aufgeworfenen Frage der Qualitätssicherung in der Weise

verknüpft, dass sich das neue Verfahren der Akkreditierung im wesentlichen auf die neuen Studiengangsformen mit den Abschlüssen Bachelor und Master konzentriert (und nur im Ausnahmefall andere Abschlüsse umfasst). Die Einführung von Akkreditierungsmechanismen ist hier vor dem Hintergrund zu sehen, dass in Deutschland die Besorgnis weit verbreitet war, ein Studium an deutschen Hochschulen sei angesichts der international wenig kompatiblen Studiengangsstrukturen und Abschlüsse und der fehlenden Kontrolle und Transparenz hinsichtlich der Qualität der Lehre und der Studienangebote für ausländische Studienbewerber/-innen wenig attraktiv.

Gerade Akkreditierung und Zertifizierung von Studiengängen galten neben weltweit anerkannten Abschlüssen als eine wesentliche Voraussetzung für hohe internationale Akzeptanz und Konkurrenzfähigkeit. International kompatible Abschlüsse und akkreditierte Studiengänge sollten die erwünschte Mobilität von Studierenden nach Deutschland („incoming“) fördern. Tatsächlich haben sich die deutschen Hochschulen schon vor Bologna erfolgreicher auf dem globalen Bildungsmarkt behaupten können, als es den Befürchtungen entsprach. Zahl und Anteil ausländischer Studienanfänger/-innen sind seit Anfang der 1990er Jahre stark angewachsen, auch unter Abzug der sogenannten „Bildungsinländer/-innen“. Seit 2004 ist die Zahl, ironischerweise mit fortschreitender Implementation der neuen Studiengänge, leicht rückläufig (vgl. DAAD/HIS 2007, S. 24; Konsortium Bildungsberichterstattung 2006, S. 273). Akkreditierung kann aber für die Hochschulen durchaus eine Art Marketinginstrument sein, mit dem sie auf dem internationalen Bildungsmarkt werben können.

*Differenzierung und Profilbildung:* Die herkömmliche Form der staatlichen Genehmigung von Studiengängen, normiert an Rahmenprüfungsordnungen, führte nicht nur zu einer ausgeprägten Standardisierung der Programme, sondern auch zu umständlichen und langwierigen Prozeduren für Studiengangsreformen. Die Entwicklung besonderer institutions- oder programmspezifischer Studiengangsprofile war unter diesen Umständen kaum möglich; vielmehr folgten Studiengänge im Wesentlichen dem Muster „to offer more of the same“. Dieser Zwang zur Vereinheitlichung biss sich mehr und mehr mit der gegenläufigen Erwartung einer stärkeren Differenzierung innerhalb des deutschen Hochschulsystems, die auf der Ebene der Institutionen und der Studienangebote die Möglichkeit stärkerer Profilbildung einschließt. Grundsätzlich können zwei „Achsen“ der Differenzierung zwischen Hochschulen unterschieden werden: eine horizontale und eine vertikale (vgl. Teichler 1999, 2002b).

- *Horizontale Differenzierung* bezieht sich in erster Linie auf eine Vielfalt von (gleich-rangigen) Hochschultypen oder auf bestimmte Schwerpunkte im Fächerspektrum und in den Studienangeboten, auf besondere curriculare „Profillinien“ (z.B. Praxisbezug, Internationalität, Interdisziplinarität oder spezielle Lern- und Lehrformen wie Projektstudium oder medienbasiertes Studieren) oder andere institutionelle Varianten. Diese Form der Vielfalt wird häufig auch mit dem Begriff der Diversifizierung bezeichnet.
- *Vertikale Differenzierung* richtet sich dagegen primär auf die Position einer Hochschule innerhalb eines formell oder informell stratifizierten Hochschulsystems und

beinhaltet vorrangig Unterschiede in der Qualität oder Reputation von Hochschulen, die in einer Rangskala abgebildet werden können („ranking“). In Deutschland sind Ansätze vertikaler Differenzierung vor allem mit der Exzellenzinitiative verbunden. Auch die bisherige Differenzierung zwischen verschiedenen Hochschultypen (Universität, Fachhochschule) entspricht diesem Typus, weicht jedoch mit dem Bologna-Prozess möglicherweise einer stärkeren Konvergenz.

Horizontale Differenzierung muss nicht zwingend mit solchen Rangabstufungen einhergehen, wird aber öffentlich häufig so wahrgenommen (z.B. gelten herausragende Forschungsleistungen immer noch mehr als entsprechende Lehrleistungen). Angesichts eines steigenden Wettbewerbs- und Exzellenzdrucks führt das dazu, dass Hochschulen primär solche Schwerpunkte, Angebote und Leistungen fördern, die ihnen einen erkennbaren Vorteil im Wettbewerb und im „ranking“ versprechen. Profilbildung ist zwar zunächst nur eine Strategie der horizontalen Differenzierung, erfolgt aber zumeist mit der Absicht, Status und Prestige einer Hochschule zu erhöhen. Ein häufiger Anlass für Profilbildung ist die durch stagnierende staatliche Finanzierung erzwungene stärkere Konzentration knapper Ressourcen. Aber nicht jede Form von Mittelkonzentration durch fachliche Schwerpunktbildung bedeutet per se auch schon Profilbildung.

Differenzierung und fachliche Konzentration führen erst dann zur Profilbildung, wenn bestimmte zusätzliche Ansprüche und Anforderungen erfüllt werden. So lässt sich nach Teichler (vgl. 2001, 2002b) von einem *Profil* erst dann sprechen, wenn drei Bedingungen erfüllt sind: (1) es muss ein gewisses Maß an horizontaler Besonderheit vorhanden sein, das mit vertikalen Einstufungen in einer Rangfolge verbunden sein kann; (2) die das Profil bestimmenden Merkmale gelten für die Institution als ganzes oder zumindest überwiegend; und (3) müssen diese Charakteristika auch in den Leistungen der Hochschule zum Ausdruck kommen. Teichler prognostiziert, dass entlang dieser Dimensionen zukünftig „(...) die Konfiguration der Differenzierung des deutschen Hochschulwesens immer komplexer (...)“ (Teichler 2002c, S. 33) wird. Unter diesen Voraussetzungen versagen aber die klassischen Mechanismen Input-orientierter Qualitätssicherung unter anderem durch weitgehend einheitliche Studiengangsnormierung. Statt Einheitlichkeit ist Vergleichbarkeit oder Gleichwertigkeit angesagt. Externe Akkreditierung von Studiengängen ist einer jener Mechanismen, dieser neuen Komplexität und dem Gebot der Profilbildung gerecht zu werden und den Hochschulen einen größeren Handlungsraum für innovative Studienangebote oder didaktische Organisations- und Realisierungsformen zu eröffnen.

### 3. Akkreditierung in Deutschland: Formen, Institutionen und Verfahren

In Deutschland lassen sich zurzeit vier verschiedene Formen der Akkreditierung unterscheiden: (1) die Akkreditierung von *Studiengängen* primär in den Varianten der Programm- oder Clusterakkreditierung, (2) die Akkreditierung von *Agenturen* der Qualitätssicherung und Akkreditierung, (3) die Akkreditierung von *Qualitätsmanagement-*

*konzepten* von Hochschulen (häufig auch als Systemakkreditierung bezeichnet) sowie (4) die institutionelle Akkreditierung von *Hochschulen*. Die ersten drei Formen sind eng miteinander verknüpft, sie stehen im Zentrum des deutschen Akkreditierungssystems und der folgenden Ausführungen. Die vierte Variante, die *institutionelle Akkreditierung* nimmt insofern eine Sonderstellung ein, als sie sich bislang auf den Sektor der nicht-staatlichen Hochschulen (vulgo, rechtlich aber nicht ganz präzise, „private“ Hochschulen genannt) beschränkt. In Deutschland sind zurzeit etwa 100 Hochschulen, das sind 30 % aller Hochschulen, in nicht-staatlicher Trägerschaft, davon etwas weniger als die Hälfte in kirchlicher, die übrigen in anderer nicht-staatlicher Trägerschaft. Insbesondere der Sektor der nicht-kirchlichen, oft wirtschaftsnahen Hochschulen weist eine hohe Dynamik auf. Der Studierendenanteil des gesamten nicht-staatlichen Sektors fällt mit ca. 4 % deutlich geringer aus. Nicht-staatliche Hochschulen können sich als Voraussetzung ihrer staatlichen Anerkennung durch den Wissenschaftsrat akkreditieren lassen, wofür dieser inzwischen ein relativ elaboriertes Verfahren entwickelt hat (vgl. Wissenschaftsrat 2000; Goll 2006).

Die hochschulpolitische Funktion der *Studiengangsakkreditierung* besteht vorrangig in der Verringerung des staatlichen Einflusses auf die Studiengangsentwicklung. Die Organisation des Akkreditierungsprozesses umfasst vier Kernelemente. Erstens überträgt der Staat seine bisherige Genehmigungsaufgabe partiell auf neue unabhängige intermediäre Einrichtungen, den Akkreditierungsrat (AR) und externe Agenturen. Zweitens kooperieren bei der Akkreditierung die relevanten Akteure, nämlich Staat (vertreten durch die Kultusministerkonferenz), Hochschulen einschließlich der Studierenden und Repräsentanten/-innen der Berufspraxis miteinander. Die Aufgabe des Staates besteht dabei drittens primär darin, den rechtlichen und prozeduralen Rahmen des Akkreditierungsprozesses durch entsprechende Vereinbarungen bzw. Beschlüsse der Kultusministerkonferenz zu regeln, die Prozesszuständigkeit aber zu delegieren. Und viertens wird die Akkreditierung selbst in einem mehrstufigen Verfahren organisiert. Dieses komplexe System der Akkreditierung soll zugleich der Deregulierung bisheriger Staatsfunktionen, der Professionalisierung der Studiengangsbewertung durch Fachexperten/-innen, der Stärkung der Verantwortlichkeit der Hochschulen und der Mitwirkung gesellschaftlicher Akteure dienen.

Das Akkreditierungssystem in Deutschland ist zweistufig aufgebaut; unter Einschluss der für die Studiengänge verantwortlichen Hochschulen müsste man eigentlich von einem dreistufigen Aufbau sprechen (detaillierter Überblick bei Bretschneider/Wildt 2007). Der AR akkreditiert Studiengänge nicht selbst (bzw. nur in Ausnahmefällen), dies übernehmen Akkreditierungsagenturen, die für die Akkreditierung der Studiengänge zuständig sind und untereinander im Wettbewerb stehen sollen, obgleich von einem „Akkreditierungsmarkt“ (vgl. Lüthje 2004) nicht wirklich gesprochen werden kann. Der AR wurde 1998 von der Kultusministerkonferenz (KMK) zunächst probeweise als unabhängige Einrichtung mit der Aufgabe eingerichtet, Standards und Anforderungen für die Akkreditierung von Studiengängen und externen Agenturen im Rahmen der Vorgaben der KMK zu entwickeln und solche Agenturen zu begutachten und zu akkreditieren. Er nahm im Juli 1999 seine Arbeit auf. Nach einer dreijährigen Pilotphase

wurde der AR nach einer Evaluation im Jahr 2002 auf Dauer eingerichtet (vgl. Schade 2004). Einer Vereinbarung der KMK entsprechend erfolgte im Jahr 2005 die Umwandlung des AR, bis dato eine nicht rechtsfähige Einrichtung, in eine rechtsfähige Stiftung des öffentlichen Rechts nach nordrhein-westfälischem Landesrecht.<sup>3</sup> Damit waren die rechtlichen Voraussetzungen dafür geschaffen, dass der AR die den Ländern obliegende Verpflichtung zur Gewährleistung der nationalen und internationalen Vergleichbarkeit der Studienabschlüsse übernimmt. Der AR ist im Übrigen in internationale Netzwerke der Qualitätssicherung eingebunden.

Dem AR gehören Vertreter/-innen der Hochschulen, der Länder, der beruflichen Praxis, Studierende, internationale Experten/-innen sowie (mit beratender Stimme) die Agenturen an. Zu seinen wichtigsten Beschlüssen zählen die Festlegung der Anforderungen an die Akkreditierung von Studiengängen, die sich an den Kriterien für die Einrichtung von Bachelor- und Masterstudiengängen orientieren, wie sie von der KMK beschlossen wurden, sowie die Entwicklung eines Verfahrens zur Akkreditierung und Re-Akkreditierung der Akkreditierungsagenturen (gelegentlich als „Meta-Akkreditierung“ bezeichnet). Da wissenschaftlich und hochschulpolitisch keineswegs unumstritten ist, was akademische „Qualität“ eigentlich ist und wie sie gemessen werden kann, „Qualität“ vielmehr wie alle theoretischen Konstrukte überhaupt erst durch unterschiedliche Indikatoren und multiple Messverfahren operationalisiert werden muss, kommt solchen Entscheidungen durch die Agenturen eine zentrale operative Bedeutung zu. Zu den Regeln der Akkreditierung gehört, dass sowohl Agenturen als auch Studiengänge jeweils nur für einen festgelegten Zeitraum akkreditiert werden und dann – etwa nach fünf Jahren – der Re-Akkreditierung bedürfen.

Derzeit sind vom AR sechs Agenturen akkreditiert. Drei der Agenturen sind auf einzelne Fachrichtungen spezialisiert (FIBAA: Wirtschaftswissenschaften; ASIIN: Ingenieur- und Naturwissenschaften; AHPGS: Heilpädagogik, Gesundheit, Pflege, soziale Arbeit). Die anderen drei (AQAS, ACQUIN und ZEvA) sind fächerübergreifend angelegt und eher regional verortet, wenn auch grundsätzlich offen (ZEvA: Niedersachsen; AQAS: vorwiegend Hochschulen aus Nordrhein-Westfalen und Rheinland-Pfalz; ACQUIN: überwiegend Hochschulen aus Bayern, Baden-Württemberg und den neuen Ländern). Die Agentur ZEvA, hervorgegangen aus der Evaluationsbewegung, bietet als einzige Agentur zugleich Akkreditierungs- und Evaluationsdienstleistungen an. Die Agenturen haben einen ähnlichen Aufbau. Mit Ausnahme der FIBAA (Schweizer Stiftung) und der ZEvA (gemeinsame Einrichtung der niedersächsischen Hochschulen und des Landes) handelt es sich um eingetragene Vereine, mit Hochschulen und/oder Verbänden als Mitgliedern. Alle Agenturen sind in ihrer Arbeit an die Vorgaben, Anforderungen und Kriterien gebunden, die der AR festgelegt hat und natürlich an die Rahmenregelungen der KMK für die neuen Studiengänge (vgl. ausführlich Schade 2004).

3 Durch das Gesetz zur Errichtung einer ‚Stiftung zur Akkreditierung von Studiengängen in Deutschland‘ vom 15. Februar 2005.

Die Hauptaufgabe der Akkreditierung von Studiengängen besteht in der Gewährleistung fachlich-inhaltlicher Mindeststandards, der Berufs- und Arbeitsmarktrelevanz, der viel zitierten Beschäftigungsfähigkeit – eine keineswegs unproblematische Schlüsselkategorie des Bologna-Prozesses (vgl. Schindler 2004) – sowie in der Sicherstellung international vergleichbarer bzw. gleichwertiger Studienabschlüsse. Innerhalb der Agenturen entscheidet eine Akkreditierungskommission auf der Grundlage eines Gutachtertutums über die Akkreditierung, die auch mit Vorbehalt und Auflagen ausgesprochen werden kann. Sie ist ähnlich wie der AR zusammengesetzt. Die Akkreditierungsverfahren für die einzelnen Studiengänge selbst werden von Gutachterteams durchgeführt, die von der Akkreditierungskommission berufen werden. Diese bestehen aus unabhängigen Fach- und Hochschulvertretern sowie Repräsentanten der Berufspraxis und der Studierenden. Das Verfahren selbst ist relativ standardisiert und folgt im Wesentlichen dem schon bekannten Modell mehrstufiger Evaluation mit Hilfe von „peer reviews“ (vgl. Daniel/Mittag/Bornmann 2003; Bornmann/Mittag/Daniel 2006). Auf einen schriftlichen Akkreditierungsantrag der Hochschule mit einer Selbstdokumentation des Studiengangs wird in der Regel zunächst geprüft, ob und inwiefern die formalen Bedingungen an eine Akkreditierung erfüllt sind. Nach einer vertraglichen Vereinbarung zwischen Hochschule und Agentur – Akkreditierung ist für die Hochschulen kostenpflichtig – erfolgt die eigentliche externe Begutachtung durch das Gutachterteam, das anschließend einen Bericht mit einer Empfehlung für die Akkreditierungskommission erarbeitet. Nach der Entscheidung der Akkreditierungskommission vergibt die Agentur ein Gütesiegel. In vielen Ländern ist allerdings der letzte Schritt immer noch die abschließende staatliche Genehmigung.

Der relativ hohe Aufwand für Akkreditierungen führt dazu, dass bis Mitte 2007 nur etwa ein Drittel der Bachelorstudiengänge und etwa 40 Prozent der Masterstudiengänge akkreditiert wurden (vgl. Tabelle 1), wobei sich zwischen neu eingeführten und reformierten Studiengängen statistisch nicht unterscheiden lässt. Dabei lassen sich große Unterschiede zwischen den Hochschularten, den Fächer(-gruppen) und den Ländern beobachten. An den Fachhochschulen, die mit der Umstellung auf gestufte Studiengänge etwas früher begonnen haben, wurde bereits ein größerer Teil der Studiengänge akkreditiert. Bei ihnen gab es auch gegenüber den Vorjahren eine deutliche Zunahme in der Akkreditierungsquote. An den Universitäten hingegen änderte sich nur wenig (vgl. Kerst/Heine/Wolter 2005, S. 46). Vergleichsweise selten sind bisher Studiengänge in den Sprach- und Kulturwissenschaften sowie im Bereich Kunst und Musik akkreditiert worden. In den Ingenieurwissenschaften liegt der Anteil akkreditierter Studiengänge sowohl bei den Bachelor- als auch den Masterprogrammen an der Spitze. Auch die Länder unterscheiden sich im Tempo der Akkreditierung beträchtlich.

Akkreditierungsvorhaben werden in der Regel erfolgreich zu Ende geführt. Bis Mitte 2006 wurden nur 30 Akkreditierungen abschlägig entschieden. Allerdings wurde etwa die Hälfte aller Akkreditierungen mit Auflagen vergeben (vgl. Akkreditierungsrat 2007a, b). Danach verlässt jeder zweite Studiengang den Akkreditierungsprozess offensichtlich in anderer Form, als er hineingegangen ist. Insofern scheint Akkreditierung die ihr zugeschriebene Erwartung, ein Instrument der Qualitätsentwicklung zu bilden, durchaus

zu erfüllen. Auf der anderen Seite scheint die Dynamik des Studienreformprozesses an den deutschen Hochschulen die Akkreditierungsinstanzen zu überfordern, da zur Zeit deutlich mehr nicht-akkreditierte als akkreditierte neue Studiengänge angeboten werden.

Tab 1: **Stand der Akkreditierung im Juli 2007<sup>1)</sup>**

	<b>Anzahl</b>	<b>Anteil akkreditierter Studiengänge (in %)</b>
<b>Grundständiges Studienangebot: Bachelorstudiengänge<sup>2)</sup></b>		
Insgesamt	4.038	31,3
an Fachhochschulen	1.677	46,9
an Universitäten <sup>3)</sup>	2.361	20,2
Nach Fachrichtungen <sup>4)</sup>		
Agrar-, Forst-, Haushalts- und Ernährungswissenschaften	93	31,2
Gesundheitswissenschaften, Medizin	213	30,5
Ingenieurwissenschaften	1.031	41,0
Kunst, Musik	176	13,1
Mathematik, Naturwissenschaften	825	35,4
Rechts-, Wirtschafts- und Sozialwissenschaften	955	37,3
Sprach- und Kulturwissenschaften	1.308	18,0
<b>Weiterführendes Studienangebot: Masterstudiengänge</b>		
Insgesamt	2.641	40,8
an Fachhochschulen	921	56,2
an Universitäten <sup>3)</sup>	1.720	32,5
Nach Fachrichtungen <sup>4)</sup>		
Agrar-, Forst-, Haushalts- und Ernährungswissenschaften	90	36,7
Gesundheitswissenschaften, Medizin	164	45,7
Ingenieurwissenschaften	672	47,0
Kunst, Musik	89	34,8
Mathematik, Naturwissenschaften	527	44,2
Rechts-, Wirtschafts- und Sozialwissenschaften	785	44,6
Sprach- und Kulturwissenschaften	636	29,1

1) Tag der Datenbankrecherche: 2. Juli 2007 entsprechend des Manuskripteingangs

2) Ohne Studiengänge, für die Studienanfänger nicht zugelassen sind

3) Einschließlich Kunst-/Musikhochschulen

4) Studiengänge können mehreren Fachrichtungen zugeordnet sein.

(Quelle: HRK-Studiengangsdatenbank)

#### 4. Diskurse und Perspektiven der Akkreditierung

Gerade wenn, wie in Deutschland, Evaluation und Akkreditierung erst auf eine relativ kurze Geschichte zurückblicken können und ihre Institutionalisierung noch fragil ist, kommt eine intensive Diskussion über Institutionen, Verfahren, Erfahrungen, Kosten und Nutzen nicht überraschend. Ergebnisse empirischer Forschung über Bedingungen und Folgen, Stärken und Schwächen des bislang praktizierten Akkreditierungsverfahrens sind mit Ausnahme einer international-vergleichenden Studie zur Implementation von Akkreditierungsverfahren in 20 Staaten (vgl. Schwarz/Westerheijden 2004b) kaum verfügbar. Zurzeit stehen vor allem zwei Fragen im Zentrum: die nach dem geeigneten Akkreditierungsverfahren (zugespitzt auf die Alternative zwischen Programm- und Systemakkreditierung) sowie die nach dem Verhältnis zwischen Evaluation und Akkreditierung. Beide Fragen hängen wiederum eng mit der gegenwärtig noch geringen Quote akkreditierter Studiengänge zusammen.

Da die Einrichtung neuer Bachelor- und Masterstudiengänge (bzw. die Umwandlung alter Studiengänge) noch voll im Gang ist und in den nächsten Jahren eine sehr hohe Zahl an Akkreditierungsanträgen ansteht, wird zunehmend der enorme zeitliche und personelle Aufwand, den Hochschulen und Akkreditierungsagenturen treiben (müssen), problematisiert. Zugleich ist damit zu rechnen, dass schon vor dem Ende des ersten Akkreditierungsdurchgangs die erste Welle der Re-Akkreditierung kommt. Bislang wurde, mit Ausnahme gebündelter Verfahren, jeder Studiengang einzeln begutachtet und akkreditiert (*Programmakkreditierung*). Dies bindet erhebliche Ressourcen bei den Akkreditierungsagenturen und den Gutachtergremien. Es wird damit gerechnet, dass nach erfolgter Umstellung auf das zweistufige Studiensystem (unter Einschluss der Studiengänge, die nicht umgestellt werden, wie z.B. in der Medizin und Rechtswissenschaft) mehr als 10.000 Studiengänge an den deutschen Hochschulen angeboten werden, die fast alle zu akkreditieren und zu re-akkreditieren wären.

Alternativ zur Programmakkreditierung werden von den Agenturen inzwischen auch gebündelte Akkreditierungsverfahren (*Clusterakkreditierung*) durchgeführt – mit dem Ziel, den Aufwand für die Agenturen und die Hochschulen zu senken. Dabei werden mehrere, fachlich zusammengehörige Studiengänge gemeinsam akkreditiert (vgl. Künzel 2006). Dieses Verfahren bietet insbesondere bei affinen Fächern oder Kombinationsstudiengängen (z.B. im Lehramt) eine Reihe von Vorteilen nicht nur im Blick auf die erforderlichen Ressourcen, sondern auch auf die Studiengangsstruktur; eine Einzelakkreditierung aller Kombinationen ist kaum praktikabel und sinnvoll (vgl. Künzel 2006, S. 19). Die Akkreditierung der „gebündelten“ Studiengänge erfolgt dann in einem einzigen Verfahren durch eine größere Gutachtergruppe.

Eine weitergehende Alternative zur Programm- und Clusterakkreditierung mit dem Ziel, ein vereinfachtes Verfahren mit reduziertem Aufwand durchzuführen, besteht in der *System- oder Prozessakkreditierung*, die inzwischen durch Beschlüsse der Kultusministerkonferenz und des Akkreditierungsrats eine festere Gestalt angenommen hat. Hier verschiebt sich der Brennpunkt von der Qualität des Studiengangs – einzeln oder als Cluster – auf die jeweilige Hochschule mit ihrem Qualitätssicherungs- bzw. -manage-

mentkonzept, auf den gesamten Prozess der Qualitätsentwicklung an einer Hochschule. Neben dem hohen Aufwand für die Hochschulen und Agenturen, insbesondere für die Rekrutierung von Gutachtern, sowie der bislang relativ geringen Akkreditierungsquote wird an der Programmakkreditierung vor allem kritisiert, dass sie als eine Art „top-down“-Verfahren kaum als Instrument hochschulinterner Qualitätssicherung funktioniert und daher nur einen geringen Beitrag zur institutionellen Qualitäts- und Hochschulentwicklung leistet. In diesem Zusammenhang ist die Idee entstanden, nicht eine Vielzahl einzelner Studiengänge, sondern hochschuleigene Prozesse und Verfahren der Qualitätssicherung, also ein hochschulinternes Qualitätsmanagementsystem zu akkreditieren, das wiederum gewährleisten soll, dass die einzelnen Studiengänge den Anforderungen an Vergleichbarkeit und Qualität der Abschlüsse entsprechen.

Dazu wurde in einem Pilotprojekt, das die Akkreditierungsagentur ACQUIN an vier Hochschulen unter Mitwirkung der Hochschulrektorenkonferenz durchgeführt hat, ein Verfahren der Prozessakkreditierung entwickelt und erprobt. Die Ergebnisse dieses Modellversuchs waren nach Einschätzung der Akteure positiv. Die teilnehmenden Hochschulen hätten „...ihre internen Organisations- und Entscheidungsstrukturen qualitätsorientiert optimiert (Qualitätsmanagement)“ (ACQUIN 2006, S. 6). Um Prozessakkreditierung flächendeckend zu realisieren, müssten allerdings alle Hochschulen solche funktionierenden Systeme der Qualitätssicherung aufbauen, die zudem regelmäßig intern und extern zu evaluieren und zu akkreditieren wären. Ein internationaler Vergleich zeigt, dass in den Hochschulsystemen anderer Staaten erst nach einer längeren Erfahrung mit Programmakkreditierung ein Übergang zu bzw. eine Ergänzung durch Prozessakkreditierung erfolgte (vgl. Banscheraus 2006, S. 24).

Im Anschluss an die Ergebnisse des Modellversuchs und nach einer entsprechenden Empfehlung des Akkreditierungsrats hat die KMK im Juni 2007 beschlossen, die bisherige Programmakkreditierung durch eine Systemakkreditierung zu ergänzen. Die KMK geht davon aus, dass auf absehbare Zeit beide Systeme nebeneinander bestehen werden. Sie hat den AR beauftragt, bis zum Herbst 2007 ein Umsetzungskonzept vorzulegen. Der AR ist dem nachgekommen und hat im Oktober 2007 die Einführung und Kriterien der Systemakkreditierung und die Zulassung der derzeit für Programmakkreditierung zuständigen und zertifizierten Agenturen auch zu diesem neuen Verfahren beschlossen (vgl. Akkreditierungsrat 2007a, b). Gegenstand einer Systemakkreditierung ist danach das interne Qualitätssicherungssystem einer Hochschule, in Ausnahmefällen auch von Teileinheiten (z.B. Fakultäten). Zu den Voraussetzungen für die Zulassung einer Hochschule zur Systemakkreditierung zählt u.a., dass sie in bestimmtem Umfang bereits über erfolgreich abgeschlossene Programmakkreditierungen verfügt. Der Beschluss formuliert mehrere Kriterien, welche an die Anerkennung eines internen Qualitätssicherungssystems zu stellen sind. Sie sollen sicherstellen, dass Qualitätssicherung innerhalb einer Hochschule als zentraler Teil der strategischen Hochschulentwicklung und in einer verbindlichen, nachprüfbaren und transparenten Weise unter Einschluss verschiedener evaluativer Komponenten verankert ist. In einem gesonderten Beschluss werden genaue Regeln für die Durchführung einer Systemakkreditierung aufgestellt. Im

Februar 2008 folgten Beschlüsse über die Kriterien und Verfahrensregeln der Systemakkreditierung.

Das neue Verfahren der System- bzw. Prozessakkreditierung wird zwar von vielen Akteuren, vor allem in den Hochschulen und auf Seite einiger Länder, begrüßt, aber von anderen Akteuren, insbesondere auf Seite der Studierenden und der im AR vertretenen gesellschaftlichen Organisationen, heftig kritisiert. Auf der einen Seite stärkt Prozessakkreditierung die institutionelle Verantwortlichkeit der Hochschule für Qualitätssicherung und -verbesserung als Teil ihrer Entwicklungsstrategie. Auf der anderen Seite verlieren die Akteure des Akkreditierungsprozesses an Einfluss auf die Gestaltung und Beurteilung von Studiengängen. Auch wird bemängelt, dass es sich bei Prozessakkreditierung weniger um ein Instrument inhaltlicher Studienreform handele, sondern eher um ein neues Konzept von Hochschulmanagement und Organisationsentwicklung. Weder Programm- noch Systemakkreditierung liefern, so wird kritisiert, Impulse für eine inhaltliche Studienreform (vgl. Banscherus/Staack 2007). Ein Qualitätssicherungssystem auf der Ebene einer Hochschule verhindert keineswegs, dass innerhalb der Hochschule ein Qualitätsgefälle zwischen den Studienangeboten und Studiengängen bestehen kann. Wie schon die Programmakkreditierung, so werde auch die Systemakkreditierung mit dem Problem zu kämpfen haben, dass sich die Bewertungsmaßstäbe der Agenturen erheblich voneinander unterscheiden und diese Diffusion in der Praxis zu mangelnder Vergleichbarkeit der Verfahren und Ergebnisse führt (vgl. ebd.). Müller-Böling sieht in beiden Varianten die Gefahr eines aufgeblähten, sich verselbständigenden Systems, in dem „fremde Agenturen“ den Hochschulen vorschreiben wollen, was „Qualität“ sei. Tatsächlich verfügten die Agenturen zurzeit gar nicht über gesicherte Kriterien, „(...) was denn ein gutes ‚Qualitätsmanagement-System‘ überhaupt auszeichnet“ (Müller-Böling 2007, S. 16).

Eine recht lebhaftete Debatte entzündete sich in Deutschland darüber hinaus an der möglichen Doppelgleisigkeit von Evaluation und Akkreditierung und der Entstehung einer eigenständigen Akkreditierungskultur neben dem hier immer noch zarten Pflänzchen einer Evaluationskultur, deren Verankerung sich von Hochschule zu Hochschule erheblich unterscheidet. Handelt es sich bei Evaluation und Akkreditierung eher um ein Verhältnis der Konkurrenz, der Kooperation bzw. Kombination oder der Kongruenz? Verfechter/-innen einer möglichst weitgehenden Annäherung stehen neben solchen einer möglichst strikten Separierung beider Verfahren (vgl. Hochschulrektorenkonferenz 2004, 2005). In der Tat gibt es eine Reihe substanzieller und methodischer Gemeinsamkeiten und Ähnlichkeiten zwischen beiden Verfahren. Bei beiden handelt es sich um Instrumente der externen und internen Qualitätssicherung. Beide sind mehrstufig nach dem Modell des „peer-review“-Verfahrens aufgebaut. Beide setzen eine Verständigung über die anzulegenden Qualitätskriterien und -standards voraus. Begreift man Evaluation nicht allgemein als Form der Bewertung, sondern als datenbasierten, ergebnis- und wirkungsbezogenen Beurteilungsprozess, dann kann zumindest eine Re-Akkreditierung nicht ohne Evaluation der Ergebnisse eines Studiengangs auskommen. Von daher wären Evaluation und Akkreditierung nicht nur wegen ihres Aufwandes, sondern auch wegen ihrer methodischen Kongruenz inhaltlich und zeitlich miteinander zu verknüpfen.

Auf der anderen Seite lassen sich Unterschiede in den Zielen und Funktionen, in den Adressaten/innen und Konsequenzen nicht übersehen (vgl. Hochschulrektorenkonferenz 2005, S. 61ff., 122ff., 159ff.). Bei Programmakkreditierung geht es im Kern um die Zulassung eines Studiengangs einschließlich der Anerkennung des damit verbundenen Titels. Dagegen steht bei einer Evaluation eher die erfahrungsgestützte Analyse der Stärken und Schwächen eines Studienprogramms als Teil eines kontinuierlichen Prozesses der Qualitätsverbesserung im Vordergrund. Akkreditierung orientiert sich an Mindeststandards, Evaluation an weitergehenden Qualitätsansprüchen. Während Akkreditierung stärker nach außen wirkt, wirkt Evaluation primär nach innen und soll in einer Fakultät bei den Programmverantwortlichen und Lehrenden einen selbstreflexiven Prozess der Qualitätsprüfung und -verbesserung initiieren. Akkreditierung ist auf den Erfolg in einer formalen Prozedur ausgelegt, Evaluation basiert geradezu auf einem selbstkritischen Effekt. Während (Erst-)Akkreditierung immer ein prognostisches Element über die Bewährung eines Studiengangs und die angestrebten „outcomes“ enthält, kann Evaluation dies genau im Nachhinein prüfen. Das Ergebnis einer Akkreditierung besteht in der autoritativen formalen Entscheidung über die Existenz eines Programms, das der Evaluation dagegen in differenzierten Empfehlungen zur qualitativen Verbesserung eines Lehr-/Lernprozesses und Programms. Die Unterschiede zwischen beiden Verfahren verschwimmen jedoch wieder ein wenig, sobald stärker vom Konzept der Prozessakkreditierung her gedacht wird.

## **5. Schlussbemerkungen**

In Deutschland wie in der Mehrzahl der anderen europäischen Staaten sind in den letzten Jahren Akkreditierungsverfahren eingeführt worden (vgl. Schwarz/Westerheijden 2004a, S. 17). In der Mehrzahl der Staaten spielt dabei der Bologna-Prozess eine zentrale Rolle. Allerdings lassen sich erhebliche Unterschiede in der nationalen Ausgestaltung beobachten, so schon bei der Frage, ob Programme, Institutionen oder Verfahren akkreditiert werden. In manchen Staaten werden nur bestimmte Studiengangstypen, in anderen dagegen alle Programme und Hochschulen akkreditiert. Auch der Typus der für Akkreditierung zuständigen Einrichtungen reicht von staatlichen Stellen über intermediäre Instanzen bis zu Fachverbänden oder Berufsorganisationen. Die Verfahren und Kriterien variieren ebenso wie das Ergebnis einer Akkreditierung und die jeweilige Verknüpfung zwischen Akkreditierung und Evaluation. Schließlich gibt es Unterschiede in den Funktionen, die Akkreditierung innerhalb eines nationalen Hochschulsystems erfüllen soll. In Deutschland stehen bei der Akkreditierung, ungeachtet weiterer Zielsetzungen, die Ziele „Deregulierung“, „Qualitätssicherung“ und „Profilbildung“ im Vordergrund. Das deutsche System der Akkreditierung ist gerade acht Jahre alt und steht schon vor einer Umstrukturierung. Noch gibt es kaum empirische Forschung zur Bewährung dieses neuen Verfahrens. Spätestens im Jahr 2010, dem angepeilten vorläufigen Abschluss des Bologna-Prozesses, wäre es an der Zeit, eine umfassende, auch empiriegestützte Bilanz zu ziehen.

Die genannten drei Ziele werden Prüfsteine für die Bewährung des spezifisch deutschen Akkreditierungssystems sein. Erstens gilt es genau zu beobachten, ob Akkreditierung tatsächlich dazu führt, bisher staatlich wahrgenommene Zuständigkeiten in intermediärer und einer sich selbst regulierenden Form so zu organisieren, dass die Hochschulen dies nicht als Fremdbestimmung, sondern als produktive fachbasierte Unterstützung und Anregung wahrnehmen, oder ob – wie manche befürchten – hier nur eine neue Qualitätsbürokratie mit einer subtileren Variante hierarchischer Steuerung entsteht. (Hier sei noch einmal Teichlers schon zitierte Warnung vor einem „superkomplexen System“ aufgegriffen). Das zukünftige Verhältnis zwischen staatlicher Zuständigkeit (Landesvorbehalte bei Studiengängen), intermediär organisierter Akkreditierung und institutioneller Studiengangsverantwortung ist nicht wirklich geklärt. „The history of accreditation (in Germany) is to be read as a constant fight for organisational independence against regulative attempts on behalf of the federal states” (Serrano-Velarde, zitiert nach Kehm 2007, S. 89).

Die zweite kritische Frage lautet, ob Akkreditierung tatsächlich zu einem Moment kontinuierlicher Qualitätsverbesserung wird, der das Qualitätsbewusstsein an deutschen Hochschulen stärkt und Impulse für die Studienreform liefert, oder ob sie nur zu einer mehr oder weniger duldsam ertragenen formalen Prozedur wird, der man nicht entkommen kann. Die Tatsache, dass das Reformtempo an den Hochschulen – historisch eher ungewöhnlich – gegenwärtig schneller ist als die nachlaufende Praxis der Akkreditierung, könnte darauf schließen lassen, dass die Akkreditierung eine zwar formal unumgängliche, aber keine substanzielle Funktion für Qualitätsentwicklung ausübt, was im übrigen durch den Wechsel zur Systemakkreditierung verstärkt würde. Drittens wird zu fragen sein, ob Akkreditierung zur individuellen Profilbildung von Hochschulen und Studiengängen und damit zu Vielfalt und Innovation beiträgt oder ob hier eher ein neues engmaschiges Netz aus mehr oder weniger vereinheitlichenden Anforderungen und Standards übergestülpt wird. Offen ist auch, wie sich die neuen Exzellenzbestrebungen auf die Akkreditierung auswirken: Wird es mit einer stärkeren vertikalen Differenzierung von Hochschulen und Studiengängen auch eine solche der Akkreditierung geben? Mit allen drei Aspekten ist letztlich die Frage verbunden, ob sich Akkreditierung zu einer lernenden Institution entwickelt, die sich selbst kritisch an diesen hochschulpolitischen Zielen evaluiert, oder ob sie diesen Ansprüchen selbst nicht gerecht wird.

## Literatur

- ACQUIN (2006): Reakkreditierungsantrag an den Akkreditierungsrat. Verfügbar unter: [http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/Beschluesse\\_AR/ACQUIN\\_Reakkreditierungsantrag.pdf](http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/Beschluesse_AR/ACQUIN_Reakkreditierungsantrag.pdf) (Letzter Zugriff 02.10.2008).
- Akkreditierungsrat (2007a): Allgemeine Regeln zur Durchführung von Verfahren zur Akkreditierung und Reakkreditierung von Studiengängen. Verfügbar unter: [http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/Beschluesse\\_AR/08.02.29\\_Regeln\\_Studiengaenge.pdf](http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/Beschluesse_AR/08.02.29_Regeln_Studiengaenge.pdf) (Letzter Zugriff 02.10.2008).

- Akkreditierungsrat (2007b): Kriterien für die Akkreditierung von Akkreditierungsagenturen. Verfügbar unter: [http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/Beschluesse\\_AR/07.10.08\\_Kriterien\\_fuer\\_Agenturen.pdf](http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/Beschluesse_AR/07.10.08_Kriterien_fuer_Agenturen.pdf) (Letzter Zugriff 02.10.2008).
- Banscherus, U. (2006): Entwicklungsmöglichkeiten des deutschen Akkreditierungssystems im internationalen Vergleich. Düsseldorf: Hans-Böckler-Stiftung (unveröffentlicht).
- Banscherus, U./Staack, S. (2007): Reformchance oder Abwicklung? Zur Zukunft des Akkreditierungssystems. In: Forum Wissenschaft 24, H. 1, S. 38–41.
- Bornmann, L./Mittag, S./Daniel, H.-D. (2006): Quality Assurance in Higher Education – Meta-Evaluation of Multi-Stage Evaluation Procedures in Germany. In: Higher Education 52, S. 687–709.
- Braun, D./Merrien, F. (1999) (Hrsg.): Towards a New Model of Governance for Universities? A Comparative View. London: Kingsley.
- Bretschneider, F./Wildt, J. (Hrsg.) (?2007): Handbuch Akkreditierung von Studiengängen. Bielefeld: Bertelsmann.
- Carstensen, D./Reissert, R. (1995): Interne und externe Evaluation – Modell und Praxis. HIS-Kurzinformation A 16/95. Hannover: HIS.
- Clark, B. (1983): The Higher Education System. Academic Organization in Cross-National Perspective. Berkeley: University of California Press.
- DAAD/HIS (2007): Wissenschaft weltweit. Daten und Fakten zur Internationalität von Studium und Forschung in Deutschland. Bielefeld: Bertelsmann.
- Daniel, H.-D./Mittag, S./Bornmann, L. (2003): Mehrstufige Evaluationsverfahren für Studium und Lehre. In: Berendt, B./ Voss, H.-P./Wildt, J. (Hrsg.): Neues Handbuch Hochschullehre. Stuttgart: Raabe, I 2.2, S. 1–26.
- Daxner, M. (1999): Evaluation, Indikatoren und Akkreditierung. Auf dem Wege in die Rechtfertigungsgesellschaft. In: Hochschulrektorenkonferenz (Hrsg.): „Viel Lärm um nichts?“ Evaluation von Studium und Lehre und ihre Folgen. Bonn: HRK, S. 41–50.
- Goll, D. (2006): Institutionelle Akkreditierung nicht-staatlicher Hochschulen durch den Wissenschaftsrat. In: Hochschulrektorenkonferenz (Hrsg.): Von der Qualitätssicherung der Lehre zur Qualitätsentwicklung als Prinzip der Hochschulsteuerung. Band II. Bonn: HRK, S. 53–63.
- Hartwig, L. (2003): Quality Assessment and Quality Assurance in Higher Education Institutions in Germany. In: Beiträge zur Hochschulforschung 25, H. 1, S. 64–83.
- Hochschulrektorenkonferenz (Hrsg.) (2004): Evaluation und Akkreditierung: bluffen – vereinheitlichen – profilieren? Bonn: HRK.
- Hochschulrektorenkonferenz (Hrsg.) (2005): Akkreditierung und Evaluation. Zwei Ziele, ein Verfahren? Bonn: HRK.
- Hochschulrektorenkonferenz (Hrsg.) (2006): Standards und Leitlinien für die Qualitätssicherung im Europäischen Hochschulraum. Bonn: HRK.
- Hochschulrektorenkonferenz (2007): Weitere Entwicklung der Systemakkreditierung. Empfehlung des 104. HRK-Senats am 12.06.2007.
- Hüfner, K. (1991): Accountability. In: Altbach, P. (Hrsg.): International Higher Education. An Encyclopedia. Chicago/London: St. James Press, S. 47–58.
- Jarausach, K. (Hrsg.) (1983): The Transformation of Higher Learning 1860-1930. Stuttgart: Klett.
- Kehm, B. (2007): Struktur und Problemfelder des Akkreditierungssystems in Deutschland. In: Beiträge zur Hochschulforschung 29, H. 2, S. 78–97.
- Kehm, B./Lanzendorf, U. (2005): Ein neues Governance-Regime für die Hochschulen – mehr Markt und weniger Selbststeuerung? In: Teichler, U./Tippelt, R. (Hrsg.): Hochschullandschaft im Wandel. 50. Beiheft der Zeitschrift für Pädagogik, S. 41–56.
- Kehm, B./Lanzendorf, U. (Hrsg.) (2007): Reforming University Governance. Bonn: Lemmens.
- Kerst, C./Heine, C./Wolter, A. (2005): Implementation der neuen Studiengänge - Stand und Auswirkungen auf Studierbereitschaft und Studierchancen. In: Gützkow, F./Quaißer, G. (Hrsg.): Jahrbuch Hochschule gestalten 2005. Denkanstöße zum Bologna-Prozess. Bielefeld: Weblar, S. 39–61.

- Kogan, M./Bauer, M./Bleikie, I./Henkel, M. (Hrsg.) (2006): *Transforming Higher Education. A Comparative Study*. Dordrecht: Springer.
- Konsortium Bildungsberichterstattung (2006): *Bildung in Deutschland*. Bielefeld: Bertelsmann.
- Künzel, R. (2006): Qualitätssicherung durch Akkreditierung. Vortrag auf der Jahrestagung des Verbands der Hochschullehrer für Betriebswirtschaft am 8. Juni 2006 in Dresden.
- Leszczensky, M. (2004): Paradigmenwechsel in der Hochschulfinanzierung. In: *Aus Politik und Zeitgeschichte*, H. 25, S. 18–25.
- Lüthje, J. (2004): Eröffnungsvortrag. In: Hochschulrektorenkonferenz (Hrsg.): *Evaluation und Akkreditierung: bluffen – vereinheitlichen – profilieren?* Bonn: HRK, S. 23–30.
- Müller-Böling, D. (2007): Ein Fest für Bürokraten. In: *Süddeutsche Zeitung*, Nr. 143, S. 16.
- Myers, R./Frankel, M./Reed, K./Waugaman, P. (1998): *Die Akkreditierung amerikanischer Hochschulen*. Bonn: BWFT.
- Nickel, S. (2006): Anforderungen an die Weiterentwicklung des Akkreditierungssystems aus der Sicht des CHE. Gütersloh: Centrum für Hochschulentwicklung.
- OECD (2003): *Education Policy Analysis*. Paris: OECD.
- Orr, D. (2007): More Competition in German Higher Education. In: Enders, J./Jongbloed, B. (Hrsg.): *Public-Private Dynamics in Higher Education*. Bielefeld: Transcript, S. 157–186.
- Orr, D./Jäger, M./Schwarzenberger, A. (2007): Performance Based Funding as an Instrument of Competition in German Higher Education. In: *Journal of Higher Education Policy and Management* 29, S. 3–23.
- Reissert, R. (1994): Evaluation der Lehre – interne Selbstevaluation und externe Begutachtung durch Peers. HIS-Kurzinformation A 8/94. Hannover: HIS.
- Schade, A. (2004): The Case of Germany. Country Report Germany. In: Schwarz, S./Westerheijden, D. (Hrsg.) (2004b): *Accreditation in the Framework of Evaluation Activities*. Frankfurt: GEW, S. 163–180.
- Schindler, G. (2004): Employability und Bachelor-Studiengänge – eine unpassende Verbindung. In: *Beiträge zur Hochschulforschung* 26, H. 4, S. 6–27.
- Schwarz, S./Westerheijden, D. (2004a): *Accreditation in the Framework of Evaluation Activities. Synopsis of the Current Situation and Dynamics in Europe*. In: Schwarz, S./Westerheijden, D. (Hrsg.): *Accreditation in the Framework of Evaluation Activities*. Frankfurt: GEW, S. 11–39.
- Schwarz, S./Westerheijden, D. (Hrsg.) (2004b): *Accreditation in the Framework of Evaluation Activities*. Frankfurt: GEW.
- Teichler, U. (1999): Profilierungspfade der Hochschulen im internationalen Vergleich. In: Olbertz, J.-H./Pasternack, P. (Hrsg.): *Profilbildung – Standards – Selbststeuerung*. Weinheim: Beltz, S. 27–38.
- Teichler, U. (2001): Profilbildung. In: Hanft, A. (Hrsg.): *Grundbegriffe des Hochschulmanagements*. Neuwied: Luchterhand, S. 369–373.
- Teichler, U. (2002a): Hochschulbildung. In: Tippelt, R. (Hrsg.): *Handbuch Bildungsforschung*. Opladen: Leske & Budrich, S. 349–370.
- Teichler, U. (2002b): Diversification of Higher Education and the Profile of the Individual Institution. In: *Higher Education Management* 14, H. 3, S. 177–188.
- Teichler, U. (2002c): Die Zukunft der Hochschulen in Deutschland. In: *Die Hochschule* 11, S. 29–45.
- Teichler, U. (2003): Die Entstehung eines superkomplexen Systems der Qualitätsbewertung in Deutschland. In: Mayer, E./Daniel, H.-D./Teichler, U. (Hrsg.): *Die neue Verantwortung der Hochschulen*. Bonn: Lemmens, S. 213–216.
- Teichler, U. (2005a): Hochschulsysteme und Hochschulpolitik. Quantitative und strukturelle Dynamiken, Differenzierungen und der Bologna-Prozess. Münster: Waxmann.
- Teichler, U. (2005b): Towards a „European Higher Education Area“ – Visions and Reality. In: *Higher Education Forum* 2, S. 35–54.
- Trow, M. (1973): *Problems in the Transition from Elite to Mass Higher Education*. Berkeley: Carnegie Commission on Higher Education.

- Trow, M. (1996): Trust, Markets and Accountability in Higher Education: A Comparative Perspective. In: Higher Education Policy 9, S. 309–324.
- Windolf, P. (1990): Die Expansion der Universitäten 1870-1985. Stuttgart: Enke.
- Wissenschaftsrat (2000): Empfehlungen zur Akkreditierung privater Hochschulen. Köln: Wissenschaftsrat.
- Wittrock, B. (1993): The Modern University – The Three Transformations. In: Rothblatt, S./Wittrock, B. (Hrsg.): The European and American University since 1800. Cambridge: University Press, S. 303–362.
- Wolter, A. (1999): Die Transformation der deutschen Universität – Historische Erbschaft und aktuelle Herausforderungen. In: Scholz, W.-D./Schwab, H. (Hrsg.): Bildung und Gesellschaft im Wandel. Oldenburg: Bibliotheks- u. Informationssystem der Carl von Ossietzky Univ. Oldenburg, S. 19–60.
- Wolter, A. (2004): From State Control to Competition. German Higher Education Transformed. In: The Canadian Journal of Higher Education XXXIV, H. 4, S. 73–104.
- Wolter, A. (2006): Auf dem Wege zu einem Europäischen Hochschulraum? Studienreform und Hochschulpolitik im Zeichen des Bologna-Prozesses. In: Hettlage, R./Müller, H.-P. (Hrsg.): Die europäische Gesellschaft. Konstanz: UVK, S. 299–324.
- Wolter, A. (2007): From the Academic Republic to the Managerial University – The Implementation of New Governance Structures in German Higher Education. In: University of Tsukuba, Research Center for University Studies (Hrsg.): Reforms of Higher Education in Six Countries – Commonalities and Differences. Tokyo: University Studies 35, S. 111–132.
- Ziegele, F. (2002): Reformansätze und Perspektiven der Hochschulsteuerung in Deutschland. In: Beiträge zur Hochschulforschung 24, H. 2, S. 106–121.
- Ziegele, F. (2006): Zielvereinbarungen als Kern des „Neuen Steuerungsmodells“. In: Hochschulrektorenkonferenz (Hrsg.): Von der Qualitätssicherung der Lehre zur Qualitätsentwicklung als Prinzip der Hochschulsteuerung. Band I. Bonn: HRK, S. 77–105.

*Anschrift der Autoren:*

Prof. Dr. Andrä Wolter, Technische Universität Dresden, Fakultät Erziehungswissenschaften,  
01062 Dresden, andrae.wolter@tu-dresden.de.

Dr. Christian Kerst, HIS Hochschul-Informationssystem GmbH, Gosseriede 9, 30159 Hannover,  
Kerst@his.de.

*Bernhard Schmidt*

# Qualität der Lehre an Hochschulen

## 1. Aktualität und Relevanz des Themas

Mit Einführung von Studienbeiträgen und einer durch den demografischen Wandel zu erwartenden Verknappung von Schulabgängern mit Hochschulreife ist eine zunehmende Konkurrenz der Hochschulen um qualifizierte Studienanfänger absehbar. Damit wird nicht nur das nach den Studentenprotesten zu Beginn der 1990er verstärkt in der hochschulpolitischen Diskussion präsen- te Thema der Qualität von Hochschullehre (vgl. Friedrich 2004) erneut relevant, sondern wird auch die Messung und Sichtbarmachung dieser Qualität in der Außendarstellung zu einem zentralen Anliegen der Hochschulen. Als ein weiterer Motor der Diskussion um die Qualität von Hochschullehre kann die Europäisierung der Hochschulbildung im Rahmen des sog. Bologna-Prozesses gesehen werden. So wird im Kommuniké der Bologna-Folgekonferenz in Berlin 2003 festge- stellt: „*The quality of higher education has proven to be at the heart of the setting up of a European Higher Education Area*“ (Berlin-Kommuniké 2003, S. 3). Die Voraussetzungen für eine qualitativ hochwertige Lehre sind – durch eine starke Expansion der Hoch- schulen von den 1960er bis in die 1980er Jahre ohne einen entsprechenden Ausbau der staatlich bereitgestellten Ressourcen für die Hochschulbildung – keineswegs optimal (vgl. Ebel-Gabriel 2004). Dennoch sehen sich die Hochschulen unter dem Druck, sich im europäischen Bildungsmarkt zu positionieren und den Anforderungen von Studie- renden, Arbeitsmarkt und Bildungspolitik gerecht zu werden. Diese verschiedenen Inte- ressengruppen prägen auch die Diskussion darüber, was im Kontext von Hochschulleh- re unter Qualität zu verstehen sei.

## 2. Was bedeutet Qualität im Kontext Hochschullehre?

Der Qualitätsbegriff hat vor allem im industriellen Sektor eine lange Tradition und wird in der ISO 9000 allgemein als „Grad, in dem ein Satz inhärenter Merkmale Anforderun- gen erfüllt“ (Zollondz 2002, S. 152) beschrieben. Die Übertragung auf Bildungskontexte scheint bei einer so unspezifischen Festlegung zunächst wenig problematisch, wirft aber in ihrer Konkretisierung Probleme auf, die in unterschiedlichen Bezugsgrößen und Maßstäben sichtbar werden. Qualität kann nach Harvey und Green (2000) erstens als Ausnahmeerscheinung verstanden werden, die aus tradiertem Elitarismus oder dem Übertreffen auch höchster Standards resultiert. Dieser Qualitätsbegriff spiegelt sich implizit auch in der gegenwärtigen Diskussion um Elite-Universitäten wider. Qualität

kann zweitens als Perfektion im Sinne von Fehlerlosigkeit verstanden werden, ein für den Bildungssektor utopisches Konzept. Drittens kann Zweckmäßigkeit im Sinne von Kundenzufriedenheit Basis eines Qualitätsdenkens sein, wobei für den Bildungsbereich festzulegen wäre, wer als Kunde anzusehen ist (Studierende, Arbeitsmarkt, Gesellschaft). Diese Problematik ergibt sich auch, wenn Qualität viertens als „adäquater Gegenwert“ (ebd., S. 28) definiert wird – eine Perspektive, die auch Kosten und Ressourcenaufwendungen eben dieser „Kunden“ berücksichtigt. Schließlich kann sich Qualität fünftens auch auf Transformationsprozesse beziehen, womit die mitgebrachten Voraussetzungen auf Teilnehmerseite (Vorwissen, Schlüsselkompetenzen, Motivation, etc.) in die Qualitätsbewertungen einfließen. Mit Ausnahme der Fehlerlosigkeit lassen sich diese Qualitätskonzepte auch in das Hochschulsystem transferieren und finden dort in verschiedenen Kriterien der Lehrqualität ihre Operationalisierung (vgl. Carstensen/Hofmann 2004), sei es in Form von Studierendenbefragungen zu deren Zufriedenheit, der Quantifizierung von Lehrleistungen z.B. in Absolventenzahlen, in Hochschulrankings oder dem Label der Elite-Universitäten. Wesentlich ist dabei festzuhalten, dass Qualität keine dem Gegenstand immanente Eigenschaft ist, sondern erst durch die Perspektiven der Betrachter und die herangezogenen Kriterien normativ konstruiert wird (vgl. Heid 2000). Da Bildung genuin mit der persönlichen Weiterentwicklung – ausgehend vom individuellen Wissens- und Bildungsstand des Einzelnen – verbunden ist, wird hier das letztgenannte, an Transformationsprozesse geknüpfte Qualitätskonzept als Bezugspunkt herangezogen. Für den Bereich Hochschule ergibt sich jenseits verschiedener Qualitätsbegriffe zusätzlich die Problematik, dass bereits für die Festlegung von zu erbringenden Leistungen und dazu notwendigen Ressourcen keine klaren Kriterien vorliegen (vgl. Pasternack 2001) und Forschung und Lehre als zwei zentrale Aufgabenfelder in der Bewertung von Hochschulen nicht getrennt voneinander betrachtet werden können. Qualität von Hochschullehre ist entsprechend schwer zu erheben (siehe 4.), wie sie aber beschaffen sein müsste, um den Anforderungen der wissenschaftlichen Lehr-Lernforschung gerecht zu werden, ist – zumindest aus hochschuldidaktischer Perspektive – sehr viel klarer zu beschreiben.

### 3. Was macht Qualität der Lehre an Hochschulen aus?

Qualität von Hochschullehre wird bestimmt durch vier Kategorien (vgl. Carstensen/Hofmann 2004): Studienprogramme, die den Rahmen für Lehr-Lernprozesse vorgeben; Diese Prozesse selbst, deren Ausgestaltung zu erheblichen Teilen den Lehrenden obliegt; Die Institution, die den Rahmen für Studium und Lehre schafft und gleichzeitig von der Lehre mitgeprägt wird; Das Handeln einzelner Akteure – Lehrender wie Studierender – und vor allem deren Interaktion. Darüber hinaus spielen aber auch (Studien-) Dekane und Prorektoren bzw. Vizepräsidenten, die für die Lehre auf übergeordneter Ebene verantwortlich sind, eine wesentliche Rolle. Künzel (2004) fordert für Personen in diesen Schlüsselpositionen eine „enge Kopplung von Entscheidungsbefugnis und persönlicher Verantwortung“ (ebd., S. 7) um die Effizienz der Zusammenarbeit der verschiede-

nen Akteure zu verbessern. Im Folgenden soll die Ebene der Interaktion von Studierenden und Lehrenden im Mittelpunkt stehen, da dieser Bereich in den letzten Jahren am deutlichsten im Fokus erziehungswissenschaftlicher Forschung und Auseinandersetzung stand und daher empirisch gut abgesichert ist. Qualität lässt sich auf dieser Ebene an der konkreten Gestaltung von Lernarrangements, an motivationalen Faktoren und an Kommunikationsstrukturen festmachen (vgl. Schmidt/Tippelt 2005a). In diesen Bereichen lässt sich seit Mitte der 1990er Jahre ein grundlegender Paradigmenwechsel feststellen, der meist unter dem Titel „the shift from teaching to learning“ firmiert und aus der US-amerikanischen Diskussion kommend in Deutschland vor allem durch die Arbeiten von Johannes Wildt (2004) bekannt wurde.

### 3.1 *Gestaltung von Lehr-Lernsituationen*

Die Qualität eines Angebots der Hochschullehre liegt zwar zu großen Teilen in der Verantwortung der jeweiligen Dozenten, wird aber im Vorfeld von institutionellen Rahmenbedingungen geprägt. Räumliche, zeitliche und materielle Vorgaben bilden die Grenzen des didaktischen Handlungsspielraums, in dem die Lehrveranstaltungen geplant werden können. Gute Lehre zeichnet sich hier durch die Berücksichtigung der Interessen und des Vorwissens der Lernenden bereits bei der Auswahl der Inhalte und Methoden aus (vgl. Mürmann/Schulte 2003a). Während der Durchführung lassen sich qualitativ hochwertige Lehrveranstaltungen gliedern in eine orientierende, motivierende und das Vorwissen der Teilnehmer aktivierende Einstiegsphase, eine durch das Initiieren, Unterstützen und Lenken von Lernprozessen geprägte Arbeitsphase und eine resümierende, weitere Lernprozesse anregende Abschlussphase. In allen Phasen ist auf eine abwechslungsreiche Gestaltung durch Nutzung eines breiten Spektrums von Methoden und Medien zu achten, wobei sich deren Auswahl wiederum an organisatorischen Rahmenbedingungen, Vorwissen der Teilnehmer und Lernzielen orientieren sollte. Die genaue Bestimmung affektiver, psychomotorischer und kognitiver Lernziele ist Voraussetzung für gezielte Wissensvermittlung und eine spätere Überprüfung der Lernzielerreichung (vgl. AHW 1998b).

Gehen die kognitiven Lernziele über die bloße Wissensaneignung hinaus und inkludieren ein tieferes Verständnis, die Analyse von Sachverhalten oder deren Bewertung, so sind insbesondere Lernformen geeignet, die eine aktive Auseinandersetzung mit Inhalten unterstützen (vgl. Schmidt/Tippelt 2007). Während dozentenorientierte Formen in der Hochschullehre, wie z.B. die Vorlesung, v.a. zur Vermittlung von Grundkenntnissen geeignet sind, fördern lernerorientierte Unterrichtsformen das aktive Erarbeiten der Inhalte. Das Lehr-Lerngeschehen richtet sich dabei v.a. an den Prozessen der Wissenskonstruktion auf Seiten der Lernende aus, während den Lehrenden eine moderierende, anleitende, strukturierende und begleitende Funktion zukommt. Dieser Perspektivenwechsel hin zu einer Orientierung an den Lernenden (vgl. Knoll 1998, S. 27) setzt den Einbezug des Vorwissens der Teilnehmer voraus, das sich durch assoziative Methoden wie z.B. Brainstorming aktivieren lässt (vgl. Mürmann/Schulte 2003b). Auf Seiten der

Seminarleitung setzen aktivierende Methoden, wie Lerntandems, Kleingruppenarbeit oder gelenkter Dialog, auch Sensitivität für Gruppenprozesse, Akzeptanz gegenüber Widerständen und Bereitschaft zu deren Artikulation voraus (vgl. Knoll 1998). Zudem gehört die Schaffung adäquater Rahmenbedingungen für den Lernprozess zu ihren Aufgaben, wobei die genannten institutionellen Rahmenbedingungen wirksam werden. Die beschriebenen Methoden eignen sich primär für den Einsatz und die Verbesserung der Lehrqualität in Seminar- und Kleingruppen, während aktivierendes Lehren und Lernen in universitären Massenveranstaltungen nur stark eingeschränkt möglich ist. Zwar wurden aus der universitären Praxis heraus Methoden für eine teilweise Zurücknahme der Lehrerdominanz in Vorlesungen entwickelt (vgl. Möhrle 1994), deren Möglichkeiten bleiben dennoch begrenzt und können den intensiven Lernprozess in überschaubaren Seminargruppen nicht ersetzen.

Aus einer konstruktivistischen Lernphilosophie heraus wurde die Bedeutung eines anwendungsbezogenen Wissenserwerbs betont, um Transferleistungen in Anwendungskontexte zu erhöhen. Verschiedene didaktische Ansätze – z.B. Konzepte situierten (vgl. Mandl/Gruber/Renkl 1997), forschenden (vgl. Huber 2001) oder problemorientierten Lernens (vgl. Schulmeister 2002) – tragen dieser Forderung Rechnung, indem sie Wissenskonstruktion und -anwendung miteinander verbinden. Die Grenzen zwischen den drei Ansätzen sind fließend und sie sind anschlussfähig an Modelle fallbasierten, explorativen Lernens, wie Projektseminare (vgl. Tippelt 1979) oder Planspiele (vgl. Hölscher 2000), wobei sich forschendes und z.T. auch projektorientiertes Lernen durch seine Zieloffenheit von den anderen Strategien abhebt (vgl. Schulmeister 2002) und sich primär zur Wissensvertiefung eignet (vgl. Huber 2001). Insbesondere für problemorientierte und fallbasierte Angebote im Kontext mediengestützter Hochschullehre liegen Evaluationsstudien vor, die auf eine Eignung dieser lernerzentrierten Konzepte verweisen (z.B. Kopp/Balk/Mandl 2002; Schmidt 2004; Tausch 2002). Forschendes Lernen verdeutlicht dagegen wie kaum ein anderes Konzept die produktive Verknüpfung von Forschung und Lehre, die zu den fundamentalen Charakteristika von Universitäten gehört, und belegt, dass beide Tätigkeitsfelder nicht als konkurrierend missverstanden werden dürfen, sondern sich gegenseitig synergetisch befruchten können (vgl. Huber 2001). Allerdings führt die nur langsam aufbrechende Dominanz der Forschungsleistung in der Leistungsbeurteilung von Wissenschaftlern zum Nachteil der Wertschätzung von Lehrleistungen, immer noch zu einem ungleich verteilten Engagement (vgl. auch Macke 1998).

Wie für alle Lehr-Lernformen gilt auch für die vorgestellten lernerzentrierten Konzepte der Hochschullehre, dass sich die Angebote nicht für alle Lernenden und in jedem Kontext gleichermaßen eignen und sowohl curriculare Vorgaben als auch räumliche, sachliche und personelle Ressourcen in die Entscheidung über die methodische Grundkonzeption einer Lehrveranstaltung einfließen müssen, um die Qualität von Hochschullehre nachhaltig zu sichern. Darüber hinaus lassen sich mit Prosser und Trigwell (2001) vier weitere zentrale Anforderungen an Lehrende ausmachen, die wesentlich zur Qualität von Lehre beitragen.

1. Moderation und Leitung von Lernsituationen erfordert Empathie und Reflexivität von den Studierenden, um die sich ständig verändernde Lernsituationen in ihrer Komplexität zu erfassen und adäquat agieren zu können.
2. Lehrende müssen die Wirkungen von eingesetzten Medien und Technologien wahrnehmen und abschätzen können.
3. Die Lerngewohnheiten und Aneignungsstrategien Lernender können sich sehr unterschiedlich darstellen, was Dozenten wahrnehmen und in die Seminargestaltung einfließen lassen sollten.
4. Die kontinuierliche Verbesserung der Lehre durch Evaluation ist eine auf verschiedenen Ebenen wesentliche Aufgabe von Hochschuldozenten (s.u.).

Die Berücksichtigung hochschuldidaktischer Erkenntnisse und der daraus ableitbaren Anforderungen ist wesentlicher Bestandteil der Qualitätsentwicklung in der Hochschullehre und begünstigt nicht nur einen höheren Lernerfolg und besseren Transfer des aufgebauten Wissens in praktische Anwendungsfelder (vgl. Kopp/Balk/Mandl 2002), sondern trägt auch wesentlich zur Lern- und Studienmotivation der Studierenden bei (vgl. Schmidt 2004), die ihrerseits auf den Verlauf von Lehr-Lernprozessen zurückwirkt.

### 3.2 *Motivation und Kommunikation*

Sowohl die Förderung der Lernmotivation Studierender als auch die angemessene Kommunikation zwischen Lehrenden und Lernenden sind als wesentliche Bestandteile der Qualität von Hochschullehre auf der Mikroebene anzusehen. Fehlende Motivation der Studierenden kann nicht nur Lehrveranstaltungen im Ablauf behindern, sondern ist auch eine wesentliche Ursache für Studienabbrüche (vgl. Jenkins/Breen/Lindsay/Brew 2003). Allgemein wird im Kontext von Motivationsförderung meist die intrinsische Motivation von Studierenden angestrebt, die v.a. eine vertiefte Auseinandersetzung mit Inhalten begünstigt, aber ungleich schwerer von Außen zu initiieren ist als extrinsische Motivation. Gerade hinsichtlich der schnellen Aneignung reproduzierbaren Wissens (z.B. bei der Prüfungsvorbereitung) ist für Studierende extrinsische Motivation oft funktionaler und effizienter, obwohl Prüfungen und Leistungsdruck – also klassische Quellen extrinsischer Motivation in der Hochschule – für ein vertieftes Lernen sonst eher abträglich sind (vgl. Ashcroft/Foreman-Peck 1994). Eine weitere Motivationsquelle für Studierende ist das Streben nach einem berufsqualifizierenden Abschluss als Voraussetzung für eine Integration auf dem Arbeitsmarkt sowie (damit einhergehend) ökonomische Vorteile und Statusgewinn. Diese Perspektiven fassen Newstead und Hoskins (2001) unter dem Begriff der „achievement motivation“ als dritte Form der Motivation, die vor allem das Interesse an praxisnahen Inhalten und Problemstellungen erhöht. Diese Interessen aufzugreifen und mit den curricularen Vorgaben in Einklang zu bringen sowie die praktische Relevanz der Inhalte zu verdeutlichen, können dann wirksame Methoden zur Förderung studentischer Lernmotivation sein. Darüber hinaus bieten regelmäßige, faire und valide Feedbacks, aktivierende Lernformen und der gezielte Abbau

von Lernbarrieren Möglichkeiten, die intrinsische Motivation zu erhöhen, wodurch nicht zuletzt die unmittelbare Interaktion zwischen Lehrenden und Lernenden relevant wird.

Die Kommunikation zwischen Lehrenden und Lernenden – soviel lässt sich bereits aus dem pragmatischen Konzept von Schulz von Thun (1992) ableiten – beschränkt sich nicht auf den rein sachlichen Austausch, sondern reflektiert – implizit oder explizit – immer auch die Beziehung zwischen beiden. Eine fehlende Sensibilität für diese Prozesse, die auch das Aushandeln von Handlungsspielräumen implizieren, kann zu Missverständnissen und Reaktanz führen. Die Definitionsmacht für die jeweilige Lehr-Lernsituation liegt schon aufgrund der institutionell verankerten Hierarchie auf Seiten der Lehrenden. Dennoch bedarf es einer Abstimmung zwischen Lehrenden und Lernenden und eines impliziten Konsens darüber, inwieweit diese hierarchischen Strukturen im Lernprozess reproduziert oder aufgebrochen werden (vgl. AHW 1998a). Implizite Aussagen über persönliche Wertschätzung verbergen sich auch in klassischen Unterrichtsfragen, die als „echte“ offene Fragen vor allem dann Interesse an der Perspektive der Lernenden signalisieren, wenn sie schon in ihrer Formulierung explizit auf deren Sichtweise abzielen. Rückmeldungen der Lehrenden können dagegen Wertschätzung direkt zum Ausdruck bringen, aber auch verletzend wirken, wenn sie z.B. Verallgemeinerungen oder Kritik an der Person enthalten (vgl. ebd.). Insbesondere die bereits beschriebenen Formen lernerzentrierten Unterrichts, die mit einer Rollenerweiterung der Lehrenden einhergehen, die nun nicht nur als Wissensübermittler sondern auch als Lernbegleiter agieren, gehen auch mit veränderten Kommunikationsstrukturen einher und verlangen von allen Beteiligten ein hohes Maß an kommunikativer Kompetenz sowie die Bereitschaft zur Übernahme veränderter Rollen im Lernprozess. Dabei darf die Verantwortung für den Lernprozess bei aller Selbststeuerung nicht gänzlich auf die Lernenden übertragen werden (vgl. Schumacher 2003). Nicht zuletzt deshalb eignen sich lernerzentrierte Formen der Hochschullehre keineswegs für alle Lernziele gleichermaßen.

Qualität im Kontext der Hochschullehre kann – wie hier gezeigt – zunächst auf Veranstaltungsebene hergestellt werden, spiegelt sich aber auch in angemessenen Angebotsstrukturen und adäquaten institutionellen Rahmenbedingungen wider. Entsprechend muss auch die Überprüfung der Lehrqualität auf unterschiedlichen Ebenen ansetzen. Letztlich kann diese Qualitätsmessung aber auch eine zentrale Grundlage für Verbesserungen in der Hochschullehre im Sinne eines Qualitätsentwicklungsprozesses sein, wenn sie nach den im Folgenden skizzierten Standards erfolgt.

#### **4. Wie wird Qualität der Hochschullehre überprüft?**

Die gängigen Instrumente zur Sicherung der Qualität in der Hochschullehre umfassen neben verschiedenen Formen der Evaluation auch die Akkreditierung als Instrument zur Sicherung eines gewissen Standards im Rahmen der Etablierung neuer Studiengänge. Akkreditierungsverfahren überprüfen nicht die Qualität eines laufenden Studien-

programms, sondern schätzen die Tauglichkeit eines neu einzurichtenden Studienprogramms hinsichtlich dessen zu erwartender Qualität, die auf Basis von Kennzahlen und Expertisen eingeschätzt wird. Das Thema der Akkreditierungsverfahren wird gerade im Zuge der Einführung neuer konsekutiver Studiengänge aktuell intensiv diskutiert (z.B. Winter 2006), soll aber hier nicht vertieft werden. Stattdessen steht im Folgenden die Frage nach der Messung von Qualität in bestehenden Studienprogrammen und Lehrveranstaltungen im Zentrum.

Wenn nach Verfahren zur Erhebung von Qualität gefragt wird, sind zunächst einige andere Punkte zu klären. Zum einen ist nach Kriterien guter Lehre zu fragen und nach Pasternack (2001) ist generell zweifelhaft, ob sich Normierungen und Standardisierungen, die üblicherweise die Basis von Qualitätsanalysen bilden, im Hochschulsystem überhaupt festlegen lassen. Wie im ersten Teil dieses Beitrags gezeigt wurde, gibt es aber zumindest aus hochschuldidaktischer Perspektive sehr wohl einige entsprechende Kriterien auf Ebene des Handelns der Akteure, die aber nicht immer einer direkten Messung zugänglich sind. Zum anderen stellt sich die Frage nach den Adressaten der Ergebnisse einer Qualitätsanalyse – seien es Abiturienten, Ministerien oder die Universitätsleitung – und den für diese Zielgruppe relevanten Daten (vgl. Müller-Böling/Berthold/Buch 2004). Die Untersuchung von Lehrqualität bedient sich i.d.R. der verschiedenen Formen und Methoden der Evaluation, worunter in Anlehnung an Wottawa und Thierau (1998, S. 14) eine zweck- und zielorientierte Maßnahme zur Bewertung von Handlungsalternativen verstanden werden kann, die dem aktuellen Stand wissenschaftlicher Methodik angepasst ist. Evaluationen können somit als Entscheidungshilfe herangezogen werden, der Legitimierung oder Optimierung von Studiengängen oder Lehrveranstaltungen dienen oder auch als Kontrollinstrument z.B. hinsichtlich erzielter Lernerfolge eingesetzt werden (vgl. Wesseler 1999, S. 739). Festlegungen, wie Zeitpunkt und Ziel der Evaluation (formative vs. summative Evaluation), für die Durchführung Verantwortliche (intern oder extern) und Bewertungskriterien (vgl. Kromrey 2001, S. 23), müssen sich unmittelbar an den mit der Evaluation verbundenen Intentionen ausrichten sowie an dem Evaluationsgegenstand. Im Fokus der Evaluation können dabei die Gestaltung der Lehrveranstaltungen (Prozess), die Prüfungsleistungen, Absolventen- und Abbrecherzahlen (Output) oder die Chancen auf dem Arbeitsmarkt (Outcome) stehen. Die verfügbaren Ressourcen und Curricula (Input) sind, ebenso wie außerinstitutionelle gesellschaftliche Rahmenbedingungen (Kontext), in jeder Evaluation als Rahmenbedingungen von Lehre zu berücksichtigen (vgl. Wesseler 1999, S. 739). Betrachtet man die Qualität von Hochschullehre als übergreifenden Evaluationsgegenstand, so lässt sich dieser in Anlehnung an die bereits vorher eingeführten vier Kategorien von Carstensen und Hofmann (2004) weiter spezifizieren:

- Die Evaluation von Lehr- Lernprozessen sowie des Handelns der Akteure erfolgt auf Veranstaltungs- bzw. Personenebene kann im Rahmen der Qualitätssicherung der Hochschullehre auch als **Mikroebene** bezeichnet werden.
- Auf der **Mesoebene** sind dann Evaluationen zu subsumieren, die institutionelle Ressourcen, sachliche und personale Ausstattung oder curriculare Fragen eines Studien-

gangs oder Fachbereichs fokussieren, also vorrangig die Institution und Studienprogramme im Blick haben.

- Auf **Makroebene** schließlich sind Übergangs- und Verbleibsstudien anzusiedeln, die die weitere Laufbahn von Studierenden nach Abschluss ihrer Ausbildung verfolgen, also den universitären Outcome (vgl. auch Schmidt/Tippelt 2005b).

Auf Mikroebene wird häufig einer summativen Evaluation der Vorzug gegeben und eine retrospektive Veranstaltungsbewertung angestrebt, die allerdings keine korrigierenden Interventionen mehr zulässt. Auf Mesoebene dagegen sind gerade diese Interventionsmöglichkeiten im Fokus der dann zumeist formativ durchgeführten Evaluation, da hier eine kontinuierliche Qualitätsentwicklung im Bezug auf einen Studiengang oder Fachbereich im Zentrum steht (vgl. Kromrey 2001).

#### 4.1 Mikroebene

Um Aussagen über die Qualität einzelner Lehrveranstaltungen an Hochschulen treffen zu können, wird gewöhnlich auf das subjektive Urteil von Studierenden zurückgegriffen. Der dabei implizit unterstellte Wirkungszusammenhang zwischen didaktischer Gestaltung – operationalisiert durch das Urteil der Lernenden – und Lernerfolg der Studierenden kann wegen seines einseitigen Blickwinkels auf die Lehrenden und der damit verbundene Vernachlässigung der lernerseitigen Voraussetzungen (Vorwissen, Motivation, Interesse, etc.) zu Recht als nur bedingt ausreichend kritisiert werden (vgl. Zimmer 2000). Insbesondere die Feststellung des Erfolgs von Lehre in Form von Wissenszuwachs (im Sinne eines auf Transformationsprozesse hin ausgerichteten Qualitätsbegriffs) kann durch dieses Vorgehen nicht zufrieden stellend erfasst werden. Dieses Ziel wäre wohl nur mit Leistungstests (vgl. Glowalla/Glowalla/Kohnert 2002) oder Verfahren zur Kompetenzdiagnostik (vgl. Braun/Soellner/Hannover 2006) und in einem Pre-Post-Test-Design zu erreichen, wobei allerdings nur diejenigen Lernerfolge erfasst werden, die per Lernzielformulierung vorab definiert wurden und sich kaum Rückschlüsse zur Verbesserung der Lehre ableiten lassen (vgl. Lohnert/Rolfes 1997). V.a. hinsichtlich einer Qualitätsverbesserung erscheint eine Evaluation der didaktischen Qualität von Hochschullehre folglich angezeigt, wobei die Aussagekraft von Studentenurteilen seit Jahren intensiv diskutiert wird (vgl. Schmidt/Tippelt 2005b). Als Fazit dieser Diskussion kann festgehalten werden, dass studentische Veranstaltungskritik ein wesentlicher Bestandteil von Lehrevaluationen sein sollte, aber – insbesondere in standardisierter Form – der Ergänzung durch weitere Daten bedarf (vgl. Kromrey 2005). Eine hohe Relevanz kommt der Einbindung Studierender in die Veranstaltungsevaluation nicht nur deswegen zu, weil sich deren Urteile als durchaus aussagekräftig im Hinblick auf die Lehrqualität erwiesen haben (vgl. Schmidt/Tippelt 2005b), sondern auch weil durch sie die Kommunikation zwischen Lehrenden und Lernenden über die Qualität von Hochschullehre hergestellt wird, weshalb studentische Lehrveranstaltungskritik nicht nur als das gebräuchlichste sondern nach Rindermann (2001, S. 76) auch „als das bisher geeigneteste Maß

zur Bestimmung von Akzeptanz und Qualität einer Lehrveranstaltung“ angesehen werden kann.

Die Erhebung studentischer Urteile kann dabei sowohl standardisiert als auch qualitativ mit Hilfe unterschiedlicher Methoden erfolgen. Während die standardisierte Veranstaltungskritik v.a. im Hinblick auf veranstaltungsübergreifende Vergleiche – die methodisch allerdings als äußerst problematischen anzusehen sind (vgl. Kromrey 2005) – Vorteile verspricht, betonen qualitative Verfahren, wie z.B. Gruppendiskussionen o.a., stärker den Dialog zwischen Lehrenden und Lernenden (vgl. Bülow-Schramm 1995). Eine Kombination beider Erhebungsverfahren verspricht besonders fundierte Erkenntnisse zur Qualitätsentwicklung, ist aber mit höherem Aufwand verbunden.

Diese Kombination verschiedener methodischer Ansätze hat sich auch hinsichtlich der Evaluation virtueller Lehr-Lernarrangements bewährt (vgl. Schmidt 2004). Dabei können traditionelle Erhebungsverfahren in das Medium übertragen werden (z.B. Online-Fragebögen oder Diskussionen in Internet Relay Chats oder Online-Foren) oder erweiterte, dem Medium immanente Möglichkeiten der Datengewinnung (z.B. Logfile-Analysen) genutzt werden. Auch inhaltlich sind die Kriterien qualitativ hochwertiger Lehre (s.o.) im Kontext virtueller Lehre noch um einige Aspekte zu ergänzen, die sich aus den didaktischen und technischen Anforderungen des Mediums ergeben und die in zahlreichen Kriterienkatalogen zu finden sind (vgl. Tergan 2001). Computer- und internetgestützte Lehr-Lernarrangements können so einerseits auf ihre Eignung für den Einsatz in der Hochschullehre hin überprüft werden, andererseits kann die Existenz virtueller Lernangebote per se bereits als ein Indikator für innovative Lehre in einem Studiengang oder einem Fachbereich gewertet werden – also auf der Mesoebene der Lehrevaluation.

## 4.2 *Mesoebene*

Mit der Novellierung des Hochschulrahmengesetzes 1998 wurde die Evaluation von Fachbereichen erstmals auch gesetzlich verankert, entsprechende Initiativen und Verbände der Universitäten entstanden allerdings schon vor dieser Gesetzesinitiative (vgl. HRK 2000). Dabei sollten in die Einschätzung der Lehrqualität eines Studiengangs oder Fachbereichs nicht nur das Lehrangebot einfließen, sondern auch Daten zum Output (Studiendauer, Prüfungsleistungen, Absolventenzahlen), zu Transparenz und Anerkennung der zu erreichenden Abschlüsse sowie zu Zulassungsbedingungen (vgl. Kohler 2004). Zu den zentralen Fragen hinsichtlich der Qualitätssicherungsstrategien auf Ebene eines Studiengangs oder Fachbereichs gehört die Entscheidung zwischen interner und externer Evaluation, die erheblichen Einfluss auf die Ergebnisse und deren Tragweite hat (vgl. Reissert/Carstensen 1998), wobei sich im Hochschulbereich eine Kombination aus interner und externer Evaluation durchgesetzt hat (vgl. Friedrich 2004). Die interne Evaluation mündet i.d.R. in einen Selbstbericht, in dem Stärken und Schwächen, Stand und Perspektiven eines Faches oder Studiengangs dargestellt werden. Zu den wesentlichsten Schwächen dieses Instruments gehören neben der zusätzlichen zeit-

lichen Belastung der an der Berichterstellung Beteiligten, fehlendes Interesse und Engagement der Stakeholder sowie deren fehlende Bereitschaft zu einer kritischen Selbstreflexion (vgl. Mittag 2004). Auch evtl. einfließende Daten aus Veranstaltungsevaluierungen können die genannten Probleme nicht gänzlich verhindern, aber eine wesentliche Bereicherung für die interne Evaluation darstellen (vgl. HRK 2000). Diese Lehrberichte können auch im Rahmen externer Evaluationen, die meist als Peer-Review-Verfahren durchgeführt werden, herangezogen werden (vgl. Speck-Hamdan/Tippelt 2001). Im Rahmen dieser externen Begutachtung werden Dokumentenanalysen mit Gesprächen vor Ort kombiniert, wobei eine Befragung der Studierenden eher die Ausnahme bleibt. Die Evaluatoren sind also wiederum auf Auskünfte und Datensammlungen der zu Evaluierenden angewiesen, wodurch die Probleme der internen Evaluation auch hier sich nicht gänzlich auflösen (vgl. Bargel/El Hage 2000).

Um diesen möglichen Schwachpunkten von Evaluationen auf Fachbereichsebene zu begegnen, lassen sich in Anlehnung an die Kriterien des Joint Committee on Standards for Educational Evaluation (vgl. Sanders 1999) allgemeine Standards formulieren, welchen eine (Lehr-)Evaluation zu genügen hat. Die Nützlichkeit einer Evaluation kann nur dann nachhaltig gesichert werden, wenn Betroffene und Beteiligte vorab informiert werden, glaubwürdige Evaluatoren mit der Durchführung betraut werden, alle relevanten Informationen einfließen und abschließend ein klarer und transparenter Ergebnisbericht vorgelegt wird. Diese Voraussetzungen klingen ebenso plausibel wie die Forderung nach einem praktikablen und kostenwirksamen methodischen Design zur Sicherung der Durchführbarkeit einer Evaluation, dennoch sind die genannten Punkte in der Praxis oft mit Schwierigkeiten verbunden und werden nicht immer vollständig erfüllt und die gewünschten Effekte der Evaluation (z.B. Verbesserung der Lehrqualität) bleiben aus. Insbesondere von den Evaluatoren ist für die Korrektheit der Evaluation auf Basis vertraglich abgesicherter Evaluationsbedingungen und durch Offenlegung aller Ergebnisse, auch unter Berücksichtigung von Ablauf, Kontext und Zielsetzungen der Evaluation Sorge zu tragen. In der Verantwortung der Evaluatoren liegt auch die Genauigkeit der Ergebnisse, die vor allem auf der Verlässlichkeit der Quellen, der Validität und Reliabilität der Daten sowie deren methodisch korrekter Auswertung beruht. Für die Evaluation einzelner Veranstaltungen sind diese Standards selten vollständig umzusetzen und es ist daher nach pragmatischen Lösungen zu suchen. Auf der Mesoebene der Evaluation von Lehre sind diese Vorgaben dagegen als zwingend anzusehen, um als Handlungs- und Entscheidungsbasis verwertbare Ergebnisse zu produzieren. Neben den Evaluationsstandards finden sich auch international ähnliche Verfahren der Qualitätsmessung an den Hochschulen, allerdings ergeben sich erhebliche Unterschiede in deren Umsetzung, die durchaus auch quer zu Nationen und Bildungssystemen liegen und v.a. durch die Kultur der jeweiligen Hochschule geprägt sind (vgl. Brennan/Shah 2002).

International üblich sind auch Hochschulrankings, die für eine ausgewählte Zielgruppe (Schüler, Abiturienten) Informationen über die Qualität der Lehre an einzelnen Hochschulen und in einzelnen Fächern bereitstellen wollen. Zwar ist die Kritik bzgl. einer mangelnden empirischen Basis heute nicht mehr haltbar (vgl. Hornborstel/Keiner

2002), die Ergebnisse sind dennoch sehr stark von der Gewichtung einzelner Kriterien abhängig und führen zu uneinheitlichen Urteilen. Als Instrumente der Qualitätsentwicklung sind die Rankings, die sich auf die grobe Klassifizierung eines Ist-Zustands beschränken, nicht geeignet, da sie kaum Hinweise auf Verbesserungspotenzial geben.

### 4.3 Makroebene

Der methodische Aufwand zur Gewinnung von Daten zum Outcome von Hochschullehre – also zum Verbleib von Absolventen und deren Karriereverläufen – ist ungleich größer als die Auswertung i.d.R. ohnehin vorhandenen statistischen Materials zu Studenten- und Absolventenzahlen, wissenschaftlichem Personal oder Lehrangeboten. So bleiben verlässliche Daten zu der Berufseinmündung und den beruflichen Perspektiven von Hochschulabsolventen rar und liegen meist nur regional begrenzt und für einzelne Disziplinen vor. Hinzu kommt, dass auch diese aufwändig erhobenen Daten einem schnellen Veralterungsprozess unterliegen, da Arbeitsmarktbedingungen und Berufsperspektiven einerseits oft an kurzfristige wirtschaftliche Entwicklungszyklen gekoppelt sind (vgl. Kellermann 2000), andererseits aber erst einige Jahre nach dem Hochschulabschluss zuverlässig erfasst werden können. Neben den Übergängen von der Hochschule in die Erwerbsarbeit, den Tätigkeitsfeldern, den Arbeitsbedingungen und dem Einkommen wird auf Makroebene auch nach der Passung der in der Hochschule erworbenen Kompetenzen und den Anforderungen des Arbeitsmarkts bzw. des jeweiligen Tätigkeitsfelds gefragt. Diese Perspektive muss aber auch kritisch hinterfragt werden, sofern eine universitäre Ausbildung nicht allein – wenn überhaupt – berufsqualifizierende Ziele verfolgt (vgl. Kellermann 2000).

Auch bieten Untersuchungen zum beruflichen Verbleib von Hochschulabsolventen gerade für sozialwissenschaftliche Fächer die Möglichkeit, Berufsfelder zu identifizieren, Fehlallokationen aufzudecken, regionale und standortspezifische Differenzen sichtbar zu machen und diese für die Weiterentwicklung des Studiengangs heranzuziehen (vgl. Rauschenbach/Züchner 2005). Darüber hinaus werden berufliche Perspektiven sichtbar, die wiederum die Motivation der Studierenden (achievement motivation) beflügeln können.

## 5. Perspektiven der Qualität von Lehre an Hochschulen

Zur Verbesserung der Qualität von Lehre an Hochschulen leisten entsprechende Fortbildungsangebote zumindest auf Mikroebene einen zentralen Beitrag und können auch als Säule der Qualitätsentwicklung verstanden werden.

*„Quality enhancement (QE) is about getting teachers to teach better, which is what staff development is all about“* (Biggs 2001, S. 227).

In Deutschland existieren inzwischen zahlreiche hochschuldidaktische Zentren, Initiativen und Programme, deren inhaltlicher Schwerpunkt primär auf Themen liegt, die

der konkreten Unterrichtsgestaltung dienen, weniger auf Beratung und Evaluation der Lehre (vgl. Brems/Gruber 2003). Theoretisch aufgebautes Wissen zur Gestaltung von Lehr-Lern-Situationen ist aber keineswegs ein Garant für dessen Umsetzung in der Lehre (vgl. Winteler 2003). Die kollegiale Beratung kann hier einen wesentlichen Beitrag zur Umsetzung hochschuldidaktischen Know-Hows leisten, ebenso wie eine Ausrichtung hochschuldidaktischer Fortbildungsangebote am Paradigma des handlungsorientierten Lernens. Voraussetzung für hochschuldidaktisch sinnvolle Lehre sind aber auch entsprechende institutionelle Rahmenbedingungen, wie überschaubare Veranstaltungsgrößen und Zeitreserven zur Vor- und Nachbereitung. Darüber hinaus sind hochschuldidaktische Qualifikationen auch im Rahmen von Berufungsverfahren stärker zu gewichten und es müssen für die Studierenden Beratungs- und Informationsangebote geschaffen werden (vgl. Künzel 2004).

Neue Chancen aber auch Probleme hinsichtlich der Qualität der Lehre ergeben sich aus der aktuellen Entwicklung des Hochschulstandorts Europa, die unter dem Stichwort Bologna-Prozess eine breite Einführung gestufter Studiengänge beinhaltet. An die neuen Bachelor- und Master-Studiengänge sind erweiterte Qualitätskriterien anzulegen, die neben den genannten Bereichen (Gestaltung der Lehre, Studienberatung, Hochschuldidaktische Weiterbildung) auch die Studienorganisation (Modularisierung, Kreditpunkte-System) sowie multimediale Angebote einbeziehen und die Kriterien Internationalisierung und Berufsqualifizierung stärker betonen (vgl. Welbers 2004). Herausforderungen ergeben sich insbesondere aus der Modularisierung, die mit einer klaren Definition zu erwerbender Kompetenzen verbunden ist. Zwar ist die Definition von Lernzielen Grundvoraussetzung für erfolgreiche Lehre und damit auch wünschenswert, die Formulierung von Kompetenzen bringt aber in der Praxis – durch die mangelnde Operationalisierbarkeit – erhebliche Schwierigkeiten mit sich (vgl. Braun/Soellner/Hannover 2006). Terhart (2005) plädiert aufgrund praktischer Erfahrungen für pragmatische Lösungen in Form einer „*Kompromisslinie zwischen den gegebenen Ausgangsbedingungen und dem vorgestellten Ziel*“ (ebd., S. 91) und skizziert damit auch den permanenten Konflikt von Qualität und Qualitätsentwicklung der Lehre an Hochschulen zwischen wissenschaftlich fundierten und klar formulierten Anforderungen einerseits und durch Unterausstattung und durch institutionelle Anforderungen eingeschränkter Handlungsspielräumen andererseits. In diesem Spektrum gilt es im Rahmen der gegebenen Möglichkeiten den Prozess der Qualitätsentwicklung auf Veranstaltungs- wie auf Fachbereichsebene weiter voranzutreiben.

## Literatur

- AHW 1998a=Arbeitsgruppe Hochschuldidaktische Weiterbildung (Hrsg.) (1998): Besser Lehren. Heft 4: Kommunikation in Seminaren. Weinheim: Deutscher Studien Verlag.
- AHW 1998b=Arbeitsgruppe Hochschuldidaktische Weiterbildung (Hrsg.) (1998): Besser Lehren. Heft 5: Lehrstrategien: Aspekte der methodischen Strukturierung von Seminaren. Weinheim: Deutscher Studien Verlag.

- Ashcroft, K./Foreman-Peck, L. (1994): *Managing Teaching and Learning in Further and Higher Education*. London: Falmer Press.
- Bargel, T./El Hage, N. (2000): Evaluation der Hochschullehre: Modelle, Probleme und Perspektiven. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.), S. 207–224.
- Benz, W. (Hrsg.) (2004): *Handbuch Qualität in Studium und Lehre*. Stuttgart: Raabe.
- Berlin-Kommuniqué (2003): Realising the European Higher Education Area. [http://www.bologna-bergen2005.no/Docs/00-Main\\_doc/030919Berlin\\_Communique.PDF](http://www.bologna-bergen2005.no/Docs/00-Main_doc/030919Berlin_Communique.PDF) [06.03.2007]
- Biggs, J. (2001): The reflective institution: Assuring and enhancing the quality of teaching and learning. In: *Higher Education* 41/2001, 221–238.
- Braun, E./Soellner, R./Hannover, B. (2006): Ergebnisorientierte Lehrveranstaltungsevaluation. In: Hochschulrektorenkonferenz (Hrsg.): *Qualitätsentwicklung an Hochschulen. Beiträge zur Hochschulpolitik* 8/2006. Bonn, S. 60–67.
- Brems, S./Gruber, H. (2003): *Aktuelle Entwicklungen der Hochschuldidaktik*. Wien: WUV.
- Brennan, J./Shah, T. (2000): *Managing Quality in Higher Education. An International Perspective on Institutional Assessment and Change*. Buckingham: OECD, SRHE & Open University Press.
- Bülow-Schramm, M. (1995): „Wer hat Angst vor den Evaluatoren?“ Der Umgang mit Akzeptanzproblemen von Evaluationsverfahren. In: Verbeek, D. (Hrsg.) (1995): *Handbuch Hochschullehre Highlights*. Band 1: Evaluation der Lehre. Stuttgart: Raabe, S. 2–19.
- Carstensen, D./Hofmann, S. (2004): Qualität in Lehre und Studium: Begriffe und Objekte. In: Benz, W. (Hrsg.), C 1.1.
- Clemens, W./Strübing, J. (Hrsg.) (2000): *Empirische Sozialforschung und gesellschaftliche Praxis*. Opladen: Leske+Budrich.
- Ebel-Gabriel, C. (2004): Qualitätssicherung an Hochschulen – Konzepte, Prozesse, Akteure. In: Benz, W. (Hrsg.), B 1.1.
- Friedrich, H. R. (2004): Nationale und internationale Grundlagen der Qualitätssicherung an Hochschulen. In: Benz, W. (Hrsg.), A 2.1.
- Glowalla, U./Glowalla, G./Kohnert, A. (2002): Qualitätsmanagement interaktiver Studienangebote. In: Issing, L.J./Stärk, G. (Hrsg.): *Studieren mit Multimedia und Internet. Ende der traditionellen Hochschule oder Innovationsschub?* Münster: Waxmann, S. 113–128.
- Gogolin, I./Krüger, H.-H./Lenzen, D./Rauschenbach, T. (Hrsg.) (2005): *Standards und Standardisierung in der Erziehungswissenschaft. Beiheft der Zeitschrift für Erziehungswissenschaft*. Wiesbaden: VS Verlag.
- Harvey, L./Green, D. (2000): Qualität definieren. Fünf unterschiedliche Ansätze. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.), S. 17–40.
- Heid, H. (2000): Qualität: Überlegungen zur Begründung einer pädagogischen Beurteilungskategorie. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.), S. 41–54.
- Helmke, A./Hornstein, W./Terhart, E. (Hrsg.) (2000): *Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule*. Zeitschrift für Pädagogik, 41. Beiheft. Weinheim: Beltz.
- Hölscher, L. (2000): Das Planspiel als Lehrmethode für Massenveranstaltungen. In: Albach, H./Mertens, P. (Hrsg.): *Hochschulorganisation und Hochschuldidaktik*. Wiesbaden: Gabler, S. 155–178.
- Hornborstel, S./Keiner, E. (2002): Evaluation der Erziehungswissenschaft. In: *Zeitschrift für Erziehungswissenschaft* 4/2, S. 634–654.
- HRK 2000=Hochschulrektorenkonferenz (Hrsg.) (2000): *Wegweiser 2000 durch die Qualitätssicherung in Lehre und Studium*. Bonn.
- Huber, L. (2001): Forschendes Lernen. In: Technische Universität Darmstadt (Hrsg.): *Aktivierende Lehr- und Lernformen*. Darmstadt: Technische Universität, S. 11–30.
- Jenkins, A./Breen, R./Lindsay, R./Brew, A. (2003): *Reshaping Teaching in Higher Education. Linking Teaching with Research*. London: Kogan Page Limited.
- Kellermann, P. (2000): Employability als ultima ratio? Bewertung von Kompetenzen und Arbeitsanforderungen aus Sicht von Graduierten. In: Clemens, W./Strübing, J. (Hrsg.), S. 155–172.

- Knoll, J. (1998): Hochschuldidaktik der Erwachsenenbildung. In: Knoll, J. (Hrsg.): Hochschuldidaktik der Erwachsenenbildung. Bad Heilbrunn: Julius Klinkhardt, S. 13–47.
- Kohler, J. (2004): Kriterien und Standards. In: Benz, W. (Hrsg.): Handbuch Qualität in Studium und Lehre: Evaluation nutzen, Akkreditierung sichern, Profil schärfen. Stuttgart: Raabe, D 1.1.
- Kopp, V./Balk, M./Mandl, H. (2002): Evaluation problemorientierten Lernens in der Münchner Medizinerbildung – Bewertung durch die Studierenden (1997 bis 2001). Forschungsbericht Nr. 148. München: Ludwig-Maximilians-Universität.
- Kromrey, H. (2001): Evaluation von Lehre und Studium – Anforderungen an Methodik und Design. In: Spiel, C. (Hrsg.): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck. Münster: Waxmann, S. 21–60.
- Kromrey, H. (2005): Studierendenbefragungen in Lehrveranstaltungen: Instrument der Evaluation oder „nur“ der Qualitätsentwicklung? In: Hochschulrektorenkonferenz (Hrsg.): Hochschule entwickeln, Qualität managen: Studierende als (Mittel)punkt. Beiträge zur Hochschulpolitik 10/2005. Bonn: HRK, S. 57–71.
- Künzel, R. (2004): Aktive Qualitätssicherung und -entwicklung in Lehre und Studium. In: Benz, W. (Hrsg.), E 2.2.
- Lohnert, B./Rolfes, M. (1997): Handbuch zur Evaluation von Lehre und Studium an Hochschulen. Ein praxisorientierter Leitfaden. Hannover: ZEvA.
- Macke, G. (1998): „Besser Lehren“ – Grundlagen und Konzeption. In: Arbeitsgruppe Hochschuldidaktische Weiterbildung (Hrsg.): Besser Lehren. Heft 1: Grundlagen und Konzeption. Weinheim: Deutscher Studien Verlag, S. 22–50.
- Mandl, H./Gruber, H./Renkl, A. (1997): Situiertes Lernen in multimedialen Lernumgebungen. In: Issing, J.L./Klimsa, P. (Hrsg.): Information und Lernen mit Multimedia. Weinheim: Beltz, S. 167–178.
- Mittag, S. (2004): Empfehlungen für die Ausgestaltung der internen Evaluation. In: Hochschulrektorenkonferenz (Hrsg.): Metaevaluation. Evaluation von Studium und Lehre auf dem Prüfstand. Beiträge zur Hochschulpolitik 5/2004. Bonn: HRK, S. 29–42.
- Möhrle, M.G. (1994): Qualitätsverbesserung interaktiver Lehre durch das Lead-Learner-Konzept. In: Albach, H./Mertens, P. (Hrsg.): Hochschuldidaktik und Hochschulökonomie. Neue Konzepte und Erfahrungen. Wiesbaden: Gabler, S. 41–52.
- Müller-Böling, D./Berthold, C./Buch, F. (2004): Qualitätsermittlung, Qualitätssicherung und Qualitätserhöhung. In: Benz (Hrsg.), C 1.2.
- Mürmann, M./Schulte, D. (2003a): Lehrveranstaltungen lehrgerecht planen. In: Welbers (Hrsg.), S. 133–140.
- Mürmann, M./Schulte, D. (2003b): Aktivierende Methoden sinnvoll einplanen. In: Welbers, (Hrsg.), S. 237–244.
- Newstead, S.E./Hoskins, S. (2001): Encouraging Student Motivation. In: Fry, H./Ketteridge, S./Marshall, S. (Eds.): A Handbook for Teaching & Learning in Higher Education. London: Kogan Page Limited, pp. 70–82.
- Pasternack, P. (2001): Hochschulqualität: Ein unauflösbares Problem und seine Auflösung. In: Olbertz, J.-H./Otto, H.-U. (Hrsg.): Qualität von Bildung. Vier Perspektiven. Wittenberg: HoF Institut für Hochschulforschung.
- Prosser, M./Trigwell, K. (2001): Understanding Learning and Teaching. The Experience in Higher Education. Buckingham: SRHE & Open University Press.
- Rauschenbach, T./Züchner, I. (2005): Ausbildungsstandards im Lichte des Arbeitsmarkts – Anforderungen an die Reform erziehungswissenschaftlicher Studiengänge. In: Gogolin, I./Krüger, H.-H./Lenzen, D./Rauschenbach, T. (Hrsg.), S. 153–174.
- Reissert, R./Carstensen, D. (1998): Praxis der internen und externen Evaluation. Handbuch zum Verfahren. Hannover: HIS.
- Rindermann, H. (2001): Die studentische Beurteilung von Lehrveranstaltungen – Forschungsstand und Implikationen. In: Spiel, C. (Hrsg.) (2001): Evaluation universitärer Lehre – zwischen Qualitätsmanagement und Selbstzweck. Münster: Waxmann, S. 61–88.

- Sanders, J.R. (Hrsg.) (1999): Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“. Opladen: Leske+Budrich.
- Schmidt, B. (2004): Virtuelle Lernarrangements für Studienanfänger. Didaktische Gestaltung und Evaluation des Online-Lehrbuchs Jugendforschung und der begleitenden virtuellen Seminare. München: Utz.
- Schmidt, B./Tippelt, R. (2005a): Besser lehren – Neues von der Hochschuldidaktik. In: Teichler, U./Tippelt, R. (Hrsg.), S. 103–114.
- Schmidt, B./Tippelt, R. (2005b): Lehrevaluation. In: Gogolin, I./Krüger, H.-H./Lenzen, D./Rauschenbach, T. (Hrsg.), S. 227–242.
- Schmidt, B./Tippelt, R. (2007): Aktivierendes Lehren – Problemorientiertes Lernen: Herausforderungen für Hochschule und Beratung. In: Zeitschrift für Beratung und Studium 2007/1, S. 13–18.
- Schulmeister, R. (2002): Zur Komplexität Problemorientierten Lernens. In: Asdonk, J./Kroeger, H./Strobl, G./Tillmann, K.-J./Wildt, J. (Hrsg.): Bildung im Medium der Wissenschaft. Weinheim: Deutscher Studien Verlag, S. 185–202.
- Schumacher, E.-M. (2003): Der Lehrende als Coach. Schlüsselqualifikationen für eine neue Lernkultur. In: Welbers, U. (Hrsg.), S. 221–228.
- Schulz von Thun, F. (1992): Miteinander reden. Band 1: Störungen und Klärungen. Reinbek: Rowohlt.
- Speck-Hamdan, A./Tippelt, R. (2001): Evaluation im Prozess – Zeitliche, sachliche und soziale Dimensionen einer Evaluation in der Erziehungswissenschaft. In: Keiner, E. (Hrsg.): Evaluation (in) der Erziehungswissenschaft. Weinheim: Beltz, S. 155–162.
- Tausch, J. (2002): Problemorientierte Verfahren – Was können sie in Lehre und Unterricht leisten? In: Neumann, K./Osterloh, J. (Hrsg.): Gute Lehre in der Vielfalt der Disziplinen. Weinheim: Beltz, S. 97–116.
- Teichler, U./Tippelt, R. (Hrsg.) (2005): Hochschullandschaft im Wandel. 50. Beiheft der Zeitschrift für Pädagogik. Weinheim: Beltz.
- Tergan, S.-O. (2001): Qualitätsbeurteilung von Bildungssoftware mittels Kriterienkatalogen. Problemaufriss und Perspektiven. In: Unterrichtswissenschaft 2001/4, S. 319–341.
- Terhart, E. (2005): Die Lehre in den Zeiten der Modularisierung. In: Teichler, U./Tippelt, R. (Hrsg.), S. 87–102.
- Tippelt, R. (1979): Projektstudium. Exemplarisches und handlungsorientiertes Lernen an der Hochschule. München: Kösel.
- Welbers, U. (Hrsg.) (2003): Hochschuldidaktische Aus- und Weiterbildung. Grundlagen – Handlungsformen – Kooperationen. Bielefeld: Bertelsmann.
- Welbers, U. (2004): Der Beitrag der Hochschuldidaktik zur Verbesserung und nachhaltigen Entwicklung von Qualität in Studium und Lehre. In: Benz, W. (Hrsg.), E 5.5.
- Wesseler, M. (1999): Evaluation und Evaluationsforschung. In: Tippelt, R. (Hrsg.) (1999): Handbuch Erwachsenenbildung/Weiterbildung. Opladen: Leske+Budrich, S. 736–752.
- Wildt, J. (2004): The Shift from Teaching to Learning – Thesen zum Wandel der Lernkultur. In: Ehlert, H./Welbers, U. (Hrsg.): Qualitätssicherung und Studienreform. Düsseldorf: Grupello, S. 168–178.
- Winteler, A. (2003): Ein Programm zur Entwicklung und Veränderung von Konzeptionen des Lehrens und Lernens. In: Welbers, U. (Hrsg.), S. 141–150.
- Winter, M. (2006): Studienstrukturreform, Akkreditierung und Gleichstellungspolitik. <http://evanet.his.de/evanet/positionen/positionen2006/winter.pdf> [08.03.2007]
- Wottawa, H./Thierau, H. (1998): Evaluation. Bern: Huber.
- Zimmer, G. (2000): Evaluation der Lehre durch logische Rekonstruktion der Lernhandlungen. In: Clemens, W./Strübing, J. (Hrsg.), S. 113–134.
- Zollondz, H.-D. (2002): Grundlagen Qualitätsmanagement. Einführung in Geschichte, Begriffe, Systeme und Konzepte. München: Oldenbourg.

*Anschrift des Autors:*

Dr. Bernhard Schmidt, Ludwig-Maximilians-Universität München, Institut für Pädagogik, Lehrstuhl für Allgemeine Pädagogik und Bildungsforschung, Leopoldstrasse 13, 80802 München.

Rudolf Tippelt

## Qualitätsstandards und Forschungsevaluationen in der Erziehungswissenschaft

Forschungsevaluationen gibt es in verschiedenen Bereichen der Erziehungswissenschaft und es ist unbestritten, dass die Erziehungswissenschaft – wie andere akademische Fächer auch – permanent ihre Qualität und ihre spezifischen Forschungspotentiale legitimieren muss (vgl. Hornborstel/Keiner 2002). Allerdings muss schon einleitend darauf hingewiesen werden, dass sich Evaluationen von Disziplinen und fachlichen Bereichen an Hochschulen auf keinen Fall nur auf die Forschungsleistung beziehen können, denn gleichgewichtig sind die Lehrleistungen nach quantitativen und qualitativen Kriterien ebenfalls zu berücksichtigen (vgl. zu wichtigen Indikatoren in diesem Zusammenhang den Datenreport Erziehungswissenschaft der DGfE; z.B. Tillmann u.a. 2008; Tippelt 2004). Die folgende Darstellung konzentriert sich auf drei ausgewählte Bereiche:

- 1) Struktur- und Forschungsevaluationen des Faches an Hochschulstandorten analysieren die organisatorischen Rahmenbedingungen, die Weiterentwicklung erziehungswissenschaftlicher Studiengänge und das wissenschaftliche Profil der Erziehungswissenschaft an einem Hochschulstandort (vgl. EWFT 2003a; Tippelt 2004, S.89ff), wobei zu konstatieren ist, dass bisherige Evaluationen der Erziehungswissenschaft auf äußerst verschiedene Bewertungskriterien zurückgegriffen haben (vgl. Böllert/Tippelt 2004). Auf der Basis bisheriger erziehungswissenschaftlicher Evaluierungen und durch das Reflektieren und Auswerten der dabei kumulierten Erfahrungen hat sich langsam, wenn nicht ein Konsens so doch eine Annäherung über verallgemeinerbare Bewertungskriterien in der scientific community herausgebildet.
- 2) Aus Forschungsevaluationen der außeruniversitären pädagogischen und erziehungswissenschaftlichen Institute – wie insbesondere der Einrichtungen der Wissenschaftsgemeinschaft Leibniz – haben sich im Anschluss an die jahrzehntelange Evaluationspraxis des Wissenschaftsrats begründete Bewertungskriterien und Standards entwickelt, die sowohl auf naturwissenschaftliche, wirtschaftswissenschaftliche, sozial- und erziehungswissenschaftliche Forschungseinrichtungen in den letzten Jahren angewandt wurden. Im erziehungswissenschaftlichen Feld wurden beispielsweise das Institut der Pädagogik der Naturwissenschaften (IPN, Kiel), das Deutsche Institut für internationale pädagogische Forschung (DIPF, Frankfurt), das Deutsche Institut für Erwachsenenbildung (DIE, Bonn), das Institut für Wissensmedien (IWM, Tübingen), aber auch das Zentrum für Psychologische Information und Dokumentation (ZPID, Trier) und GESIS (mit den Teileinrichtungen ZUMA, Mannheim, dem Zentralarchiv, Köln und dem Zentralinstitut für Information, Bonn) nach vergleich-

baren Qualitätskriterien beurteilt (vgl. WGL Leibniz-Gemeinschaft 2003). Auch andere Forschungseinrichtungen wie z.B. das Deutsche Jugendinstitut (DJI, München) oder das Institut für Qualitätssicherung und Bildungsstandards (IQB, Berlin) werden nach zwar spezifizierten aber doch vergleichbaren Qualitätskriterien evaluiert und eingeschätzt.

- 3) Im internationalen Bereich sind die von OECD-CERI durchgeführten „Reviews of Educational Research and Development“ in verschiedenen Ländern, wie Neuseeland, England, Dänemark, Mexiko, Schweiz hervorzuheben. Diese Forschungsevaluationen werden nur erstellt, wenn das jeweilige Land einen entsprechenden Auftrag formuliert. Die bisherigen Reviews wurden im OECD-Report „New Challenges for Educational Research“ (2003) sowie als separate Länderberichte ([www.oecd.org/edu/rd](http://www.oecd.org/edu/rd).) publiziert.

## **1. Allgemeine Standards und Anforderungen**

Zunächst muss gefragt werden, warum Forschungsevaluationen überhaupt notwendig sind?

Hierzu ist festzuhalten, dass Forschungsevaluationen eine permanente Qualitätssicherung und darüber hinaus auch die autonome Selbststeuerung von Arbeitseinheiten ermöglichen (vgl. Hanft 2003; Ditton 2002). Ohne eine solche Qualitätssicherung wäre die forschungsbezogene Selbststeuerung von Arbeitseinheiten gegenüber Geldgebern und fördernden öffentlichen Institutionen kaum legitimierbar. Forschungsevaluationen sind aber auch Mittel im Wettbewerb um knappe Ressourcen, und sie sind Steuerungsinstrumente, um Excellence und innovative Leistungen zu unterstützen. Es kann also nicht übersehen werden, dass die ökonomische Rationalität, d.h. eine rationale Vergabepraxis von meist öffentlichen Geldern, eine wichtige Ausgangsbedingung für Forschungsevaluationen darstellt. Die Budgetierung von Forschungseinrichtungen garantiert in kontrollierten, überprüfbaren Grenzen den freien Einsatz von knappen finanziellen Mitteln; gleichzeitig lassen sich Forschungssubventionen, Mittel für Projekte aus öffentlicher Hand nur dann legitimieren, wenn sie nach systematischen und klaren Qualitätskriterien vergeben werden (vgl. Schenker-Wicki 1996; McKinney 1999). Forschungsevaluationen dienen also der Steuerung von relativ autonomen gesellschaftlichen Funktionsbereichen, die sich kompetitiv bewähren müssen – dies können Forschungs- und Arbeitsbereiche an Universitäten und Hochschulen, aber auch externe Forschungsinstitutionen sowie Forschergruppen oder Graduiertenschulen beispielsweise im Kontext der Mittelvergabe der Deutschen Forschungsgemeinschaft oder eines nationalen Fonds sein. Dabei sind unbedingt einige allgemeine Anforderungen einzuhalten:

Evaluationskommissionen müssen in ihren Entscheidungen unabhängig sein, d.h. die Experten, welche die externe Evaluation durchführen, dürfen nicht in einer zu großen Nähe zu den zu evaluierenden Arbeitseinheiten stehen. Evaluationen sind u.a. dem Gebot der Objektivität verpflichtet (vgl. Sanders 1999).

Forschungsevaluationen müssen aktuell sein und daher immer auch das letzte Berichtsjahr beinhalten. Sie sollten darüber hinaus beispielsweise durch die Erfassung der Entwicklungen der letzten fünf Jahre auch eine Trendbeschreibung implizieren, so dass die Kontinuität und die Dynamik der Forschungsleistungen besser eingeschätzt werden können. Evaluationen, die diese Norm der Aktualität verletzen sind nur sehr bedingt aussagekräftig, weil sich in Forschungseinheiten bereits innerhalb von zwei bis drei Jahren enorme inhaltliche und personelle Veränderungen ergeben können.

Wichtig ist selbstverständlich, dass die Validität der Informationen gewährleistet ist. Insbesondere die Informationen in Selbstreports über Drittmittel, über Geldgeber, über inhaltliche Forschungsprojekte, über das Ausmaß der Nachwuchsförderung, über Publikationsintensität etc. müssen valide sein, denn Forschungsunterlagen, die lückenhaft sind oder gar falsche Daten produzieren, sind für jegliche Forschungsevaluation irreführend und daher unbrauchbar.

Eine Herausforderung von Forschungsevaluationen ist es, die Differenzierung der Forschungsleistungen gerade in größeren Einrichtungen zu berücksichtigen, denn man muss den besonderen Departments oder Abteilungen und ihren jeweiligen Strukturen gerecht werden. Da es in Einzeleinrichtungen immer stärkere und schwächere Arbeitseinheiten und Abteilungen gibt, sind pauschale Evaluationsberichte zur Forschung von Gesamteinrichtungen ohne differenzierende Beschreibungen und Bewertungen wertlos. Keinesfalls müssen schwächer bewertete Abteilungen weniger Förderung erhalten, entsprechende Diagnosen können auch Anlass dafür sein, spezifische Empfehlungen zu formulieren, um bislang schwächere Einheiten zu stärken. Andererseits können entsprechende Informationen auch zur leistungsadäquaten Förderung verwendet werden. Die jeweiligen internen Empfehlungen folgen keinem Automatismus, sondern kompensierende oder belohnende Maßnahmen können nur durch die übergeordneten Entwicklungsziele und -strategien von Forschungseinrichtungen begründet werden.

Zur Fairness von Forschungsevaluationen, aber auch zur Objektivität gehört es, dass neben der intensiven Befragung der Institutsleitung auch wissenschaftliche Mitarbeiter, Beiräte, fördernde Einrichtungen oder auch unmittelbar kooperierende Partner aus Forschung und Praxis gehört werden und bei der Bewertung und der Ableitung von Empfehlungen mit einbezogen werden.

Die Evaluationspraxis verschiedener Forschungsgemeinschaften legt nahe, dass sich die Evaluierungsgruppen aus unabhängigen Gutachtern zusammensetzen, die jeweils die umfangreichen Unterlagen der zu beurteilenden Institutionen heranziehen. Diese Materialien werden meist nach einem systematischen Fragenkatalog von den zu begutachtenden Instituten bearbeitet, so dass eine interne „Selbstevaluation“, der von der unabhängigen Evaluationskommission durchgeführten Fremdevaluation vorausgeht. Der interne Selbstreport ist wesentliche Ausgangsbasis der Analyse und Bewertung der „peers“ (vgl. zur Fremdevaluierung Reissert/Carstensen 1998). Im Selbstreport (vgl. Daniel 1988; Wesseler 1999) werden Aspekte der wissenschaftlichen Qualität, der Produktivität, des Drittmittelvolumens, Formen der Nachwuchsförderung, des praktischen Transfers und der praktischen Relevanz sowie die forschungsstrategische Entwicklungsplanung auf der Basis überprüfbarer Daten gesichtet.

Außerdem werden ein- bis zweitägige Vorortbesuche der Einrichtungen organisiert, bei denen nach festen Regeln Gespräche mit Institutsleitungen, Nachwuchswissenschaftlern und Kooperationspartnern, vor allem auch mit den jeweils benachbarten Hochschulen geführt werden (ähnlich dem Evaluationsverfahren des Wissenschaftsrats, der Universitäten oder der Leibniz-Gemeinschaft, vgl. Speck-Hamdan/Tippelt 2001, S. 156f.).

Die bisherigen Erfahrungen zeigen, dass Freiräume für die Gespräche im Evaluierungsverfahren aber auch standardisierte Regeln ein solides wie auch faires Urteil über die Forschungsleistung von Forschungseinrichtungen ermöglichen. Forschungsevaluierungen sind einerseits deskriptiv, andererseits beinhalten sie immer auch substantielle Empfehlungen für die künftige Weiterentwicklung einer wissenschaftlichen Einrichtung.

## **2. Forschungsevaluationen der Erziehungswissenschaft im Hochschulkontext**

Allgemein ist festzuhalten, dass Forschungsevaluationen sowohl das innovative Potential der zu evaluierenden Einrichtungen, die Originalität der Fragestellungen und der methodischen Erarbeitung, das Gewicht der Forschungsarbeiten zur Weiterentwicklung eines Forschungsfeldes, aber auch den scientific und public impact von Forschungseinheiten bewertet (vgl. Boffo u.a. 1999; Fries 2003; Gogolin u.a. 2005).

*Pluralität methodologischer Zugangsweisen:* Es muss selbstverständlich sein, dass Forschungsevaluationen einen allgemeinen und weiten Forschungsbegriff für sich in Anspruch nehmen, denn Forschung darf auf keine spezifische methodische Vorgehensweise eingeschränkt werden, d.h. Feldstudien, Einzelfallanalysen, die Begleitevaluation von Entwicklungsprojekten, experimentelle Forschung etc. haben prinzipiell eine vergleichbare Bedeutung im Kontext von Forschungsevaluationen in der Erziehungswissenschaft. Notwendig allerdings ist bei jeder forschungsmethodischen Strategie die Orientierung an Validität, Reliabilität und Objektivität der Forschung. Insbesondere müssen Forschungsergebnisse nachvollziehbar und nachprüfbar sein und die Ergebnisse müssen in gut zugänglichen, angesehenen, teilweise auch peer-reviewten und internationalen Zeitschriften und Buchbeiträgen veröffentlicht werden.

*Strategisches Wissensmanagement:* Erziehungswissenschaftliche Forschung ist mit der Dissemination von Ergebnissen verbunden, so dass sie an Wert gewinnt, wenn die Öffentlichkeit und insbesondere pädagogische Praxisfelder oder auch bildungspolitische Entscheidungsbereiche auf entsprechende Ergebnisse erziehungswissenschaftlicher Forschung eingehen können. Auch dieser Wissenstransfer im Zusammenhang von Bildungs- und Wissensmanagement, die Fähigkeit zur Kommunikation von Forschungsergebnissen, von Speicherungsformen und Transferpotentialen sind wichtige Kriterien für die Evaluation und die Qualität erziehungswissenschaftlicher Forschung (und Lehre). Wissensmanagement und insbesondere Wissenstransfer richtet sich auf die Weitergabe und Kooperation mit der Scientific Community, der Politik und der Praxis, insofern handelt es sich um den „public impact“ als ein zusätzliches Qualitätskriterium.

*Kontinuität und Kohärenz:* Eine Forschungseinheit muss ihre Arbeitsschwerpunkte, ihre zentralen Aufgaben und auch den Wandel ihres Auftrags erläutern können. Sofern erziehungswissenschaftliche Institute bereits in der Vergangenheit nach klaren und systematischen Kriterien evaluiert wurden oder beispielsweise bei Leibniz-Instituten bereits Evaluationen durch den Wissenschaftsrat erfolgt sind, kann aufgezeigt werden, wie auf die Empfehlungen einer vorausgehenden Evaluation eingegangen wurde und inwieweit eine Umsetzung von Empfehlungen tatsächlich erfolgt ist. Die nationale und internationale Bedeutung des Forschungsprogramms, aber auch einzelner Mitglieder von Forschungseinheiten gilt es genauso zu erfassen, wie das Entwicklungspotential des Forschungsfeldes selbst. Dabei geht es sowohl um die Kohärenz als auch die Kumulation der Forschungsarbeiten, denn es muss in einer heute hochdifferenzierten Forschungslandschaft ein Fokus deutlich erkennbar sein.

*Originalität, Innovation und internationale Kooperation:* Selbstverständlich wird von den Experten einer Evaluation auch die Originalität von Ansätzen und Fragestellungen im jeweiligen Forschungsfeld beurteilt. In pädagogischen Forschungsbereichen mag dies nicht immer leicht konsensuell feststellbar sein, weil verschiedene theoretische und methodische Strömungen konkurrieren, aber es gibt doch anerkannte operationalisierbare nationale und internationale Indikatoren, die über die Bedeutung eines Forschungsansatzes und einer -einheit Aufschluss geben: Höhe und Zusammensetzung von Drittmitteln, Mitgliedschaft in Herausgebergremien von angesehenen wissenschaftlichen Zeitschriften, DFG-Gutachtertätigkeit, Gutachtertätigkeit für internationale Forschungsorganisationen, herausragende Funktionen im Wissenschaftssystem sowie die Präsenz bei internationalen und nationalen Kongressen und Konferenzen sind nur einige Indikatoren, die doch klar benannt und eingeschätzt werden können. Allerdings sollte bei der Bewertung internationaler Kontakte und Aktivitäten die gesamte Breite möglicher internationaler Kontakte und Kooperationen berücksichtigt werden und ein „bias“ sollte vermieden werden, denn sowohl Kooperationen mit anglo-amerikanischen wie mit spanischen, französischen, chinesischen etc. Partnern haben gegebenenfalls hohen internationalen Wert. Sicher aber ist, dass die internationale Sichtbarkeit und Vernetzung von Arbeitseinheiten im Prozess der Globalisierung kontinuierlich an Bedeutung gewinnen.

*Nachwuchsförderung:* Das wichtigste Instrument für die langfristige Sicherstellung eines Forschungsbereichs oder auch einer Disziplin, ist die Förderung des wissenschaftlichen Nachwuchses. Besonders in den formativen Perioden der Entwicklung des Potentials von Wissenschaftlern werden Normen, Werte und Wissen aufgebaut, die es im weiteren Verlauf der je individuellen wissenschaftlichen Entfaltung ermöglichen in der internationalen scientific community anerkannt zu werden. Es ist zu konstatieren, dass es im Bereich der Erziehungswissenschaft Defizite bei der Nachwuchsförderung gibt und dass die Erziehungswissenschaft immer wieder darauf angewiesen ist, aus anderen Disziplinen Experten in die eigenen Forschungsbereiche zu integrieren. Dies hat einerseits sinnvoll zur interdisziplinären Bearbeitung pädagogischer Fragestellungen beigetragen, ist andererseits gleichzeitig aber auch ein Problem mangelnder kontinuierlicher Qualifizierung von Nachwuchskräften. Es darf nicht übersehen werden, dass sich bei Probleme-

men der Nachwuchsförderung implizit und explizit auch Schwierigkeiten und strukturelle Probleme der jeweiligen Disziplin zeigen. Wenn andere Fächer beispielsweise günstigere und bessere quantitative Relationen zwischen Professoren und wissenschaftlichen Mitarbeitern haben als die Erziehungswissenschaft (vgl. Tippelt/Rauschenbach/Weishaupt 2004) wirkt sich dies auf die Möglichkeiten der Nachwuchsförderung aus. Sicher sind die Anzahl der abgeschlossenen Promotionen, der Habilitationen, der Diplome und Magister wichtige Indikatoren für die Leistungsfähigkeit einer Forschungseinrichtung bzw. -einheit. Zusätzlich können Graduiertenkollegs, Sommerschulen, nationale und internationale Workshops, Kolloquien oder auch Kongresse Nachwuchskräfte gezielt motivieren und integrieren. Dies prägt die fördernden Aktivitäten von fach- oder disziplinbezogenen Forschungsgesellschaften – wie z.B. die DGfE – wesentlich.

Besonders wichtig dürfte es sein, auch kooperativ und in kleinen Gruppen zu publizieren. Senior scientists können ihre Erfahrungen einbringen, junior scientists überzeugen häufig durch ihren Enthusiasmus und ihre hohe Motivation und die Bereitschaft zur Innovation. Zu bewerten sind die Möglichkeiten, Gastaufenthalte an internationalen Einrichtungen oder an Einrichtungen im Ausland durchzuführen; wichtig sind die ausländischen Stipendiaten und ausländischen Wissenschaftler in der eigenen Forschungseinheit. EU-Projekte haben auch bei der Nachwuchsförderung eine wachsende Bedeutung.

Ohne die historische und theoretische Dimension der Erziehungswissenschaft zu vernachlässigen, dürften im jetzigen Stadium der Entwicklung der Disziplin gezielte Methodenkurse – sowohl der quantitativen wie auch der qualitativen empirischen Fundierung – von allergrößter Bedeutung sein. Die solide und gleichzeitig innovative Nutzung von Methodenkompetenzen sichert den Nachwuchskräften eine gewisse Flexibilität bei der Bearbeitung von wissenschaftlichen Themen und damit zugleich Mobilität beim Zugang zu wissenschaftlichen Einrichtungen.

An dieser Stelle ist erneut zu betonen, dass erziehungswissenschaftliche Evaluationen an Hochschulen nicht nur die Forschungsleistungen einer Einrichtung bewerten, sondern immer auch Lehrevaluationen (vgl. Schmidt/Tippelt 2004) und Strukturrevaluationen (vgl. Böllert/Tippelt 2004; EWFT 2003a) beinhalten. Insofern sind Forschungsevaluationen nur ein Teil der Qualitätssicherung an Hochschulen (Hochschulrektorenkonferenz 2003) – freilich kann mit guten Gründen die Position vertreten werden, dass gute Forschungsleistungen eine notwendige, wenn auch nicht hinreichende Bedingung für exzellente Lehrleistungen – insbesondere im Master- und Graduiertenbereich – darstellen.

### **3. Evaluation von außeruniversitären Forschungseinrichtungen**

Als exemplarisches Beispiel von außeruniversitären Forschungsevaluationen werden die in verschiedener Hinsicht modellhaften Forschungsevaluationen der Leibniz-Gemeinschaft dargestellt (zu berücksichtigen sind jeweils spezifisch auch das Max-Planck-

Institut für Bildungsforschung, das Deutsche Jugendinstitut, das Bundesinstitut für Berufsbildung, Einrichtungen der Helmholtz- oder Fraunhofer-Gesellschaft u.a.). Bedeutungsvoll sind bis zum Jahr 2007 die Ergebnisse des Evaluationsausschusses der Bundesländer-Kommission für Bildungsplanung und Forschungsförderung (BLK), der mindestens alle sieben Jahre auf der Grundlage von strengen und unabhängigen Gutachten überprüft, ob beispielsweise die Einrichtungen der in der Leibniz-Gemeinschaft zusammengeschlossenen über 80 wissenschaftlichen Einrichtungen weiterhin die Förder Voraussetzungen erfüllen (vgl. WGL Leibniz-Gemeinschaft 2003). Früher legte die BLK ihren Entscheidungen Stellungnahmen des Wissenschaftsrates zugrunde, aber seit 2003 stützt sich der Ausschuss „Forschungsförderung“ der BLK in der Regel auf Stellungnahmen des Senats der Leibniz-Gemeinschaft, der seinerseits einen Senatsausschuss „Evaluierung“ eingesetzt hat, um entsprechende Forschungsevaluationen durchzuführen. Auch wenn die Aufgaben der BLK aufgrund der veränderten Föderalismusregelungen künftig von anderen Institutionen bearbeitet werden, so bleibt doch die Notwendigkeit einer finalen Entscheidung über Förderempfehlungen auch künftig bestehen.

In der Wissenschaftsgemeinschaft Leibniz befinden sich ca. 80 wissenschaftliche Einrichtungen aus den Bereichen der Ökonomie, Medizin, Jurisprudenz, Naturwissenschaft, Sozial- und Geisteswissenschaft und auch vier explizit erziehungswissenschaftliche Einrichtungen, die nach den gleichen Kriterien evaluiert werden wie die Einrichtungen aus den anderen wissenschaftlichen Disziplinbereichen (siehe oben): das Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel (IPN), das Deutsche Institut für Erwachsenenbildung in Bonn (DIE), das Deutsche Institut für Internationale Pädagogische Forschung in Frankfurt (DIPF), das Institut für Wissensmedien in Tübingen (IWM). Zu nennen wäre auch die Gesellschaft sozialwissenschaftlicher Infrastruktureinrichtungen in Mannheim und Bonn (GESIS), die in vielfältiger Weise auch pädagogische Forschungsprojekte berät und unterstützt, sowie das Zentrum für psychologische Information und Dokumentation (ZPID) in Trier, das u.a. ein breites Arsenal psychologisch und pädagogisch relevanter Tests bearbeitet und für Nutzer bereitstellt.

Diesen Instituten ist eine permanente interne Evaluation auferlegt (Audit alle zwei Jahre) und ein wissenschaftlicher Beirat – eine Empfehlung des Wissenschaftsrats – soll permanent die wissenschaftliche Arbeit der Einrichtungen kritisch beraten.

Die alle sieben Jahre sich wiederholende externe Evaluation findet jeweils unter Mitwirkung unabhängiger Gutachter durch nationale und internationale wissenschaftliche Experten statt, die ihrerseits einen Evaluationsbericht mit Bewertung und mit Empfehlungen erarbeiten. Diese Empfehlungen und Bewertungen werden immer in einer unabhängigen Abschlussbesprechung abgestimmt, dann dem Senatsausschuss „Evaluation“ der WGL vorgelegt, im Senat beraten und abgestimmt und dann der BLK (so bis in das Jahr 2008) sowie den betroffenen Landesministerien vorgelegt (vgl. Wissenschaftsgemeinschaft GW Leibniz 2003). Der siebenjährige Evaluationszyklus hat sich bewährt, weil man den Instituten Gelegenheit geben muss, Empfehlungen sinnvoll und nachhaltig umzusetzen. Lediglich bei Sichtung problematischer Entwicklungen wird von diesem 7-Jahresrhythmus abgegangen und es können dann kurzfristige Zwischenevaluationen vereinbart werden.

Speziell für Institutionen der Leibniz-Gemeinschaft gilt es zu klären, in welchem Verhältnis außeruniversitäre und universitäre Forschung stehen und es gilt die Frage zu beantworten, ob Arbeiten, die in außeruniversitären Einrichtungen durchgeführt werden, von Universitäten ebenfalls leistbar wären (sogenanntes „Alleinstellungsmerkmal“). Auf die Kooperation und Vernetzung mit universitären Forschungseinrichtungen wird explizit eingegangen.

### *Strukturmerkmale von Forschungseinrichtungen*

Zur Bewertung und zur Einordnung von Forschungsleistungen ist die Kenntnis des organisatorischen Aufbaus einer Institution, also ein Organigramm unabdingbar und es gilt in außeruniversitären Einrichtungen auch das Verhältnis von Forschung und den ebenfalls als wichtig zu betrachtenden Serviceleistungen zu beurteilen. Es wäre unfair eine Einrichtung mit relevanten Serviceeinrichtungen ausschließlich nach „harten“ Indikatoren der Forschung (Drittmittel, internationalen Publikationen etc.) zu beurteilen. Serviceleistungen oder die sogenannten Querschnittsaufgaben an Hochschulen sind eigenständig zu berücksichtigen. Zu den Strukturmerkmalen gehören die Maßnahmen eines regelmäßigen Qualitätsmanagements (z.B. gezielte Zwischenevaluationen, Zielvereinbarungen mit Mitarbeitern und insbesondere dem wissenschaftlichen Nachwuchs etc.), die tatsächliche und kontinuierliche Kooperation mit einem aktiven wissenschaftlichen Beirat und im Falle von Leibniz-Einrichtungen mit starken wissenschaftsbasierten Dienstleistungen für Praxis und Politik auch einem Nutzerbeirat. Eigens zu erwähnen ist die in allen Forschungseinrichtungen wichtige Norm der Gleichstellung von Männern und Frauen bei den Forschungs- und Führungspositionen.

### *Mittelausstattung und Mittelverwaltung*

Grundsätzlich bedarf es einer guten und genauen Übersicht über Personal-, Sach- und Investitionsmittel, auch die räumliche Ausstattung ist zu dokumentieren und insbesondere müssen räumliche Engpässe bei den abschließenden Empfehlungen explizit behandelt werden. In Institutionen (dies gilt analog auch für pädagogische Hochschuleinrichtungen), die nicht hinreichend über personelle und sachliche Infrastruktur verfügen, kann nur die Einwerbung von Drittmitteln eine kontinuierliche Forschungstätigkeit sichern. Allerdings ist dies gerade unter solchen Bedingungen besonders schwierig. Bei der Einwerbung von Drittmitteln ist zu berücksichtigen, dass Nachwuchskräfte hier eine besondere Förderung erfahren können, was zusätzlich positiv zu bewerten ist. Jede Forschungseinrichtung ist gut beraten, einerseits kompetitiv bei der DFG oder anderen forschungsfördernden Institutionen Mittel einzuwerben; aber zur Förderung spezifischer und manchmal nur in komplexen Vernetzungen durchzuführender Projekte auch internationale Mittel, Bundes-, Landes-, Stiftungsmittel und auch Mittel der privaten Wirtschaft zu akquirieren. Ein besonderes Gewicht dürften zukünftig die nur in internationalen Kooperationen und Vernetzungen einzuwerbenden EU-Mittel erlangen.

Es gehört zum „hidden curriculum“ der Drittmittelinwerbung, dass kompetitiv eingeworbene Mittel (insbesondere der peer-reviewten DFG) immer ein bestimmtes Niveau der Forschung signalisieren und daher das Image der Einrichtung positiv prägen. Allerdings lassen sich selbstverständlich auch mit den anderen Mitteln aus Bundes- oder Landesetats, aber auch den bereits genannten privaten Mitteln der Wirtschaft sehr interessante Forschungsfragen, manchmal auch in einem breiten Umfang über längere Zeiträume, bearbeiten.

### *Personal*

Die Rekrutierung von wissenschaftlichem Personal, von Professuren aber auch von leitenden Personal in Forschungseinrichtungen muss dem Leistungsprinzip unterworfen sein. Nur eine internationale und bundesweite Rekrutierung des leitenden wissenschaftlichen Personals ist akzeptabel. Die Rufe an Hochschulen und an etablierte Forschungseinrichtungen sowie der Übergang von Habilitierten und Privatdozenten in Professuren sind ebenfalls wichtige Indikatoren für hervorragende wissenschaftliche Leistungen und eine anzustrebende nationale und internationale Sichtbarkeit. Generell ist auf die Altersstruktur zu achten, denn nur wenn wissenschaftlicher Nachwuchs in hinreichender Anzahl vorhanden ist, ist auch eine gezielte Nachwuchsförderung und damit die forschungsorientierte Weiterentwicklung eines Arbeitsgebiets möglich. In diesem Zusammenhang waren vor allem die Evaluationen der 90er Jahre bestrebt, den Anteil von flexiblen Stellen zu Lasten von lebenszeitlichen Stellen zu verändern, damit einerseits durch Rotation das Innovationspotential von Forschungseinrichtungen gestärkt wird und andererseits der jüngere wissenschaftliche Nachwuchs kontinuierlich eine Beschäftigungs- und Entwicklungschance erhalten kann. Die Relation von flexiblen und festen Stellen haben sich entsprechend deutlich verändert.

Positive Wirkungen zeigen auch Personalentwicklungskonzepte, d.h. die bedarfsgerechte Fortbildung – auch unabhängig von der Nachwuchsförderung – von technischem Personal, wissenschaftlichem Personal und den Sekretariaten.

### *Arbeitsergebnisse: Forschung und Entwicklung, Service*

Die Produktivität erziehungswissenschaftlicher Forschung spiegelt sich in der Anzahl und der Qualität wissenschaftlicher Publikationen. Bei der Qualitätsbeurteilung ist eine Forschungsevaluierung auf einfache Sachverhalte angewiesen: Sind die Publikationen in renommierten, einflussreichen oder auch peer-reviewten Zeitschriften erfolgt, ist die Internationalität der Publikationstätigkeit gesichert. Gleichzeitig wäre es falsch, bei Forschungsevaluationen ausschließlich diese Form der Publikation zu berücksichtigen, denn wenn man den Transfer von wissenschaftlichen Erkenntnissen in die Politik, die pädagogische Praxis und die Beratung ernst nimmt, sind auch praxisbezogene Zeitschriften, die stärker von Praktikern oder Planern gelesen werden, bei Forschungseva-

lationen zu berücksichtigen. Es darf nicht das Ziel von Forschung sein, ein träges Wissen in dem Sinne zu erzeugen, dass dieses Wissen nur einer eher kleinen scientific community vorbehalten bleibt. Insofern ist auch hier eine Diversifizierung der Publikationsstrategie von Vorteil. Dennoch ist es zutreffend, dass es häufig vor allem die nationalen und internationalen peer-reviewten journals sind, an denen die Qualität von Publikationsleistungen eingeschätzt wird.

Für die Erziehungswissenschaft gilt darüber hinaus – wie für zahlreiche Sozial- und Geisteswissenschaften – noch immer die Tatsache, dass Monographien, insbesondere bei angesehenen Verlagen, eine hohe Bedeutung zuzusprechen ist, während Monographien im Eigenverlag oder im engen regionalen Kontext häufig als Publikationen mit geringerer Qualität interpretiert werden. Man kann Publikationstätigkeit auch als eine Form der Öffentlichkeitsarbeit ansehen: Zielgruppen hierbei sind sowohl die wissenschaftliche, die politische wie die praktische Öffentlichkeit. Publikationen in diesen drei sehr unterschiedlichen Bereichen, die häufig eine unterschiedliche Form der Präsentation und sogar der sprachlichen Darstellung erfordern, müssen also differenziert beurteilt werden. Online-Publikationen (free access) werden künftig zunehmend beachtet werden, aber derzeit wird deren Bedeutung bei Forschungsevaluationen noch zurückhaltend gewichtet.

Sofern es sich um Serviceeinrichtungen handelt, müssen unbedingt Nutzerstatistiken und Einnahmen aus Serviceleistungen in die Beurteilung von Leistungen einer Einrichtung eingehen.

Sicher ist auch der „public impact“ von Abteilungen oder auch von Einzelpersonen zu prüfen, d.h. es geht in Forschungsevaluationen ein, ob einflussreiche externe Veranstaltungen oder Kongresse organisiert und geleitet wurden, ob wichtige Vorträge national und international gehalten, ob wichtige wissenschaftliche Ämter und Funktionen bekleidet werden oder auch Auszeichnungen und Ehrungen erfolgt sind. Diese Aspekte werden als eher weiche Indikatoren definiert.

#### **4. OECD-reviews – Länderberichte des Centers of Educational Research and Innovation (CERI, Paris)**

OECD/CERI hat in den letzten Jahren auf Anforderung von New Zealand, Denmark, Mexico, United Kingdom und Switzerland Evaluationen durchgeführt, „to review the extent to which the educational R&D system within a country is contributing effectively to the creation, collation and distribution of a knowledge base on which practitioners and policy-makers can draw. The aim of this reviews is broader than that of an educational R&D review exclusively focused on the quality of the research delivered. The focus of this reviews is on the evaluation of the contribution of educational R&D to the knowledge base of education in the emerging learning society, to improve both the process of decision-making at policy level and educational practices“ (OECD/CERI 2006, S. 4).

### *Analytischer Prozess*

Die analytische Herangehensweise dieser internationalen Forschungsevaluationen basiert auf aktuellen Arbeiten von OECD/CERI zu Formen des Wissensmanagements und seiner Anwendung auf Erziehungswissenschaftliche Forschung und Entwicklung (OECD/CERI, 2003; 2004a). Drei der zunächst durchgeführten reviews wendeten instrumentell ein Evaluationsdesign an, das im Kontext der Evaluation des dänischen erziehungswissenschaftlichen Forschungs- und Entwicklungssystems entstanden war (OECD/CERICERI, 2004b). Das Evaluationsteam für die Schweiz schloss an das Evaluationsdesign unmittelbar an, berücksichtigte aber die Kritik und die Anregungen, die an den zurückliegenden Forschungsevaluationen von OECD/CERI formuliert wurden (vgl. Schuller u.a. 2006), und nahm die immer notwendigen regionalen und nationalen Anpassungen und Adjustierungen an die besondere Situation eines Landes (z.B. der Schweiz) vor. Es entstand ein restrukturiertes analytisches Instrument, das der besonderen Situation der Schweiz besser angemessen war als die zurückliegenden Evaluationsdimensionen (vgl. OECD/CERI 2006). In einem nach der Evaluation durchgeführten internationalen OECD/CERI Seminar 2007 in Bern wurden die neuen analytischen Dimensionen und Indikatoren zur Evaluation von Forschung und Entwicklung international vergleichend diskutiert und es wurde noch einmal der breitere Kontext, in dem diese besonderen Länderberichte entstehen, von einer internationalen Expertengruppe reflektiert. Die wichtigsten Ziele wurden dabei festgehalten: Prüfung des Beitrags von Bildungsforschung und -entwicklung für eine evidenzbasierte Bildungspolitik, für eine Verbesserung der pädagogischen Praxis und für die Weiterentwicklung der theoretischen Wissensbasis in den Bereichen Bildung, Erziehung und lebenslanges Lernen.

Exemplarisch soll kurz der Evaluationsablauf am Beispiel der Schweiz dargestellt werden: Der nationale Vertragspartner, in diesem Fall "Swiss Council for Educational Research (CORECHED)", beauftragte das „Swiss Coordination Centre for Research in Education (SKBF)“ das gesamte Programm zu planen und insbesondere einen „Country Background Report“ (vgl. Denzler-Schircks, 2006) auszuarbeiten, welcher der unabhängigen internationalen Kommission (A. Pollard von der Universität London, M. van der Wende von der Universität Twente und der Freien Universität Amsterdam, R. Tippelt von der LMU München und F. Pedro von CERI) zur Orientierung dienen sollte. S. Wolter von der Universität Bern und SKBF Director agierte als nationaler Koordinator.

„This background report provided a good deal of information on the management and steering of knowledge in education, the organisation and production of such knowledge, and its applications and outcomes in Switzerland. As a result, the reviewers were able to better prepare themselves to understand and analyse the Swiss system. At the same time, this background report was widely circulated in Switzerland and most of the people interviewed, if not all, had the chance to familiarise themselves with it, supplement the data provided in it and respond to its main conclusions.“ (OECD/CERI 2006, S. 5)

Auch im Rückblick der Erstellung des Länderberichts wird man sagen können, dass die Inhalte, Aktualität und Qualität dieses Hintergrundberichts für die Durchführung

eines Länderberichts von großer Bedeutung ist. Dieser Bericht sensibilisiert – ähnlich den in anderen Kontexten entstehenden Selbstreports – ein internationales review-team für Stärken und Schwächen der zu analysierenden Forschungs- und Entwicklungslandschaft in einem Land – entsprechend aufmerksam und sensibel wird eine nationale scientific community darauf reagieren. Zu keiner Zeit geht es bei diesen reviews um die Evaluierung einzelner Institutionen oder Disziplinen, immer richtet sich der Blick auf die Leistung des gesamten Forschungssystems im Kontext von Bildung, Erziehung und Lernen (grundsätzlich siehe Keiner 2005; Keiner/Schriewer 2000).

Neben dem „background report“ sind es vor allem die Informationen und Interpretationen der Akteure in einem Forschungs- und Entwicklungssystem, die während eines fünftägigen Aufenthalts im Land von den internationalen Experten interviewt werden können. In zahlreichen Interviews mit Akteuren aus den (1) Universitäten, Pädagogischen Hochschulen, Fachhochschulen, (2) verschiedenen Bundesämtern und statistischen Ämtern, (3) relevanten Wissenschafts- und Berufsorganisationen, (4) der nationalen Forschungsgesellschaft, (5) der Konferenz der kantonalen Erziehungsdirektoren und einschlägiger Ministerien, (6) der Rektorenkonferenz und (7) in spezifischen Gesprächen mit Studierenden und Wissenschaftlern an mehreren Standorten versuchte sich die internationale Expertengruppe ein unabhängiges Bild zu machen (vgl. Tippelt 2007).

Um die Einschätzungen zur Qualität, zu den Stärken und den Schwächen eines nationalen Forschungs- und Entwicklungssystems zu präzisieren, sind Definitionen für Grundlagenforschung, angewandter Forschung und erziehungswissenschaftlicher Entwicklungsarbeit notwendig. Im Schweiz-review wurden folgende Definitionen zu Grunde gelegt:

„*Research* is defined as the process of knowledge creation that conforms to the agreed scholarly standards intended to warrant its validity and trustworthiness. In this report, *basic research* is differentiated from *applied research*. The former is driven by curiosity and an inherent interest in a phenomenon or problem, while the latter is consciously designed to solve a problem in policy or practice. In both cases, the process of knowledge creation is carried out within the framework of a theory, which might be either validated or challenged by new research. *Development* is defined as any form of knowledge creation designed to improve practice. Thus, the main purpose of development is to facilitate change in a particular context. A number of educational developments are teacher-lead activities and consist of *enquiry-based activities* that take place within schemes for the professional development of teachers“ (OECD/CERI 2006, S. 5).

Die Grundannahme eines Länderberichts von OECD/CERI besteht darin, dass ein erziehungswissenschaftliches Forschungssystem relevantes Wissen herausarbeitet, organisiert und disseminiert, das geeignet ist Veränderungen des Bildungs- und Erziehungssystems frühzeitig zu erkennen, damit auf der Basis empirischer und theoretischer Erkenntnis Steuerungs- und Orientierungsprozesse für das politische und das praktische Handeln eingeleitet und begründet werden können.

Die Teilnahme eines Landes am OECD-review impliziert eine Reihe von grundlegenden (auch normativen) Prinzipien der politischen Steuerung und des rationalen

Planens: eine partizipative Demokratie, eine kritische Öffentlichkeit und Offenheit für kritische Debatten, das Bewusstsein für regionale und kulturelle Unterschiede und Diversität, die Anerkennung von demokratischen Rechten und Pflichten von Bürgern, die Transparenz der Vergabe von finanziellen Mitteln aus öffentlichen Haushalten und insbesondere der Konsens darüber, dass Bildung und Erziehung der personalen Entwicklung jedes einzelnen, der sozialen und kulturellen Inklusion und der ökonomischen Sicherheit dient.

Vor dem Hintergrund solcher Ziele erscheinen nach der oben erwähnten intensiven Diskussion und Auseinandersetzung mit den bereits abgeschlossenen Länderberichten (vgl. Schuller u.a. 2006) sowie nach der ebenfalls erwähnten internationalen Konferenz von CERI in Bern 2007 sechs Untersuchungsbereiche (generic template) sinnvoll, um die Bildungs- und Erziehungsforschung unter der Vorgabe qualitativer Verbesserungen zu analysieren und auch vergleichend zu verstehen: Rahmenbedingungen (contextual issues), strategisches Bewusstsein (strategic awareness), Grundlagenforschung (basic research), angewandte Forschung (applied research), Entwicklung und professionelle Erkundung (development and professional enquiry), systemspezifische Themen (systemic issues). Um die konkreten Anliegen eines Reviews zu verdeutlichen, werden im folgenden die diesen Untersuchungsbereichen zugeordneten Leitfragen benannt (vgl. OECD/CERI 2006, S. 7ff.):

### *Contextual issues*

- What is the political, economic, social and cultural context of the country?
- What are the country's aspirations and strategies for development?
- What is the nature of the country's existing educational R&D?
- What are the major contemporary challenges to the country's educational R&D?

### *Strategic awareness*

- What is the extent and quality of the country's current knowledge about its own educational system?
- What provision is there for the accumulation and organisation of existing educational knowledge (basic, applied and developmental) in the country?
- How committed are the country's key stakeholders to the introduction of a national system for managing the production and use of educational evidence and knowledge?
- Does the country have a national policy or strategy for educational R&D, with a clear definition of what counts as basic and applied research and what counts as forms of development by practitioners and others?
- Basic research
- Does the country have appropriate provisions and incentives for the production of high quality and innovative basic research?

*Applied research*

- Does the country have appropriate provisions and incentives for the production of high quality and relevant applied research?
- How are researchers, policy-makers, practitioners and other appropriate stakeholders in the country engaged in the identification, development, application and evaluation of national priorities for applied research and for development?
- Development and professional enquiry
- Does the country have appropriate provisions and incentives for the production of high quality and relevant development work, professional enquiry and improvement, and how is it embedded in the provision for the education and training of practitioners?

*Systemic issues*

- How are the country's various research and development activities distributed, networked and coordinated nationally?
- How is the country's research and development work linked to appropriate international networks, centres and activities?
- What quality assurance and accountability procedures are in place for the country's educational research and development?
- What provision is there for the communication and dissemination of research findings to the country's stakeholders, including the general public, and how effective is this knowledge transformation and transfer?
- Is there adequate capacity building to sustain the country's complementary forms of educational research and development?

Mit diesem „generic template“ kann einem Teil der zu Recht gegenüber den ersten Länderberichten geäußerten kritischen Aspekten begegnet werden (vgl. Schuller, u.a. 2006; auch OECD/Mexico International conference 2006). Dem zurückliegenden ‚generic template‘ wurde vorgehalten, dass folgendes unzureichend sei:

- „The criticism included suggestions that the generic template;
- Offered a single model, with the implication that this represents ‚best practice‘ for all countries – thus contributing to international homogeneity;
- Was over-determined by the OECD's emphasis on economic priorities and concern for technical efficiency in markets, and did not relate education knowledge production and its application to democratic and other political contexts;
- Assumed that the existence of shared aims, values and commitment to education was an applied ‚consumer-led‘ discipline rather than reflecting a ‚curiosity-led‘ interest;
- Was not contextualised in relation to any particular country or its educational research – for instance, concerning the working conditions and or the motivations of researchers“ (OECD/CERI 2006, S. 6).

Das Untersuchungsdesign der neuen Länderberichte, die die Qualität von erziehungswissenschaftlichen Forschungs- und Entwicklungssystemen verbessern sollen, fordert eine Kultur des Forschungsmanagements, die effektive und demokratische Bildungspolitik unterstützt und die Forschungskommunikation zwischen Wissenschaft einerseits und Politik und Praxis andererseits fördert. Eine innovative, kreative Grundlagenforschung ist dabei notwendig, allerdings wird auch an die Verantwortung der Bildungsforschung appelliert, an der Lösung von aktuellen Problemen im Bildungs- und Erziehungssystem dadurch mitzuwirken, dass eine evidenzbasierte Bildungs- und Erziehungsplanung durch die Erarbeitung belastbarer empirischer Erkenntnisse unterstützt wird und sich Bildungs- und Erziehungsforscher aktiv in den Prozess der Wissenskommunikation und des Wissenstransfers einmischen. Freilich impliziert diese Forderung an die Erziehungswissenschaft korrespondierende Forderungen an die Politik: Interesse an freier und offener Forschungskommunikation, Finanzierung von Projekten und aktives Reagieren auch auf kritische Ergebnisse der Bildungs- und Erziehungsforschung.

## **5. Zur Konstruktion von Qualitätsstandards als Kriterien für exzellente Forschung**

Wenn man abschließend noch einmal die wichtigsten Kriterien für eine exzellente Forschungsleistung von erziehungswissenschaftlichen Arbeitseinheiten resümiert, wird man formal festhalten, dass das Drittmittelvolumen (insbesondere hierbei auch die kompetitiv eingeworbenen Mittel), die Publikationsintensität (insbesondere der Anteil auch in renommierten oder peer-reviewten Zeitschriften), die internationale Präsenz, die Nachwuchsförderung (hier insbesondere die Anzahl und Dauer von Qualifizierungsleistungen), die internationale Vernetzung einer Forschungseinheit wie auch der Wissenstransfer in Bereiche der Politik und der Praxis wichtige Kriterien erziehungswissenschaftlich exzellenter Forschung sind. Diesen Standards und Kriterien stellen sich die erziehungswissenschaftlichen Institute, die in der Leibniz-Gemeinschaft vertreten sind, diese Standards haben aber auch ihre Berechtigung zur Beurteilung erziehungswissenschaftlicher Forschung in universitären Einrichtungen.

Allerdings muss deutlich darauf hingewiesen werden, dass universitäre Einrichtungen auch der Lehre verpflichtet sind und es ist explizit zu betonen, dass auch die Lehre Qualitätsstandards entsprechen soll (vgl. Schmidt/Tippelt 2005). Insofern ist jede erziehungswissenschaftliche Evaluation, die entweder die Lehre oder die Forschung einseitig berücksichtigt, im Hochschulkontext nicht hinreichend. Eindeutig müssen Belastungs- und Zeitfaktoren in Bewertungen einbezogen werden, denn wenn lediglich idealtypische Qualitätserwartungen an Lehre und Forschung herangetragen werden ohne deren Kontext- und Realisierungsbedingungen fair zu prüfen und zu reflektieren, dürfte dies nachvollziehbare Abwehrhaltungen der evaluierten Einrichtungen zur Folge haben.

Die Dreiteilung von Evaluationsprogrammen in der Erziehungswissenschaft, in Struktur-, Lehr- und Forschungsevaluation hat sich bereits durchgesetzt. Sicher werden alle Formen dieser Evaluationen, insbesondere aber auch die hier speziell behandelten

Forschungsevaluationen nur zu wirklichem Feedback führen, wenn Konsequenzen zur Verbesserung der Prozessqualität erziehungswissenschaftlicher Forschung in den jeweiligen Arbeitseinheiten realisiert werden. Die mehrmalige Evaluation in kurzen Abständen, also eine „Evaluationsverdichtung und -überfrachtung“ (es gibt Einheiten, die innerhalb von zwei Jahren vier Forschungsevaluationen erfahren haben) ist kontraproduktiv. Forschungsevaluationen in der Erziehungswissenschaft sollen Qualität und Effektivität erhöhen und dies wird dann erreicht, wenn nach den hier explizierten Bewertungskriterien einer Forschungseinheit eine eindeutige und klare Rückmeldung gegeben wird, evtl. auch Anregungen und Empfehlungen und sogar Auflagen formuliert werden, die dann in einem überschaubaren Zeitraum von ca. fünf bis sieben Jahren von der jeweiligen Einheit umgesetzt werden können.

## Literatur

- Arvanitis, S./Wörter, M. (2005): *The Swiss Innovation: Governance, Public Policy, Performance and Assessment of Strengths and Weaknesses*. Background Report to the OECD Country Review of Switzerland's Innovation Policy Zurich: Swiss Federal Institute of Technology Zurich/Swiss Institute for Business Cycle Research.
- Boffo, S./Chave, D./Kaukonen, E./Opdal, L.R. (1999): *The Evaluation of Research in European Universities*. In: *European Journal of Education* 34/3, S. 325–334.
- Böllert, K./Tippelt, R. (2004): *Evaluationsverfahren in der Erziehungswissenschaft*. In: Tippelt, R./Rauschenbach, T./Weishaupt, H. (Hrsg.): *Datenreport Erziehungswissenschaft 2004*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 139–152.
- Daniel, H.-D. (1988): *Evaluation von Forschung: Methoden, Ergebnisse, Stellungnahmen*. Konstanz: Univ.-Verlag.
- Denzler-Schircks, S. (2006): *Educational Research and Development in Switzerland*. Country Background Report. Aarau: Swiss Coordination Center for Research in Education (SCCRE).
- Ditton, H. (2002): *Evaluation und Qualitätssicherung*. In: Tippelt, R. (Hrsg.): *Handbuch Bildungsforschung*. Opladen: Leske + Budrich, S. 775–790.
- EWFT (2003a): *Erziehungswissenschaftlicher Fakultätentag (Hrsg.): Evaluation der Erziehungswissenschaft. Bestandsaufnahme, Selbstevaluationsleitfaden, Erhebungsinstrumente*. Münster.
- EWFT (2003b): *Erziehungswissenschaftlicher Fakultätentag (Hrsg.): Vorschlag für einen Frageleitfaden für den Selbstreport als Vorbereitung einer Peer-Evaluation*. Münster.
- Fries, M. (2003): *Evaluation von Forschung und Lehre in Bayern*. In: Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung (Hrsg.): *Beiträge zur Hochschulforschung München*, 25 (2003), 1, S. 6–31.
- Gogolin, I. u.a. (2005): *Standards und Standardisierungen in der Erziehungswissenschaft*. editorial, In: *Zeitschrift für Erziehungswissenschaft*, Beiheft 4/2005, Wiesbaden: VS Verlag für Sozialwissenschaften, S. 5–12.
- Gretler, A. (2000): *Die Schweizerische Bildungsforschung der Nachkriegszeit im Spiegel ihrer Institutionen und ihrer Themen – von Geschichte zu aktuellen Fragestellungen*. Schweizerische Zeitschrift für Bildungswissenschaften, 22(1), S. 111–144.
- Hanft, M. (2003): *Seelsorge im Umbruch: Gedanken zur Zukunft der Pfarrgemeinde*. In: *650 Jahre Pfarrei Erzengel Michael Thüngersheim, Thüngersheim*, S. 53–5.
- Hornborstel, S./Keiner, E. (2002): *Evaluation der Erziehungswissenschaft*. In: *Zeitschrift für Erziehungswissenschaft*, H. 4, S. 634–654.
- Hochschulrektorenkonferenz (Hrsg.) (2003): *Wegweiser 2003. Qualitätssicherung an Hochschulen*. Bonn.

- Keiner, E. (2005): Zur Konstruktion erziehungswissenschaftlicher Forschung aus der Perspektive der OECD. In: Zeitschrift für Erziehungswissenschaft, Beiheft 4/2005, S. 13–24.
- Keiner, E. & Schriewer, J. (2000): Erneuerung aus dem Geist der eigenen Tradition? Über Kontinuität und Wandel nationaler Denkstile in der Erziehungswissenschaft. Schweizerische Zeitung für Bildungswissenschaften, 22(1), S. 27–49.
- McKinney, K. (1999): Attaining Excellence: A Handbook on the Standards for the Conduct and Evaluation of Research Carried Out by the Office of Educational Research and Improvement. Columbia.
- Merkens, H. (Hrsg.): Evaluation in der Erziehungswissenschaft. Wiesbaden.
- OECD (2002): Review of Tertiary Education Policy in Switzerland. Paris.
- OECD (2006): Economic Survey of Switzerland, 2005. Paris.
- OECD/CERI (2000): Knowledge Management in the Learning Society. Paris.
- OECD/CERI (2003): New Challenges for Educational Research. Paris.
- OECD/CERI (2004a): Innovations in the Knowledge Economy: Implications for education and learning systems. Paris.
- OECD/CERI (2004b): National Review on Educational R&D. Examiners report on Denmark. Paris.
- OECD/CERI (2006): OECD Review of Switzerland's Educationale R&D System. Paris.
- Reissert, R./Carstensen, D. (1998): Praxis der internen und externen Evaluation. Handbuch zum Verfahren. HIS Kurzinformation.
- Sanders, J.R. (Hrsg.) (1999): Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“. Opladen: Leske + Budrich.
- Schenker-Wicki, A. (1996): Evaluation von Hochschulleistungen. Leistungsindikatoren und Performance Measurements. Wiesbaden: Dt. Univ.-Verlag.
- Schmidt, B./Tippelt, R. (2005): Lehrevaluation. In: Zeitschrift für Erziehungswissenschaft, Beiheft 4/2005, S. 227–242.
- Schuller, T./Jochems, W./Moos, L./Van Zanten, A. (2006): Evidence and Policy Research. European Educational Research Journal 5(1), S. 57–70.
- Speck-Hamdan, A./Tippelt, R. (2001): Evaluation im Prozess – Zeitliche, sachliche und soziale Dimensionen einer Evaluation in der Erziehungswissenschaft. In: Keiner, E. (Hrsg.): Evaluation (in) der Erziehungswissenschaft. Weinheim: Beltz, S. 155–162.
- Tillmann, K.-J./Rauschenbach, T./Tippelt, R./Weishaupt, H. (Hrsg.): Datenreport Erziehungswissenschaft 2008 – erstellt im Auftrag der DGfE, Opladen 2008.
- Tippelt, R. (2004): Standards und Kriterien von Forschungsevaluationen in der Erziehungswissenschaft. In: Merckens, H. (Hrsg.), Evaluation in der Erziehungswissenschaft. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 89–100.
- Tippelt, R. (2007): The demand side of educational Research & Development. In: Joint OECD-CORECHED-International Expert meeting. Bern (unveröffentl. Manuskript).
- Tippelt, R./Rauschenbach, T./Weishaupt, H. (Hrsg.) (2004): Datenreport Erziehungswissenschaft 2004 – erstellt im Auftrag der DGfE. Opladen.
- Vogel, C. (2006): Recherche et développement au sein des hautes écoles pédagogiques et des établissements universitaires de formation des enseignants. Berne: Conférence suisse des recteurs des hautes écoles pédagogiques.
- Wesseler, M. (1999): Evaluation und Evaluationsforschung. In: Tippelt, R. (Hrsg.): Handbuch Erwachsenenbildung/Weiterbildung. Opladen: Leske + Budrich, S. 736–752.
- WGL Leibniz-Gemeinschaft (2003): Die Leibniz-Gemeinschaft. Jahrbuch 2003. Bonn.
- Wottawa, H./Thierau, H. (1998): Lehrbuch Evaluation. Bern. Huber.

*Anschrift des Autors:*

Prof. Dr. Rudolf Tippelt, Lehrstuhl Allg. Pädagogik und Bildungsforschung, Institut für Pädagogik, Ludwig-Maximilians-Universität, Leopoldstr. 13, 80802 München.



# **Implementierung von Qualitätssicherung**

Helmut Fend

## Dimensionen von Qualität im Bildungswesen

Von Produktindikatoren zu Prozessindikatoren am Beispiel der Schule

### 1. Qualitätsentwicklung seit 2000

Es ist hilfreich, bei einer Bilanzierung der Qualitätssicherung im Bildungssystem auf den Stand zurückzublicken, der etwa im 41. Beiheft der „Zeitschrift für Pädagogik“ beschrieben wurde, und daran zu messen, was sich in diesem Handlungsfeld seit dem Jahre 2000 entwickelt hat. Dieses letzte Beiheft zur Qualitätssicherung ist noch vor dem PISA-Schock, aber nach den TIMSS-Studien erschienen.

Seit dem Jahre 2000 ist keine evolutionäre Entwicklung eingetreten. Es hat sich vielmehr eine Revolution vollzogen, die als Sattelzeit in die Geschichte des Bildungswesens eingehen wird. Auf der bildungspolitischen und administrativen Ebene hat sich außerordentlich viel bewegt. Um nur einige Beispiele zu nennen:

- Im Anschluss an den ersten PISA-Bericht sind die berühmten sieben *Handlungsfelder* von der Kultusministerkonferenz beschlossen worden.
- Die Bundesländer haben systematisch versucht, die *Qualitätssicherungsmaßnahmen* zu eruieren und zu vergleichen.
- Fast alle Bundesländer sind auf externe Leistungsprüfungen in der Gestalt des *Zentralabiturs* umgestiegen.
- Klassenübergreifend und schulextern konzipierte *Vergleichsarbeiten* zur Überprüfung des Leistungsstandards sind in den meisten Bundesländern üblich.
- Die Entwicklung von neuen, kompetenz- und standardbasierten *Bildungsplänen* ist in einigen Bundesländern bereits abgeschlossen.
- Einige Bundesländer haben die interne und externe *Schulevaluation* im Schulgesetz verankert.
- Die Stärkung der *Autonomie* der Schulen, insbesondere des Leitungsauftrages an Schulen, ist weit fortgeschritten. Die Schulleitungen haben weitgehend Aufgaben des Personalmanagements und der Schulaufsicht wahrgenommen.
- Durch die Einrichtung einer *Agentur der Qualitätssicherung* wird die Entwicklung von Testaufgaben auf eine breite professionelle Basis gestellt.
- Die Intensivierung der *Frühförderung* im Kindergarten, die Diagnose und Förderung der deutschen Sprachfähigkeiten und die Anbindung an die Grundschule sind vom Programm- ins Entwicklungsstadium getreten
- Die Lehrerbildung wird an vielen Stellen neu konzipiert und ebenfalls standard- bzw. kompetenztheoretisch restrukturiert.

Nicht weniger atemberaubend sind die Arbeiten auf der Forschungs- und Entwicklungsebene. Durch die Beteiligung an den internationalen Vergleichsstudien sind qualitativ hochwertige international und intranational ausgerichtete Berichte entstanden. Dank der Initiative der deutschen Konsortien und der nationalen Forschungsinstitute (MPI, IPN, DIPF), ist es gelungen, zwei Schwerpunktprogramme bei der DFG einzurichten. Über das erste, „Bildungsqualität von Schule“, liegen bereits zwei zusammenfassende Auswertungsberichte vor. Hier wird das gesamte Spektrum der Möglichkeiten, Bildungs- und Lernprozesse zu verbessern, bearbeitet (Doll/Prenzel 2004; Prenzel/Allolio-Näcke 2006). Der zweite Schwerpunkt zur „Kompetenzmodellierung“ hat 2007 begonnen.

Daneben gibt es intensive Bemühungen zur *Förderung der empirischen Bildungsforschung*.

Die Erziehungswissenschaft hat aber auch wesentlich zur Entwicklung von Konzepten und Programmen der Evaluation beigetragen. Meilensteine sind dabei das sogenannte „Klieme-Gutachten“ zu den Bildungsstandards (Klieme et.al. 2003), der Aufbau eines Konsortiums zur nationalen Bildungsberichterstattung und die Einrichtung eines institutionellen Ortes der Entwicklung von Standards und Testaufgaben (IQ – Institut für Qualitätsentwicklung).

Die schwindelerregende Fülle von Aktivitäten weckt das Bedürfnis nach ordnenden Konzepten. Ist die neue Phase der Entwicklung der deutschen Bildungssysteme auch begleitet von neuen *Einsichten in die Funktionsweise der deutschen Bildungssysteme* und von einem *klaren Qualitätsbewusstsein*, in welche Richtung sich die Bildungssysteme entwickeln sollten, geleitet?

Im Folgenden soll illustriert werden, dass Qualitätsdimensionen *im Rahmen theoretischer Konzeptionen des Bildungswesens* präziser gefasst werden können. In unterschiedlichen Schultheorien kommen auch unterschiedliche Qualitätsdimensionen zum Vorschein. Am Beispiel von zwei Phasen theoretischer Paradigmen in den letzten Jahrzehnten soll dies gezeigt werden (Fend 2006b). Beide Entwicklungen, jene der theoretischen Konzepte zum besseren *Verstehen* des Bildungswesens und jene der *ordnenden Qualitätskriterien*, sollen dabei bilanziert werden. Schon im erwähnten Beiheft hat Ditton auf die Theorieabstinenz hingewiesen, welche die Qualitätsforschung belastet. Was hier helfen könnte, sei im Folgenden kurz skizziert (Ditton 2000).

## **2. Theoretische Modelle des Bildungswesens als Ordnungshilfen für die Einordnung von Qualitätsdimensionen**

Je vielfältiger die einzelnen Maßnahmen sind, mit deren Hilfe das Bildungswesen positiv gestaltet werden soll, umso bedeutsamer wird ein klares Verständnis dessen, was Schule eigentlich ist und tut. Welche *theoretischen Innovation* (Fend 2006b) zu einem solchen Verständnis beitragen könnten, soll hier eingangs expliziert werden.

Ein Land, das in der Moderne für alle Kinder eines bestimmten Alters Lernprozesse über mindestens neun Jahre veranstaltet, dafür Milliarden ausgibt und Tausende von

Lehrpersonen beschäftigt, schafft damit eine der größten sozialen Organisationen, die wir kennen (s. zur historischen Rekonstruktion Fend 2006a). Tagtäglich sind in Deutschland ca. 670.000 Lehrpersonen damit beschäftigt, den Masterplan der schulischen Bildung mit den Lernmöglichkeiten von mehreren Millionen heranwachsender Kinder zu synchronisieren. Wenn Lehrpersonen dies ein volles 40-jähriges Berufsleben lang tun, dann investieren sie dafür ca. 32.000 Unterrichtsstunden.

Wie kann man ein solches Unternehmen verstehen?

Waren dafür in den 60er und 70er Jahren struktur-funktionalistische Modelle adaptiert worden, die zu einer bis heute wichtigen Funktionsanalyse und Strukturdarstellung des Bildungssystems geführt haben, so stehen heute Konzepte im Vordergrund, welche die handlungstheoretische Grundlagen der Schultheorie betonen. Strukturen lehren Kinder nicht und Funktionen schaffen noch keine realen Beziehungen (Fend 2006b).

Die neue handlungstheoretische, institutionsorientierte und akteursensitive Konzeption des Bildungswesens (Fend 2006b) blendet Strukturen jedoch nicht aus, sondern integriert sie in der Gestalt institutioneller Regelzusammenhänge. Handeln von Akteuren in Institutionen ist danach regelgeleitetes Zusammenhandeln. Im Bildungswesen ist es auf die Erfüllung der zentralen Aufgabe ausgerichtet, die das Bildungswesen leitet. Sie besteht darin, ein großes *kulturelles Inhaltsprogramm* an die *kognitiven und emotionalen Strukturen* heranwachsender Kinder und Jugendlicher *anschlussfähig* zu machen, letzteren also zu einem kulturgeleiteten Lern- und Aufbauprozess von Denken, Handeln und Empfinden zu verhelfen. Dieses komplexe Unternehmen wird auf *mehreren Handlungsebenen* realisiert und gestaltet; auf der Ebene von Bildungspolitik und Verwaltung, der einzelnen Schule und des Unterrichts von Lehrpersonen.

Die Produktionsfunktion des Bildungswesens muss also handlungstheoretisch rekonstruiert werden, um sie aus der abstrakten Konzeption von Input und Output, die als Schema den OECD Studien zugrunde liegt, herauszuführen. Für ländervergleichende Darstellungen mag dieser Abstraktionsgrad notwendig sein, für länderspezifische Konzepte der Gestaltung des Bildungswesens bedürfen wir stärker handlungstheoretisch ausgearbeiteter Modelle.

Eine erste handlungstheoretische Präzisierung bietet das *Angebots- Nutzungsmodell*. Es macht in einem ersten Schritt auf den Sachverhalt aufmerksam, dass an den Lernergebnisse zwei Seiten beteiligt sind: Akteure, die ein Angebot gestalten und Akteure, die dies über Lern- und Rezeptionsprozesse zu ihrem geistigen Habitus und Kompetenzprofil umformen.

Wie kann auf den *verschiedenen Handlungsebenen* des Bildungswesens die Erfüllung des zentralen Auftrages der Schule ausbuchstabiert werden? Dazu bedarf es eines Konzeptes, das die verschiedenen Handlungsebenen verbindet, wie dies in den neueren Ansätzen des Educational Governance versucht wird (Altrichter/Brüsemeister/Wissinger, 2007; Kussau/Brüsemeister 2007). Mit jenem der *Rekontextualisierung* erscheint dies möglich. Es geht davon aus, dass der Bildungsauftrag des institutionellen Akteurs auf jeder Handlungsebene vorgegeben wird und an die jeweils ebenenspezifischen Bedingungen der Auftrags Erfüllung angepasst, umgeformt und kreativ gestaltet wird. Auf allen folgenden Ebenen kann unterschiedlich gut gehandelt werden.

### 1. Ebene: Die Konstruktion kultureller Traditionen und von Zukunftsszenarien

Bildungssysteme sollen die Kultur eines Gemeinwesens übertragen, sie folgen also einem Inhaltsprogramm. Sie können dies nur im Nachgang zur Entstehung einer Kultur tun, also sind sie gegenüber dem Primat des Kulturschaffens sekundäre Institutionen. Es bedarf der „Weimarer Hochklassik“, bevor in der Schule Goethe „durchgenommen“ werden kann. Eine genauere Betrachtung enthüllt aber, dass hinter der Definition von „Kultur“ Konstruktionsleistungen stehen. Wenn in der Schule die „wertvolle Kultur“ an die nachwachsende Generation weitergegeben werden soll, dann stellt sich die Notwendigkeit zu definieren, was die „wertvolle Kultur“ ist. Sie ergibt sich nicht „von selbst“ (s. dazu auch Bourdieu 1983; Foucault 1976).

### 2. Ebene: Selektion und Organisation von Inhalten für die schulische Übermittlung

Fragen der Traditionsbildung und der Auszeichnung von wichtigen kulturellen Inhalten und Kompetenzen entzünden sich notwendigerweise an der Frage der Auswahl von Bildungsinhalten für deren Weitergabe in Schulen. Da sich heute niemand mehr die gesamte Kultur aneignen kann, stellt sich die Frage der Selektion und der Organisation dieses Wissens für einen langjährigen Lernprozess. „Kulturwissen“ muss zu „Schulwissen“ transformiert werden. Der Prozess der Selektion und Organisation von Lerninhalten ist heute hochgradig ausdifferenziert und selbst noch in mehrere Ebenen gegliedert: in die Ebene der politisch-administrativen Entscheidungen, in die Ebene der *Lehrplankommission*, der *Lehrmittelproduktion* und im Rahmen neuer Lehrpläne sogar in die Ebene des *Schulcurriculum*.

### 3. Ebene: Rekontextualisierung auf der Unterrichtsebene – die Umsetzung durch professionelle Akteure

Auf Unterrichtsebene müssen die „operativen Akteure“ des Bildungswesens die inhaltlichen und rechtlichen Vorgaben auf einen neuen situativen Kontext von Unterrichten und Lernen adaptieren. Lehrpersonen müssen, in der Alltagssprache formuliert, die Prüfungsanforderungen, die Lehrpläne und Lehrziele „umsetzen“. Institutionelle Vorgaben bilden die eine Umwelt ihres Handelns, die situativen Bedingungen der Schulklasse und die Merkmale der vor ihnen sitzenden Schüler die andere.

Wie transformieren sie dabei die inhaltlichen Vorgaben? Der wichtigste Bezugspunkt der Rekontextualisierung besteht in der Feinanalyse der zu vermittelnden Inhalte, also in der Aufgabenanalyse, die über Jahrespläne, Wochenpläne, Tagespläne und Stundenpläne schließlich so detailliert erfolgt, dass daraus Erarbeitungsschritte von komplexen Inhalten in *Lektionen* resultieren.

Ein zweiter Rekontextualisierungsschritt wird durch die Lernbedingungen und Lernmöglichkeiten der vor den Lehrpersonen sitzenden Schülerinnen und Schüler vorgegeben.

Die Lehrpersonen als „operative Akteure“ müssen dabei die vorgegebenen Lerninhalte auf die Bedingungen ihrer Lernbarkeit in einem begrenzten Zeitraum abstimmen. Diese Lernbarkeit wird durch die kognitiven und motivationalen Voraussetzungen der Schüler mitbestimmt. In jahrzehntelangen Abstimmungsprozessen hat sich im Bildungswesen so etwas wie ein „Wissen im System“ darüber herausgebildet, was Schüler in welchem Alter lernen können.

Um diese anspruchsvollen Rekontextualisierungen vornehmen zu können, werden die operativen Akteure, also die Lehrpersonen, heute in mehreren Ausbildungsjahren trainiert. Sie können dabei die historischen Erfindungen zur Unterrichtsmethodik kennen lernen und sich in ihren Gebrauch einüben.

#### *4. Ebene: Schülerschaft und Eltern als Nutzungsakteure*

Die *Schülerschaft* als *Nutzungsakteur* übernimmt das schulische Angebot nicht spiegelbildlich. Schülerinnen und Schüler legen eigene Maßstäbe an, sie wählen aus, was ihnen wichtig und unwichtig ist. Vieles, was die Schule für zentral hält, bleibt ihnen fremd. Mit den hier entwickelten Begriffen heißt dies: Sie „rekontextualisieren“ das schulische Angebot nach neuen Gesichtspunkten. Das Lehrangebot der Schule, vermittelt über die Anforderungen der Lehrperson, ist für sie die eine Umweltgröße. Die Erwartungen der Eltern und Mitschüler bilden die andere. Alle diese Erwartungen und Handlungsbedingungen werden durch die Nutzungsakteure mit der „Innenseite“, mit ihren Bedürfnissen, Wünschen und Motivationen sowie mit den Lernkapazitäten „abgeglichen“. Zu dieser Innenseite gehören also die realen und wahrgenommenen Leistungsressourcen und die realen und wahrgenommenen Motivationen. Sie führen zu einer erneuten Rekontextualisierung des schulischen Angebotes durch die Nutzungsakteure, durch Schülerinnen und Schüler.

Das Konzept der „Rekontextualisierung“ soll darauf verweisen, dass die Adressaten schulischen Lernens sich nicht mit einer unstrukturierten Umwelt auseinandersetzen, sondern mit einer hochgradig gestalteten schulischen Lehre. Dieser Kontext wird aber neu interpretiert, wird von der Schülerschaft in einer eigenständigen Leistung zu „ihrer Welt“ angeeignet. Ihre Reaktionen, ihre Ziele und „beliefs“, ihre Freuden und Ängste sind danach jeweils der Niederschlag davon, wie sie Schule rekontextualisiert haben. Bei dieser Rekontextualisierung sind sie aber nicht einsame und autonome „Nutzer“. Sie stehen vielmehr in einem kommunikativen Austausch mit Mitschülern und – vor allem – mit Eltern. Gerade letztere versuchen aktiv, eine günstige Rekontextualisierung zu fördern, da sie die über schulischen Erfolg führenden Wegweiser zu einem produktiven Erwachsenenleben in der Moderne nur zu gut kennen. Aber auch Mitschüler definieren die schulische Situation mit: ob es prestigehaltig ist, schulisch gut zu sein, sich den Anforderungen der Lehrer möglichst zu entziehen, oder sich ihnen sogar entgegenzustellen.

### Resümé: Educational Governance

Auf jeder Handlungsebene – auf jener der Politik, der Verwaltung, der Schule, der Lehrpersonen und der Schülerschaft – sind die Vorgaben auf der jeweils übergeordneten Ebene Umwelten des Handelns, die eine Umsetzung an die jeweils neuen, ebenenspezifischen Besonderheiten der Umwelten des Handelns erfordern (s. Abb. 1).

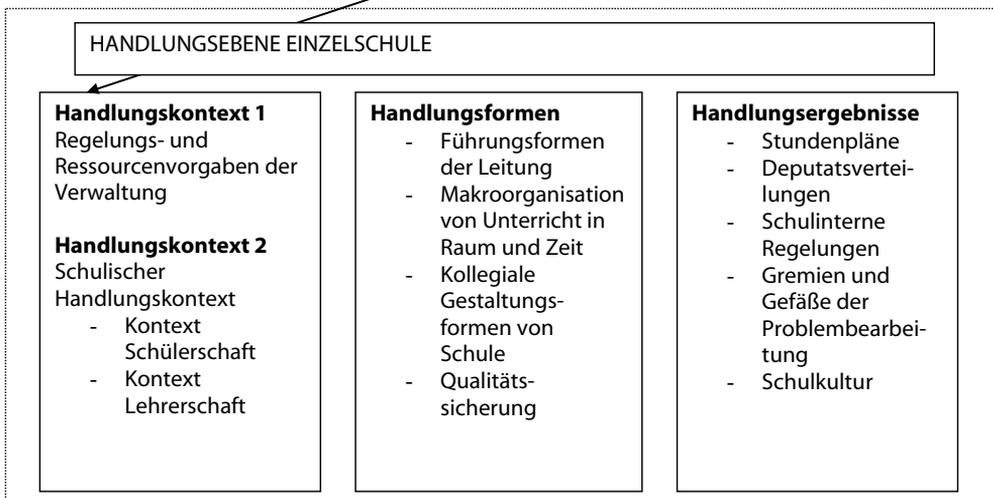
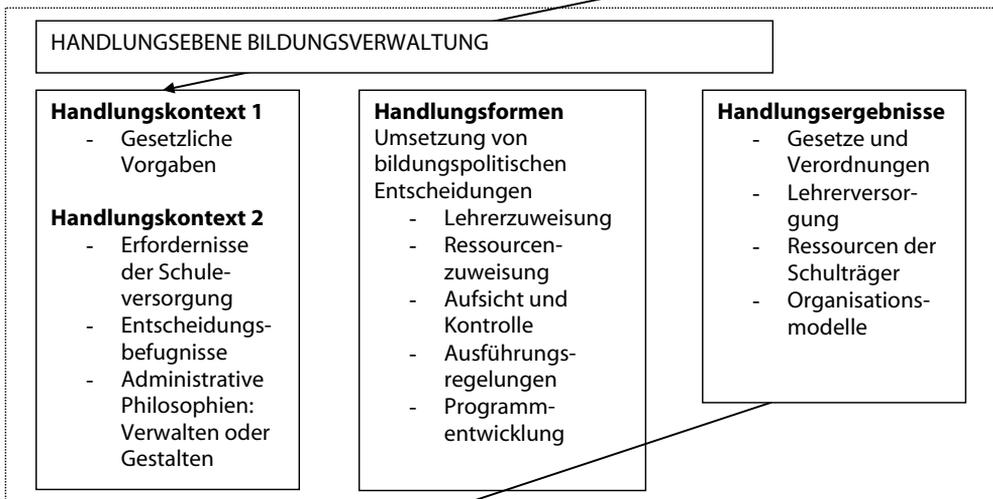
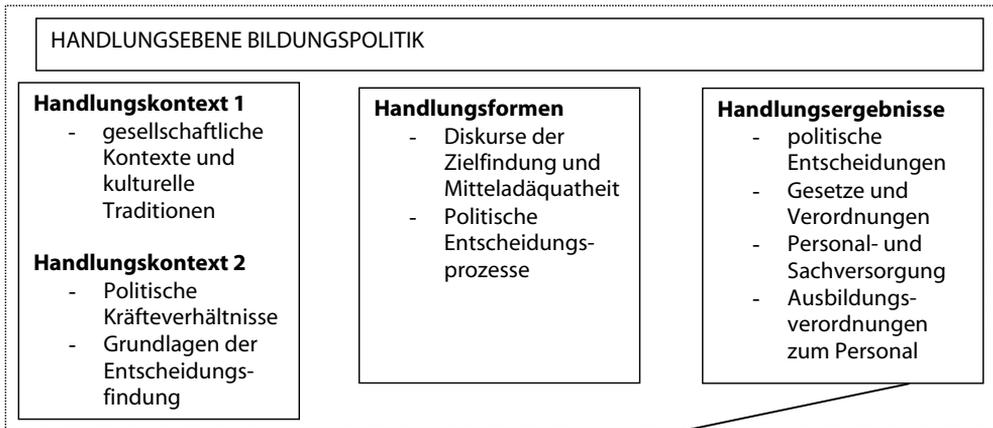
Auf jeder Ebene finden wir dabei Kontexte, denen gegenüber Akteure handeln. Der erste Kontext wird immer durch *übergeordnete* Rahmenbedingungen vorgegeben, die es zu interpretieren und umzusetzen gilt (Handlungskontext 1). Die Akteure handeln aber auch im jeweiligen Feld und unter *ebenenspezifischen Handlungsbedingungen* (Handlungskontext 2). Politische Akteure etwa im Rahmen politischer Machtverhältnisse und institutioneller Verfahren der Entscheidungsfindung.

Auf jeder Gestaltungsebene arbeiten am Bildungswesen Akteure, in dem sie ihre je ebenenspezifischen Aufgaben erfüllen und Handlungsergebnisse erzeugen, die für die jeweils „darunter“ liegende Handlungsebene wieder einen Kontext des Handelns repräsentieren. Für die einzelne Schule sind die politischen und administrativen Regelungen solche Kontexte. Die Schule agiert aber wiederum auch einem zweiten Kontext gegenüber, der sich auf der Schulebene etwa aus den Merkmalen der lokalen und schulspezifischen Schülerschaft bzw. aus Merkmalen des Kollegiums ergibt. So blickt jede Ebene nach „oben“ und nach „unten“, bzw. auf die ebenenspezifischen Aufgaben.

Die jeweilige Reaktion dem Handlungskontext 1 und dem Handlungskontext 2 gegenüber erfolgt nicht „automatisch“, sie ist vielmehr geleitet von den Wahrnehmungen, Interpretationen und Kompetenzen der jeweiligen Akteure. Es bedarf somit auch handlungstheoretischer Modelle der Bewältigung von Schule auf der jeweiligen Ebene.

Auch wenn ein gemeinsames Ziel vorgegeben ist – die „Menschenbildung“ im Sinne der „Gestaltung“ der heranwachsenden Generation –, so optimiert jede Ebene ihre eigenen Handlungsmöglichkeiten und folgt damit einer ebenenspezifischen „Belohnungslogik“:

- Die Bildungspolitik optimiert z.B. die rechtlichen Rahmenbedingungen, den Ressourceneinsatz und die Durchsetzbarkeit ihrer Richtlinien auf dem Hintergrund der politischen Akzeptanz ihrer Politik in der Bevölkerung.
- Auf Schulebene optimiert z.B. ein Kollegium die Sicherung der Arbeitsbedingungen, etwa die Solidarität der Kollegen und die Autonomie des Handelns im Klassenzimmer.
- Auf der konkreten Unterrichtsebene optimieren Lehrpersonen z.B. die Bedingungen, unter denen Klassenunterricht möglichst ungestört und effektiv durchgeführt werden kann und möglichst unangreifbare Leistungsbeurteilungen vergeben werden können.
- Auch die Schülerinnen und Schüler optimieren ihre Situation: Sie investieren so in Lernanstrengungen, dass optimale Lernergebnisse resultieren und negative Sanktionen von verschiedenen Seiten, von der Seite der Eltern, der Lehrer und der Mitschüler minimiert und positive maximiert werden.



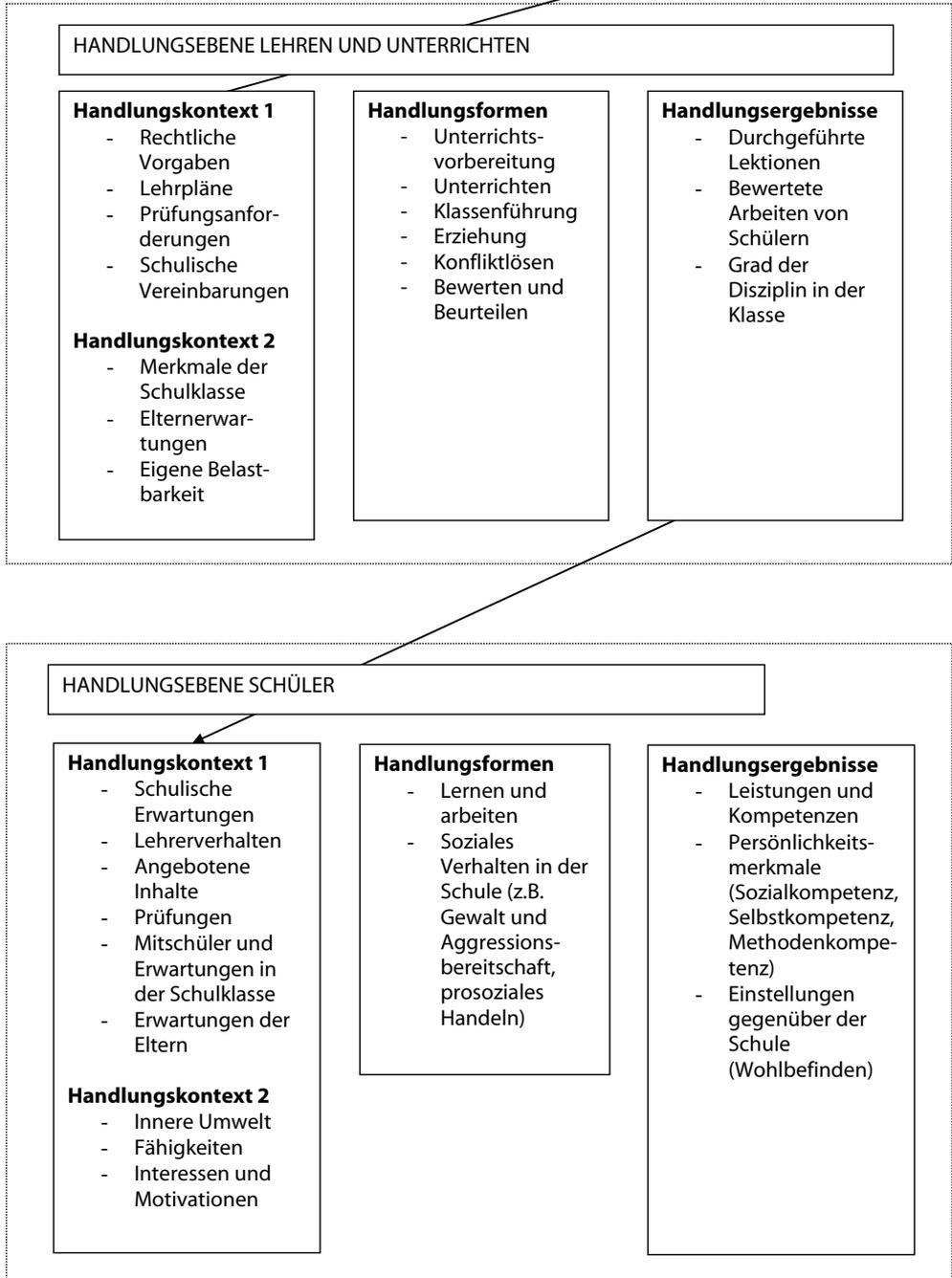


Abb. 1: Rekontextualisierung und Handlungsebenen

Diese Optimierungsprozesse sind Ausdruck rationaler Aufgabenbewältigung und Ausdruck der Interessen, die auf der jeweiligen Handlungsebene im Mittelpunkt stehen. Bei diesen Optimierungen werden aber jeweils die institutionellen Regeln ausbuchstabiert bzw. bei der Aufgabenerfüllung angewendet. Auch wenn auf der jeweiligen Ebene individuelle Interessen wirksam sind, so werden die jeweiligen Optimierungsbemühungen durch die Präsenz der institutionellen Vorgaben auch zu regelgesteuertem Zusammenwirken.

### 3. Ordnungshilfen zum Qualitätsbewusstsein

Es ist naheliegend, dass an alle Handlungsfelder der Gestaltung eines schulischen Lernangebotes und seiner bestmöglichen Nutzung *Qualitätskriterien* angelegt werden können. Die handlungstheoretische Konzeption legt nahe, an allen Stellen zu schauen, wie gut die Akteure ihre Arbeit machen.

Was ist aber „gut“? Hier ist Heid in seinem Beitrag zum 41. Beiheft der Zeitschrift für Pädagogik zuzustimmen, dass dies ein Attribut ist, das sich nicht notwendigerweise aus Merkmalen der „Sache“ ergibt, sondern das Ergebnis eines evaluativen Zusatzes, eines Auszeichnungsprozesses ist, der anderen Regeln und Prämissen folgt als die Feststellung eines Sachverhaltes (2000). Folgt letzterer der Logik wissenschaftlicher Forschung, so folgt die Qualitätsfeststellung der Logik von sozialen Herstellungs- und demokratischen Einigungsprozessen.

Wie kommt man dabei auf „Qualität“? Die Grunderfahrung ist wohl die, dass man etwas unterschiedlich „machen“ kann. Dies setzt wiederum viele Vergleichsmöglichkeiten voraus. Die unmittelbare Anschauung kann aber durch einen systematischen Suchprozess ergänzt werden, der auch wissenschaftliche Studien über Wirkungszusammenhänge im Bildungswesen einschließt.

Wichtig erscheint mir die Erfahrung, dass sich Qualitätsbewusstsein nicht von selbst ergibt. Es ist vielmehr das *Ergebnis eines langen, eigenständigen Erfahrungs- und Suchprozesses*.

In einer ersten Suchphase, die bis zum Beginn der großen internationalen Schulleistungsstudien reichte, haben sich Qualitätsdimensionen herauskristallisiert, die eine *Zustandsbeschreibung* des Bildungswesens im Auge hatten. Sie waren von Indikatoren geleitet, die den Input und vor allem den Output beschreiben sollten. Das implizite Denken war *wirkungsorientiert* und insbesondere an kognitiven Lernergebnissen ausgerichtet. Die in Abb. 2 beispielhaft aufgeführten Qualitätsdimensionen sollen dies dokumentieren.

Im Kriterienkatalog der Abb. 2 stehen unübersehbar *Wirkungsindikatoren* im Vordergrund. Sie dienen als Anhaltspunkte für eine systematische Messung der Qualität im Bildungswesen auf verschiedenen Ebenen (s. z.B. Fend 2001). *Theoretisch* sind sie mit *Modellen des Bildungswesens* verbunden, die ebenfalls *wirkungsorientiert* ausgerichtet sind, die Input, Strukturen, und Outputs im Auge haben. Strukturelle Beschreibungen des Bildungswesens stehen im Hintergrund dieser Qualitätsdimensionen.

SCHULSYSTEM:	Kognitive Lernergebnisse Persönlichkeitswirkungen (fachübergreifende Wirkungen) Chancengleichheit Flexibilität der Schullaufbahngestaltung Versorgungsdichte Ausstattung				
SCHULEBENE:	Leistungsniveau der Schule Pädagogisches Ethos des Kollegiums Arbeitsethos des Kollegiums Verhaltensweisen der Schülerschaft: <ul style="list-style-type: none"> <li>– Vandalismus und Aggression oder prosozialer Verhaltensstil</li> <li>– Schuldistanz oder Vertrauen der Schülerschaft</li> </ul>				
KLASSENEBENE:	Leistungsniveau der Klasse Motivation und Leistungsbereitschaft Problembelastung: Disziplinprobleme, Rauchen, Alkoholkonsum, Delinquenzbelastung Qualität der sozialen Beziehungen der Schüler untereinander: <ul style="list-style-type: none"> <li>– Konformitätszwang</li> <li>– Rivalitätsgrad und Ostrazierungstendenzen</li> <li>– Isolation oder Vercliquung</li> </ul>				
PERSONEBENE:	<table style="width: 100%; border: none;"> <tr> <td style="width: 100px; vertical-align: top;">Lehrer:</td> <td>Kompetenzen Pädagogische Einstellungen Burnout Hilflosigkeit</td> </tr> <tr> <td style="vertical-align: top;">Schüler:</td> <td>Wissen und Kompetenzen Leistungsbereitschaft und Arbeitshaltungen Zuverlässigkeit und Pflichtbewußtsein Ich-Stärke Soziale Integrationsfähigkeit und Führungsfähigkeit Verantwortungsbereitschaft</td> </tr> </table>	Lehrer:	Kompetenzen Pädagogische Einstellungen Burnout Hilflosigkeit	Schüler:	Wissen und Kompetenzen Leistungsbereitschaft und Arbeitshaltungen Zuverlässigkeit und Pflichtbewußtsein Ich-Stärke Soziale Integrationsfähigkeit und Führungsfähigkeit Verantwortungsbereitschaft
Lehrer:	Kompetenzen Pädagogische Einstellungen Burnout Hilflosigkeit				
Schüler:	Wissen und Kompetenzen Leistungsbereitschaft und Arbeitshaltungen Zuverlässigkeit und Pflichtbewußtsein Ich-Stärke Soziale Integrationsfähigkeit und Führungsfähigkeit Verantwortungsbereitschaft				

Abb.2: Qualitätsdimensionen

### Von Produktkriterien zu Prozesskriterien

Verändert nun eine stärker handlungs- und akteurbezogene Konzeptualisierung des Bildungswesens den Kriterienkatalog? Ermöglicht sie neue Suchprozesse für Qualitätsdimensionen? Dies ist unübersehbar der Fall. Es kommen nämlich zusätzlich Qualitätsmerkmale von *Prozessen* und von *Instrumenten* der Schulgestaltung ins Blickfeld. Sie können aber weiterhin auf verschiedenen Ebenen angesiedelt werden.

So stellt sich jetzt die Frage nach der Qualität der Systemsteuerung, oder nach der Qualität des *Educational Governance*. Die *Kerninstrumente* sind hier die Gestaltung des *Masterplanes* der Bildung, die zugeordneten Schullaufbahnen und *Prüfungssysteme*, die

Verfahren zur *Allokation von Ressourcen*, die *rechtlichen Verfahren von Verantwortungszuschreibung* und vor allem die *Strategien der Qualitätssicherung*. Die Qualitätsdiskussion richtet sich entsprechend auf die Qualität dieser Verfahren.

An einigen Beispielen soll dies illustriert werden. Was ist z.B. ein „guter Masterplan“ der Bildung?

### *Qualitätsdimensionen von Bildungsplänen*

Der „Masterplan“ dazu, was eine Schule an Lerninhalten anbietet, ist eine sehr komplexe „Partitur“ dessen (s. für das umfassendste Forschungsprogramm Künzli et al. 1999; Künzli/Hopmann, 1998), was in der Schule „aufgeführt“ werden soll. Er enthält heute

- Leitziele,
- Pädagogische Konzepte der Schule,
- Lernzielvorgaben für Fächer,
- Inhalte für Fächer und Fächergruppen,
- fachübergreifende Ziele und Inhalte,
- allgemeindidaktische und fachdidaktische Verfahren,
- Planungsvorgaben für die Schuljahre,
- zugeordnete Lehrmittel,
- Prüfungsformen,
- Niveauvorgaben,
- Standards,
- Testaufgaben.

Blickt man auf die Geschichte dieser inhaltlichen Steuerung des Bildungswesens zurück, die heute durch die Konzentration auf *rechtliche Verfahren* nicht selten vernachlässigt wird, dann stellen sich z.B. folgende Qualitätsfragen:

- Gelingt es dem Masterplan, eine *stimmige Rekonstruktion der kulturellen Traditionen* vorzunehmen, die es allen ermöglicht, in der eigenen Kultur „heimisch“ zu werden?
- Gibt es eine *einsichtige Verbindung* zwischen der *Konstruktion kultureller Traditionen* und der *Ausrichtung auf Kompetenzen und Qualifikationen*, die für die Lebensbewältigung notwendig sein werden?
- Enthält der Lehrplan ein stimmiges Ganzes, das so etwas wie ein *Kerncurriculum* repräsentiert?
- Gibt es Bemühungen, Profile von Lernangeboten zu formulieren, die auf die *Heterogenität der Schülerschaft* und die *unterschiedlichen zukünftigen Lebensverhältnisse* eingehen?
- Sind die formulierten Inhalte und Standards auf die *faktischen Lernmöglichkeiten* von Kindern und Jugendlichen abgestimmt?
- Ermöglicht der Bildungsplan, genauer zu sehen, welche Kompetenzen Schülergruppen *erreicht haben sollen*, wenn sie einen Bildungsgang durchlaufen?

- Sind die Masterpläne in ihrer Langzeitperspektive mit verschiedenen Ländern *harmonisiert*, so dass die Mobilität der Bevölkerung nicht beeinträchtigt ist?
- Sind Bildungspläne, Lehrbücher und Aufgabensammlungen so abgestimmt, dass eine stimmige und wirksame inhaltliche Steuerung möglich ist, die sowohl Vielfalt ermöglicht als auch Gemeinsamkeiten stiftet?
- Enthalten sie *nachvollziehbare* und *erprobte* Aufgabensammlungen?

Unübersehbar enthalten diese Qualitätsmerkmale von Bildungsplänen Konzepte, wie über Bildungspläne guter Unterricht vor Ort gesteuert werden kann. Bildungspläne dienen aber auch übergeordneten Aufgaben des Bildungswesens, wie der Schaffung kultureller Gemeinsamkeiten, der sinnvollen Gestaltung von Langzeitbildungsgängen und der Ermöglichung von Mobilität. Dabei stellt sich die Frage, was auf welcher Handlungsebene entschieden werden sollte, wo z.B. Autonomie der Schule und des Lehrers angebracht ist und wo nicht.

### *Qualitätsdimensionen der Qualitätssicherung auf Systemebene*

Noch präziser als im Jahre 2000 wird heute wahrgenommen, worin die traditionelle Gestalt der Qualitätskontrolle im deutschen Bildungswesen bestand und teilweise immer noch besteht.

Das deutsche Bildungswesen weist nur eine sehr dünne Fachaufsicht aus, Regelbesuche erfolgen oft nur alle vier Jahre, meist gibt es aber nur anlassbezogene Dienstaufsichtbesuche und entsprechende Beurteilungen. Die Aufsicht ist zudem intern reguliert, das System überprüft sich gewissermaßen immer selber, und die Ergebnisse werden nicht im Sinne von lernprozessförderlichen Rückmeldungen an die operativen Akteure genutzt. Ein solches Bildungswesen hat die Tendenz zur Schließung, zum Ausschluss aller nicht im System arbeitenden Akteure. Außenevaluationen im Sinne klassen- bzw. schulübergreifend vergleichender Evaluationen sind schwach ausgeprägt. Diese politische Organisation von Aufsicht und Evaluation ist gleichermaßen bedeutsam und politische schwer veränderbar.

Auf diesem Hintergrund werden fehlende Qualitätsdimensionen sichtbar, die mit folgenden Fragen angedeutet werden können:

- Wie gut weiß ein Bildungswesen über sich selber Bescheid?
- Wie wird auf allen Ebenen des Bildungswesens eine realistische Evaluation sicher gestellt?

Die Generierung von Wissen ist jedoch nur ein Anfang, um zu Qualität in der Qualitätssicherung zu kommen. Die Probleme hat schon Stryck im Beiheft des Jahres 2000 differenziert beschrieben. Die Informationen können schlicht folgenlos „versickern“, oder sie werden systematisch abgewehrt. Wenn Qualitätsinformationen zur Qualitätsverbesserung genutzt werden sollen, dann bedarf dies eines wohlüberlegten Vorgehens,

das Stryck im Idealfall in einer *systemischen Kooperationsstrategie* sieht. Sie muss darauf ausgerichtet sein, eine wichtige Qualitätsdimension im Bildungswesen zu optimieren: das *Vertrauen* unter allen Beteiligten, dass sie am gleichen Ziel arbeiten und sich gegenseitig unterstützen wollen. Die Schulaufsicht als Instrument der Qualitätssicherung müsste danach eine professionelle Entwicklungsagentur werden, die „... mit den Schulen Entwicklungsverträge abschließt und

- die notwendigen Entwicklungsdienstleistungen garantiert,
- systematische Erfolgskontrollen durchführt und
- Qualitätsunterschiede folgenreich thematisiert ...“ (Stryck 2000, S. 124).

Solche Verträge können auch durch entsprechende *gesetzliche Verpflichtungen zur Selbst- und Fremdevaluation* ersetzt werden. Nicht ersetzbar wird dabei die Qualitätsdimension des *Vertrauens* zwischen den verantwortlichen Akteuren. Sie wird hier aber erweitert durch die Dimension der *Legitimität*. Wer soll und darf Bildungssysteme, Schulen und Lehrpersonen evaluieren? Wer hat dazu die *Kompetenz* und wer hat das *Recht*?

Der eine Zweig der *Legitimation* ergibt sich durch das Verfahren der Verantwortungszuschreibung. Staatliche Instanzen gehören per Gesetz dazu. Es könnten aber auch gewählte Vertreter der Öffentlichkeit sein, wie dies z.B. bei den Schulgemeinden in der Schweiz der Fall ist. Schwierig wird es, wenn private Instanzen, etwa auf dem Wege des Outsourcing die Evaluation übernehmen sollen.

Die Frage der *Kompetenz* folgt auf dem Fuß. Prinzipiell gilt hier, dass das Messinstrument so gut sein muss, dass es Qualitätsmerkmale auch registrieren kann.

Die Organisation von Informationsströmen und der Nutzung von Informationen wird hier zu einem Kernbereich schulischer Qualitätsdimensionen. Man kann den Prozess der Qualitätssicherung unterschiedlich gut organisieren. Institutionelle Regelungen, Akteure und Instrumente müssen hier eine sinnvolle Konfiguration bilden. Die Qualitätsbereiche wären daneben Offenheit der Informationen bzw. gerechtfertigte Einschränkungen, Legitimität der Verfahren und Kompetenzen der Durchführung. Sie konstituieren einen wichtigen Ausschnitt einer guten „Educational Governance“.

### *Qualitätsdimensionen der Rekontextualisierung auf Schulebene*

Die Entdeckung von neuen Qualitätsdimensionen springt in keinem anderen Bereich so ins Auge wie auf der Schulebene. Hier sind es insbesondere die sogenannten *Referenzrahmen* verschiedener Länder (z.B. Hessen, Niedersachsen), die explizit formulieren, was Qualität in den Augen der für die Systemsteuerung Verantwortlichen heißt. Dabei stehen mehr oder weniger explizit handlungstheoretischer Konzepte Pate.

So stehen am Beginn der Kriterienkataloge für die Selbst- und Fremdevaluation *Umgangsweisen mit den Vorgaben der Ministerien* (s. Abb. 3), etwa bei der Formulierung von Evaluationsfragen wie:

- Wie setzt die Schule die Bildungsstandards um?
- Wie geht die Schule mit den Kontingenzstundentafeln um?
- Hat die Schule ein Schulprogramm entwickelt?
- Hat die Schule aus dem Schulprogramm stringent ein Schulcurriculum abgeleitet?
- Wie geht die Schule mit den schulübergreifenden Rückmeldungen in der Gestalt von Vergleichsarbeiten um?
- Nutzt sie diese für die interne Qualitätssicherung?

In einem handlungstheoretischen Modell der Aufgabenbewältigung auf Schulebene reicht es nicht aus, nur die *Vorgaben* ins Auge zu fassen. Schulen handeln angesichts sehr unterschiedlicher *lokaler Umwelten*. Dazu zählen zumindest

- die lokalen Schulangebote vor Ort, wie viele Schulen z.B. in Konkurrenz um gute Schüler stehen,
- die lokale Schülerschaft, die einer Schule zugewiesen ist, etwas durch Sprengel in einem großstädtischen sozialen Brennpunkt,
- die externe Unterstützung der Schule durch Politik, Medien und Elternschaft.

Wie mit diesen Vorgaben von „oben“ und von „unten“ umgegangen wird, gilt es in der Evaluation der Einzelschule zu überprüfen, die dabei ins Blickfeld kommenden Evaluationsdimensionen sind in Abb. 3 illustriert.

<b>Baden-Württemberg</b>	<b>Hessen</b>
Vorgaben des Ministeriums (z.B. Bildungsplan, VERA)	Voraussetzungen und Bedingungen
<i>Umgang mit den Vorgaben</i>	
Qualitätsmanagement	Ziele und Strategien der Qualitätsentwicklung
Schulführung und -management	Führung und Management
Außenbeziehungen	
Professionalität der Lehrkräfte	Professionalität
Schul- und Klassenklima	Schulkultur

Abb. 3: Evaluationsdimensionen als Umgangsformen mit den Vorgaben in zwei Bundesländern

Welches sind nun die Prozessqualitäten, die genuin in der Verantwortung der Einzelschule stehen und deshalb Gegenstand der Selbstevaluation und externen Evaluation werden können? Nach dem Schweizer Referenzrahmen sind es die in Abb. 4 wiedergegebenen, hier jedoch gekürzt und leicht abgewandelt aufgeführten Qualitätskriterien. Dabei sind Qualitätsmerkmale jeweils zu Qualitätsbereichen gruppiert, in der tatsächlichen Evaluation aber getrennt zu behandeln. Zu allen Punkten sind vierstufige Ant-

wortmöglichkeiten vorgeschlagen, die unterschiedlich bezeichnet werden können (z.B.: vollständig erfüllt/teilweise erfüllt/kaum erfüllt/überhaupt nicht erfüllt). Qualität zeigt sich danach im Umgang mit Rahmenvorgaben, in der Schulführung und Schulorganisation und in Verfahren des Qualitätsmanagements. Die in Abb. 4 aufgeführten Standards machen sichtbar, dass je nach der Verfassung eines Bildungswesens viele schon auf der *Makroebene* festgelegt sind oder auf der Schulebene als rechtliche Verfahren existieren. Für die Qualität der Einzelschule ist dann vor allem der Umgang mit diesen Regelungen relevant.

Bei den Qualitätsdimensionen in Abb. 4 steht unübersehbar immer im Mittelpunkt, *was Schulen und ihre Akteure tun*, um ein wünschenswertes Ergebnis zu erzielen. Die Prüfung der *Ergebnisse* ist dann wieder ein anderer Evaluationsbereich.

### *Qualitätsdimensionen auf Unterrichtsebene*

Wenn wir Prozessqualitäten auf der Unterrichtsebene ins Blickfeld rücken wollen, dann sind wir wieder darauf verwiesen zu untersuchen, wie Lehrpersonen mit den institutionellen Vorgaben des schulischen Masterplanes, den Rechts- und Prüfungsvorgaben sowie den realen Unterrichtsbedingungen mit einer oft sehr heterogenen Schülerschaft umgehen.

Unterricht ist der Ernstfall. Hier treffen die beiden Welten aufeinander: das kulturelle Lehrprogramm und die kognitiven Entwicklungspotentiale heranwachsender Kinder und Jugendlicher. Die Qualitätsdimensionen im Sinne von Prozessqualitäten liegen dabei auf der Hand. Sie beziehen sich darauf, wie gut die *Synchronisierung* dieser beiden Wirklichkeitsbereiche gelingt. In einer rein institutionsbezogenen und fachbezogenen Lehre wird „Stoff“ schlicht durchgenommen, vorgetragen und geprüft. Seine Verankerung im Denken und den Kompetenzen der Schülerschaft wird als selbstverständlich unterstellt. Ein ausschließlich kindbezogener Unterricht orientiert sich nur an der Originalität und spontanen Erkenntnisfähigkeit von Kindern und belässt diese in den individuellen Besonderheiten ihrer Entdeckungen. Synchronisierung bedeutet „kognitives“ Wachstum an Geformtem, eigene Gestaltwerdung durch Abarbeitung an kulturellen Hochformen.

Unterrichten wird hier zu einem permanenten *Abstimmungsprozess*, *primäre Rekontextualisierung* genannt (s. in Fend 2008), in dem die Schüler „dort abgeholt werden, wo sie stehen“, und den kulturellen Anforderungen des Lehrangebotes. In dieser handlungstheoretischen Sicht kommen wieder *Prozessqualitäten* ins Spiel. Sie sind in den Beurteilungsdimensionen des Unterrichts gespeichert, die bei der Beurteilung von Lehrpersonen angewendet werden. Dies ist klassisches „Wissen im System“, also Erfahrungswissen, das im Laufe der Schulentwicklung entstanden ist und z.B. in unzähligen Artikeln von berufsbezogenen Publikationen diskutiert wurde und wird (s. z.B. Depaepe 2000). Es führt zu Qualitätskriterien, wie sie z.B. in Abb. 5 festgehalten sind. Ähnliche Qualitätsdimensionen finden sich in Bewertungsschemata von Lektionen durch Schulaufsichtsbeamte in vielen deutschsprachigen Ländern.

**Umgang mit Rahmenvorgaben**

- An der Schule gibt es ein Schulleitbild und Qualitätsleitbild, das gemeinsam erarbeitet wurde. Die Schule hat Vorstellungen entwickelt, was sie will und was ihr Qualitätsanspruch ist.
- Die Schule hat aus den Lehrplanvorgaben ein schulspezifisches Curriculum erarbeitet.
- Die Führung der Schule achtet auf das vorgeschriebene und vereinbarte curriculare Lehrprogramm und beobachtet dessen Einhaltung.

**Verfahren: Qualitätsmanagement**

- Die Schule hat ein Qualitätsleitbild in einem partizipativen Verfahren entwickelt.
- Es sind Verfahren institutionalisiert, die die Informationsbasis für das Qualitätsmanagement schaffen.
- Feed-back zum eigenen Handeln gilt als Teil der Professionskultur.
- Für die Selbstevaluation bestehen die institutionellen Verfahren und personellen Ressourcen, um sie ohne Überforderung durchzuführen.
- Evaluationsergebnisse werden in Schulentwicklungsmaßnahmen umgesetzt.

**Verantwortliche Akteure: Schulführung, Schulorganisation und Kollegium***Schulführung*

- Die Schulführung ist von Zielorientierung, Mitbestimmung, Transparenz der Information und Kommunikation sowie von Respekt gekennzeichnet.
- Entscheidungsprozesse sind für alle transparent und klar an definierte Kompetenzen gebunden.
- Schulkonferenzen werden effektiv vorbereitet, durchgeführt und nachgearbeitet

*Schulorganisation und Schuladministration*

- Es bestehen Kommunikations-, Informations- und Kooperationsgefäße.
- Lehrer einer Klasse bilden eine Organisationseinheit, die sich regelmäßig trifft.
- Die Verwaltung der Schule ist ökonomisch und zuverlässig geregelt. Sie ist zudem kundenorientiert und freundlich.

**Kollegium**

- Die Beteiligung der einzelnen Lehrpersonen am schulischen Leben gilt als Teil der Professionalität und der Dienstpflicht.
- Das Gefühl der Zugehörigkeit zur Schule und der Zusammenhalt im Kollegium werden unterstützt (z.B. durch Feste und Feiern).
- Kollegiale Anliegen werden respektvoll und fair verhandelt. Die Privaträume und auch lehrbezogenen Eigenverantwortlichkeiten werden respektiert.
- Die Schule versteht sich als Teil der regionalen und lokalen Kultur.

**Umgang mit der Schülerschaft**

- Es gibt ein gemeinsames pädagogisches Konzept des Umgangs mit der Schülerschaft, das persönlich, wertschätzend, aufrichtig, freundlich und respektvoll ist und dabei konsequent auf komplementäre Haltungen der Schülerschaft achtet.
- Die Regeln akzeptablen Verhaltens der Schüler in der Schule und in der Schulklasse werden klar dokumentiert und auch durchgehalten.
- Es gibt ein Schulkonzept gegen Mobbing und Ausgrenzung einzelner Schüler.
- Prosoziales und gemeinschaftsförderndes Handeln von Schülern wird klar kommuniziert, erwartet und auch wertschätzend beantwortet.
- Neben den fachbezogenen Leistungserfolgen werden pädagogische und erzieherische Wirkungen regelmäßig erfasst und in die Gestaltung des Schullebens einbezogen.

Abb. 4: Referenzrahmen der Qualitätsbeurteilung (nach Landwehr 2003)

<b>1. Stoffbeherrschung</b>	Sachliche Richtigkeit, Allgemeine Stoffbeherrschung, Querbezüge
<b>2. Sprachbeherrschung</b>	Korrektheit, Flüssigkeit, Wortschatz, Ausdrucksvermögen, Genauigkeit der Formulierung, Klarheit und Verständlichkeit für die Schüler
<b>3. Didaktische Präsentation</b>	Anknüpfung, Angemessenheit von Methode und Schwierigkeitsgrad, Strukturierung, Anschaulichkeit, Rhythmus, Lebhaftigkeit, Begeisterung
<b>4. Unterrichtsmittel</b>	Zweckmässigkeit im Einsatz, Handhabung, Gestaltung (Tafel, Projektion, Arbeitsblätter etc.)
<b>5. Konsolidierung</b>	Wiederholung, Anwendung, Korrektur von Fehlern, Kontrollfragen, Überwachen von Notizen
<b>6. Unterrichtsgespräch</b>	Aufnehmen von Schülerbeiträgen, gute Erklärungen bei Schwierigkeiten, Aktivierung, Einbezug der Stillen, Vermeidung enger Fragen, Ermöglichen einer Diskussion zwischen Schülern, geeignete Reaktion auf Störungen
<b>7. Umgang</b>	Echtheit, ruhige Bestimmtheit, Selbstvertrauen, menschliche Zuneigung, Wärme, Humor, Wertschätzung, Einfühlungsvermögen

Abb. 5: Kriterien für die Beurteilung von Lektionen  
(Aus Studienkommission für bildungspsychologische Fragen In Zürich)

Aber auch die Unterrichtswissenschaft hat sich, im Sinne des „Wissens über das System“, intensiv mit Qualitätsdimensionen beschäftigt. Helmke (2003) hat dies bisher unübertroffen resümiert, und dabei folgende Aspekte hervorgehoben:

- Gute Lehre ist inhaltlich klar und präzise,
- sie setzt auf Methodenvielfalt und nicht auf Methodenmonismus,
- sie adaptiert Inhalte an die individuellen Lernmöglichkeiten und versucht Überforderung und Unterforderung zu vermeiden,
- sie kann eine Vielfalt von Motivierungsstrategien einsetzen, intrinsische, aber auch extrinsische,
- sie ist effizient in der Klassenführung, nimmt die Regulierungsaufgabe effektiv und flexibel wahr,
- sie zeigt hohe diagnostische Kompetenz, kann Lernprobleme einzelner Schüler und den Leistungsstand einer Schulklasse zuverlässig beurteilen.

Damit konvergieren die Kriterien guten Unterrichts, wie sie schon Brophy resümiert hat (zit. nach Helmke 2003, S. 63f.). Sie beziehen sich einmal auf die Inhalte, wenn die Kriterien

- Orientierung am Lehrplan,
- innerer Zusammenhang der Inhalte und
- gut durchdachter Unterrichtsplan

betont werden.

Die Unterrichtsform wird akzentuiert durch die Kriterien

- Übung und Anwendung,
- Anleitung der Eigentätigkeit der Schüler,
- Gestaltung kooperativen Lernens,
- Lehren von Lernstrategien.

Schließlich wird guter Unterricht motivational gesteuert durch

- den Aufbau einer Lern- und Aufgabenorientierung,
- kriteriumsbezogene Leistungsbeurteilung, um Erfolge sichtbar zu machen, und durch den
- Aufbau von Leistungserwartungen.

In Versuchen, nicht nur Qualitätsdimensionen *summativ* aufzuzählen, sondern sie *theoretisch zu verankern*, stehen heute *ko-konstruktive Ansätze* im Mittelpunkt, die davon ausgehen, dass am Aufbau und an der Konsolidierung von kognitiven Strukturen bei Kindern und Jugendlichen beide Parteien, Lehrende und Lernende aktiv beteiligt sein müssen. In der Tradition von Piaget, Aebli und Vygotsky arbeiten heute u.a. Reusser (1998) und Klieme u.a. (2006) mit solchen Modellen. Die Qualitätskriterien verdichten sich dabei z.B. in *kognitive Aktivierungsqualitäten*, die eine große Verarbeitungstiefe und Konsolidierung sicher stellen, in eine Aufmerksamkeit maximierende *Klassenführung* und *Regelklarheit* des Unterrichts und in ein *unterstützendes Unterrichtsklima*, das motivational über die Erfahrung von Autonomie, Kompetenz und soziale Zugehörigkeit lernförderlich wirkt.

Bei einer *ko-konstruktiven Grundhaltung* wird die *Anschlussfähigkeit* der Lehre an die Lernenden zu einem Schlüsselkriterium für die Qualität des Unterrichts. Sie konkretisiert sich in der *Qualität der Wahrnehmung des jeweiligen Lernstandes* einer Schulklasse, in einer Unterrichtsweise, in der das *Verständnis des Gelehrten prozessural beobachtet werden kann* und in *adaptiven Reaktionen der Lehrperson*, durch die Schüler immer wieder genau dort abgeholt werden, wo sie stehen.

## Resümé

Das Qualitätsbewusstsein zur Gestaltung des Bildungswesens hat sich in den letzten Jahrzehnten in beeindruckendem Ausmaß differenziert (Fend 2008), auch wenn häufig implizites Qualitätswissen damit nur explizit gemacht wurde. Unübersehbar ist dies aber nicht nur eine *Orientierungshilfe* für die Praxis, sondern auch ein *Anforderungsprofil*, das die Schulen und die Lehrpersonen unter *hohen Druck* setzen kann. Wenn die Bildungspolitik und die Administration diese Qualitätskriterien formulieren, dann tun sie dies gewissermaßen am grünen Tisch und geben die Richtung vor, in die sich das Bildungswesen entwickeln sollte. Die Schärfung des Qualitätsbewusstseins ist zwar unab-

dingbar, aber nur der erste Schritt zur faktischen *Qualitätsverbesserung* des Bildungswesens. Sie ist nicht allein dadurch zu erreichen, dass Forderungen aufgestellt werden.

Damit diese nicht schlicht das Frustrationspotential in den Schulen erhöhen, ist eine kluge *Unterstützung* für die Entwicklungsprozesse unabdingbar. Frustrationen sind unausweichlich, wenn Schulen und Lehrpersonen einfach mehr leisten sollen, mehr Arbeit aufgebürdet bekommen und dies unter allgemeiner Einschränkung der Handlungsmöglichkeiten tun sollen. Qualitätsentwicklung und „Sparleistungen“ harmonisieren nicht. So ist den Lehrpersonen in den letzten Jahren mehr an Arbeit aufgebürdet worden, sei dies durch Erhöhung ihrer Deputate, durch die Vergrößerung der Schulklassen, durch die Durchführung von extern vorgegebenen Prüfungen, die Entwicklung von Sonderprogrammen an Schulen und die Verpflichtung zur Evaluation. Die Schulleitungen selber haben viele neue Arbeitspakete erhalten, die mit den Autonomieregelungen verbunden sind (eigene Personalrekrutierung, Personalführung durch die einstufige Beurteilung, die Vergabe von Leistungsstufen und die Ausschreibung von Beförderungstellen, Entwicklung von Schulprogrammen und Schulcurricula, zunehmende Differenzierung von Profilen, Einführung von neuen Bildungsgängen, Evaluation und Qualitätssicherung, Ganztagschulen), die aber einen erheblichen Mehraufwand auf der schulischen Leitungsebene zur Folge haben.

Die *Achse der Kontrolle und Qualitätssicherung* bedarf in den nächsten Jahren einer Ergänzung und verstärkten Aufmerksamkeit für *Unterstützungssysteme*, die erforderlich sind, um die Qualitätssicherung zur Geltung kommen zu lassen. Auf die Zeit der Formulierung von Qualitätsdimensionen könnte eine der Konzentration auf die Realisierungsbedingungen folgen. Dann wären wir auf dem richtigen Weg.

## Literatur

- Altrichter, H./Brüsemeister, T./Wissinger, J. (2007): Educational Governance. Handlungskoordination und Steuerung im Bildungswesen. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bourdieu, P. (1983): Ökonomisches Kapital, kulturelles Kapital und soziales Kapital. In: Kreckel, R. (Hrsg.): Soziale Ungleichheiten. Vol. Soziale Welt, Sonderband 2, S. 183–198.
- Klieme, E./Avenarius, H./Blum, W./Döbrich, P./Gruber, H./Prenzel, M./Reiss, K./Riquarts, K./Rost, J./Tenorth, H.-E./Vollmer, H. J. (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Berlin: Bundesministerium für Bildung und Forschung.
- Depaeppe, M. (2000): Order in progress. Everyday education practice in primary schools Belgium 1880–1970. Leuven: Leuven University Press.
- Ditton, H. (2000): Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Zeitschrift für Pädagogik, 41. Beiheft, S. 73–92.
- Doll, J./Prenzel, M. (Hrsg.) (2004): Bildungsqualität von Schule. Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung. Münster: Waxmann.
- Fend, H. (2001): Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung. Weinheim: Juventa Verlag.
- Fend, H. (2006a): Geschichte des Bildungswesens. Der Sonderweg im europäischen Kulturraum. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Fend, H. (2006b): *Neue Theorie der Schule. Eine Einführung in das Verstehen von Bildungssystemen.* Wiesbaden: VS Verlag für Sozialwissenschaften.
- Fend, H. (2008): *Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität.* Wiesbaden: Verlag für Sozialwissenschaften.
- Foucault, M. (1976): *Überwachen und Strafen.* Frankfurt a. M.: Suhrkamp.
- Heid, H. (2000): *Qualität. Überlegungen zur Begründung einer pädagogischen Beurteilungskategorie.* In: *Zeitschrift für Pädagogik*, 41. Beiheft, S. 41–54.
- Helmke, A. (2003): *Unterrichtsqualität erfassen, bewerten, verbessern.* Seelze-Velber: Kallmeyer.
- Klieme, E./Lipowsky, F./Rakoczy, K./Ratzka, N. (2006): *Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projektes ‚Pythagoras‘.* In: Prenzel, M./Allolio-Näcke, L. (Hrsg.): *Untersuchungen zur Bildungsqualität von Schule.* Münster: Waxmann, S. 127–146.
- Künzli, R./Bähr, K./Fries, A.V./Ghisla, G./Rosenmund, M./Seliner-Müller, G. (1999): *Lehrplanarbeit. Über den Nutzen von Lehrplänen für die Schule und ihre Entwicklung.* Chur: Verlag Rüegger.
- Künzli, R./Hopmann, S. (Hrsg.) (1998): *Lehrpläne: Wie sie entwickelt werden und was von ihnen erwartet wird. Forschungsstand, Zugänge und Ergebnisse aus der Schweiz und der Bundesrepublik Deutschland.* Chur: Rüegger.
- Kussau, J./Brüsemeister, T. (2007): *Governance, Schule und Politik. Zwischen Antagonismus und Kooperation.* Wiesbaden: VS Verlag für Sozialwissenschaften.
- Landwehr, N. (2003). *Basisinstrument zur Schulqualität. Systematische Darstellung wichtiger Qualitätsansprüche an Schule und Unterricht.* Bern: h.e.p.
- Prenzel, M./Allolio-Näcke, L. (Hrsg.). (2006): *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms.* Münster: Waxmann.
- Reusser, K. (1998): *Denkstrukturen und Wissenserwerb in der Ontogenese.* In: Klix, F./Spada, H. (Hrsg.): *Enzyklopädie der Psychologie (Vol. 6).* Göttingen: Hogrefe Verlag für Psychologie, S. 115–166.
- Stryck, T. (2000): *Qualitätssicherung in der Geisterbahn.* *Zeitschrift für Pädagogik*, 41 Beiheft, S. 111–128.

Anschrift des Autors:

Prof. Dr. Dr. h.c. Helmut Fend, Neuhauserstraße 6, D 78464 Konstanz, Tel.: 07531/31875, e-mail: helmut.fend@t-online.de

*Esther Berner/Jürgen Oelkers/Kurt Reusser*

## **Implementation von Bildungsstandards**

*Bedingungen des Gelingens (und Scheiterns) aus internationaler Sicht*

### **1. Kontext und Ansprüche des Projekts ‚Bildungsstandards‘**

Im Anschluss an die fortgeschrittenen Entwicklungsarbeiten steht die Implementation von Bildungsstandards momentan ganz oben auf den bildungspolitischen Agenden der Länder Deutschland, Schweiz und Österreich. Zentrale Bedeutung kommt Bildungsstandards als fachbezogene Leistungsstandards im Rahmen eines neuen, ergebnisorientierten Paradigmas der Bildungssteuerung und Qualitätssicherung zu, das u.a. auf dem Hintergrund von Resultaten in internationalen Schulleistungsstudien in den Blickpunkt getreten ist. In politischen Verlautbarungen und konzeptionellen Papieren mangelt es seither nicht an Hinweisen darauf, dass erfolgreiche TIMSS- und PISA-Länder in Skandinavien oder im angelsächsischen Raum ihre Anstrengungen zur Systemsteuerung seit längerem auf eine Ergebnisorientierung umgestellt hätten.

Mit der Umsetzung von Bildungsstandards im deutschsprachigen Raum verbinden sich zugleich Hoffnungen auf die Einlösung von Reformbestrebungen, die seit den 90er Jahren vermehrt die schulpädagogischen und -politischen Diskussionen bestimmt haben. Insofern lässt sich sagen, dass das Konzept der Bildungsstandards in bereits vorangehenden und aktuellen Problemwahrnehmungen und Entwicklungstendenzen einen geeigneten Boden fand; auf diesem Hintergrund erscheint die Implementation von Standards sinnvoll oder gar notwendig. Bildungsstandards bilden so etwa unter dem Vorzeichen der Rechenschaftslegung ein Komplement zu den schon länger bestehenden Absichten, den Einzelschulen mehr Eigenständigkeit zu gewähren. Bildungsstandards wollen aber auch Lösungen anbieten in der Lehrplandebatte (Problem der Überfrachtung, Lehrzielkonfusion bei erwiesenermaßen geringer Steuerungswirkung) oder hinsichtlich einer wahrgenommenen Ungerechtigkeit und Krise in der Zertifizierung von Schulabschlüssen (Fehlen eines Bezugssystems für Schulnoten, schwindende Akzeptanz des Wertes der von Schulen vergebenen Noten und Abschlüsse). Fragen der Kosteneffizienz im internationalen Vergleich verlangen nach Kriterien der Messung des Outputs, ebenso der Vergleich der durch ein System erbrachten Bildungsleistungen angesichts von Herausforderungen, wie sie mit einer sich rasch wandelnden ‚Wissensgesellschaft‘ oder dem internationalen Wettbewerb unter Bedingungen der ‚Globalisierung‘ verknüpft werden.

Bereits diese Auswahl an Ansprüchen verweist auf Spannungsfelder und Kollisionsgefahren angesichts unterschiedlicher Zielsetzungen des Projekts Bildungsstandards. Eine neuralgische Stelle liegt dabei offensichtlich dort, wo den Lehrpersonen und Schulleitungen ein Mehr an Autonomie angeboten und gleichzeitig vermehrt Verantwortung

für Ergebnisse zugewiesen und Rechenschaft abverlangt wird. Zudem wird erwartet, dass sich über die Setzung von Leistungsstandards sowie unter Bedingungen der Output-Orientierung und dezentralen Verlagerung von Verantwortung eine Diversifikation des schulischen Angebots einstellt, die den Schulen zugleich den Rahmen bietet, ihre Qualität auch auf der Unterrichtsebene in einer Weise zu verbessern, die sich dann auch noch an Standards messen und damit als Qualitätssteigerung ausweisen lässt. Soweit Erfolg an den Leistungen der Schüler/-innen gemessen wird, wird nicht nur auf die Verbesserung des durchschnittlichen Leistungsniveaus abgehoben, sondern auch auf die Verringerung soziokulturell vermittelter Leistungsvarianz.

Das sind hohe Erwartungen, und die Rede von der ‚neuen Output-Steuerung‘ birgt die Gefahr in sich, zu glauben, diese ohne großen Aufwand bezüglich des erforderlichen Inputs sowie der Investitionen in die Schaffung notwendiger Kontextbedingungen erfüllen zu können. Auch müssen etwa Trends in Richtung Deregulierung und Dezentralisierung unweigerlich an der Forderung, allen ein qualitativ möglichst gleich hoch stehendes Angebot zugänglich zu machen, ihre Grenzen erreichen. Im Rahmen solcher Ambivalenzen lässt sich zumindest das Ergebnis einer Untersuchung der curricularen Steuerungspolitik von neun verschiedenen Ländern/Bundesstaaten<sup>1</sup> deuten, die – im Sinn einer „repeating chain of action-reaction“ – gezeigt hat, dass „(...) a too strong movement towards decentralisation in the end results in a movement in the opposite direction (more top-down steering and control), and the other way around“ (Kuiper u.a. 2006, S. 76).

Wird die Implementation von Bildungsstandards nicht isoliert und lediglich technisch verstanden im Sinne der Diffusion eines neuen, abgrenzbaren Produktes an die vorgesehenen Nutzer/-innen, sondern befasst man sich damit im Kontext oben genannter Problem- und Entwicklungsfelder, zu deren Lösung bzw. Fortschritt sie beitragen sollen, so bedarf es einer integralen Betrachtungsweise; dabei kann die Analyse von Erfahrungen anderer Länder, die ähnliche Reformen bereits früher initiiert und umgesetzt haben, durchaus von Nutzen sein. Entsprechende Untersuchungen waren denn auch Bestandteil des inzwischen mit einer Publikation (vgl. Oelkers/Reusser 2008) abgeschlossenen Expertise-Projektes, das sich im Auftrag der Länder Deutschland, Luxemburg, Österreich und Schweiz mit der Frage der Einführung oder, besser, Umsetzung von Bildungsstandards beschäftigte. Einbezogen wurden die Entwicklungen in England, Schweden, den Niederlanden und den USA. Die Wahl fiel dabei auf jene europäischen Staaten, die mit Blick auf die international vergleichenden Leistungsstudien in den vergangenen beiden Jahrzehnten offenbar erfolgreiche Reformen der Bildungssteuerung und Qualitätssicherung sowie der Schulentwicklung umgesetzt haben. Dabei ist bemerkenswert, wie die jeweiligen Bestrebungen von recht unterschiedlichen Ausgangslagen und Voraussetzungen ausgehen, sich aber dennoch in ihren Grundtendenzen überschneiden. Gerade dies sollte es erlauben, trotz der selbstverständlichen Notwendigkeit, im Einzelnen die je spezifischen nationalen Kontexte zu berücksichtigen, sowohl einige

1 Es handelt sich um England, Belgien/Flandern, Finnland, Ungarn, Deutschland (Nordrhein-Westfalen, Schleswig-Holstein), die Niederlande, Portugal, Schweden, die USA/Kalifornien.

verallgemeinerbare Schlüsse zu ziehen wie auch spezifische Zusammenhänge einsehbar zu machen. Die USA sind dabei von besonderem Interesse, weil hier im Zuge standard-basierter systemischer Schulreformen eine große Zahl an Begleit- und Evaluationsstudien entstanden ist, die die Implementation auf den verschiedenen Ebenen des Systems untersuchen. Parallel dazu und gestützt auf entsprechende empirische Ergebnisse entwickelte sich ein Bereich wissenschaftlicher Theoriebildung, der sich dezidiert mit der Frage der Implementation von Innovationen im Kontext von Schulreformen und insbesondere der Umsetzung von Standards beschäftigt. Sämtliche Staaten können auf Reformen zurückblicken, die bereits mehrere Phasen der Entwicklung kennen, Auswirkungen gezeitigt haben sowie auch (unerwünschte) Nebenwirkungen, auf die zuweilen Korrekturmaßnahmen folgten.

Im Folgenden werden drei ausgewählte Problemaspekte der Implementation von Bildungsstandards erörtert. Mit der Frage nach der Bedeutung von *Kohärenz* bzw. (im Kontext dieses Beitrags synonym verwendet) *alignment* wird zuerst ein Systemaspekt angesprochen, der im Feld von Steuerungspolitiken, die auf Dezentralisierung und Deregulierung setzen, ganz besonders zum Tragen kommt. Ein zweites potentiell Konfliktfeld stellen unterschiedliche Ansprüche von *Politik* und *Profession* dar, die sich vor dem Hintergrund neuer Zuordnungen von Autonomie einerseits und Rechenschaftspflicht bezogen auf zentral vorgegebene Standards andererseits ergeben und besonders beim Thema ‚Assessment‘ zu Kontroversen führen. Und drittens und letztens geht es um die Bedeutung von Ressourcen bzw. *Kapazitäten (capacity)* gerade als Konsequenzen einer Output-Orientierung des Bildungssystems.

## 2. Die Frage der Kohärenz im Spannungsfeld zwischen zentraler Regulation und Dezentralisierung

Das Kriterium der Kohärenz lässt sich auf verschiedene Aspekte eines Bildungssystems sowie von Reformstrategien beziehen. In den USA und England, in Staaten also, in denen, ausgehend von unterschiedlichen Voraussetzungen, eine Verstärkung zentraler Steuerung stattgefunden hat bzw. angestrebt wird, wird die Kohärenz zwischen Systemkomponenten unter dem Begriff des *alignment* stark diskutiert. Umgekehrt, so die These, kann mangelnde Kohärenz unter Umständen dann als Problem hervortreten, wenn Aufgaben der Steuerung und Kontrolle im Sinne der Dezentralisierung nach unten bzw. außen delegiert werden, wie es exemplarisch in Schweden der Fall ist.

### 2.1 USA

Blickt man auf die USA, trifft man auf Zentralisierungstendenzen einerseits im Verhältnis von Einzelstaaten und Distrikten (z.B. Kalifornien), andererseits zeigen sich solche auch ausgehend von der Ebene des Bundes, indem dieser die Staaten und ihre Schulen mit der Gesetzgebung „No Child Left Behind“ (2001) für die Primar- und Sekundar-

schulen zunehmend einer Rechenschaftspflicht unterzieht. Die Forderung nach alignment steht in diesem Kontext in erster Linie für die gegenseitige inhaltliche und konzeptionelle Abstimmung zwischen ‚content standards‘ und ‚performance standards‘, den entsprechenden Assessments sowie verwendeten Unterrichtsmaterialien. An diese Elemente anzuschließen haben wiederum auch die Lehrerausbildung und Gelegenheiten zur Weiterbildung sowie Unterstützungsangebote. Übereinstimmung ist aber auch in der Vertikalen zwischen den unterschiedlichen Standards, Assessments und Accountability-Systemen gefordert, wie sie auf den verschiedenen Ebenen von Staat und Distrikt, aber auch auf bundesstaatlicher Ebene vorgegeben bzw. auf nationaler Ebene zumindest teilweise vorhanden sind. Handlungsbedarf zur Herstellung von Kohärenz ist dabei augenscheinlich gegeben, wie die Befunde einer Untersuchung von Porter (2002) nahelegen. Sie zeigten, dass das alignment von Assessments und Mathematik-Standards innerhalb mehrerer untersuchter Staaten zum Untersuchungszeitpunkt wenn, dann offenbar eher zufällig vorhanden und nicht stärker war als im Verhältnis zu den Standards der anderen Staaten oder zu den Standards des National Council of Teachers of Mathematics (NCTM) (vgl. National Council of Teachers of Mathematics 2000).<sup>2</sup>

Mit Blick auf die Frage der Kohärenz lässt sich eine zeitliche Dimension anfügen und zwar im Sinne der Forderung nach einer Kontinuität bezüglich politischer Prioritätensetzungen und Impulsgebung für schulische Reform- und Entwicklungsvorhaben. Diese Dimension spielt eine umso größere Rolle, als die Zeit, die die Umsetzung von Reformen im komplexen System- und Handlungsbereich von Schule und Unterricht beansprucht, häufig unterschätzt wird.

Im Zuge von Erfahrungen mit der Implementation von Standards und testbasierten Accountability-Systemen hat sich in den USA ein gewisser Vorrang des Kohärenzanspruchs durchgesetzt. Er lässt sich auf den Beginn der 90er Jahre zurückverfolgen und dort im Kontext der sogenannten *Standards-based Systemic Reform* ansiedeln. Dieser Ansatz ist als Reaktion auf vorangehende Versuche der Anhebung der nationalen Schülerleistungen und outcomes, v.a. mittels Formulierung anspruchsvoller Standards, zu verstehen. Daneben zielten die Maßnahmen der 80er Jahre insgesamt auf vermehrte staatliche Regulierung und zusätzliche Verordnungen; die Unterrichtszeiten wurden ausgedehnt und gleichzeitig die Lehrergehälter erhöht. Die pädagogische Fachöffentlichkeit begegnete dieser Strategie bald mit Kritik: Es handle sich dabei um eine einseitige Top-down-Politik, die lediglich zu einer „intensification of task“ (vgl. z.B. Covington 1996) im Bildungswesen führe. Gemeint ist damit, dass es sich bei den Umsetzungen um keine wirklichen Innovationen handle, vielmehr würde auf diese Weise lediglich immer mehr vom Gleichen gefordert, ohne dass damit die Schulorganisation oder der Unterricht der Lehrpersonen in irgendeiner Weise optimiert würden.

2 Eine andere Untersuchung (vgl. Blank/Porter/Smithson 2001) hat gezeigt, dass das alignment des Mathematikunterrichts von Schulen in sechs Staaten zum staatlichen Assessment geringer war als zum nationalen NAEP-Assessment (National Assessment of Educational Progress); das alignment des Unterrichts zu den Assessments der anderen Staaten war gleich hoch wie zum eigenen Assessment.

Der neue Ansatz der *Systemic School Reform*, maßgeblich mitinitiiert durch den gleichnamigen Artikel von Smith und O'Day (vgl. 1990), sieht demgegenüber vor, Top-down-Strategien mit Bottom-up-Initiativen zu verknüpfen, wobei dem *alignment* besondere Bedeutung zugemessen wird. Diese neue Perspektive unterscheidet sich nach Lusi „(...) from the reform attempts of the 1980s in at least two important ways. First, systemic school reform strives to reform the education system as a system; it works for coherence across the system's component policies, something that the piecemeal reform of the past did not achieve. Second, systemic school reform explicitly strives to support school-site efforts at redesigning teaching and learning with the goal that all students will learn ambitious content knowledge and higher-order skills. It is insufficient to promulgate mandates such as increased graduation requirements from the ‚top‘ of the education system (state). The ‚bottom‘ of the system (schools and districts) must be supported and activated to transform teaching and learning“ (Lusi 1997, S. 6).

Die kritisierte „piecemeal reform“ meint vorangehende Programme, die gewöhnlich auf bestimmte Schülerpopulationen mit speziellen Bedürfnissen gerichtet waren oder sich auf den Unterricht gewisser Fächer bzw. Kompetenzen (z.B. Förderprogramme bezogen auf ‚numeracy‘ oder ‚literacy‘) beschränkten. Das Resultat war eine Ausgliederung schwächerer Schüler/-innen mit ‚special needs‘ in Sonderprogramme, sogenannte ‚pull-out programs‘, und eine Fragmentierung des Unterrichts. Damit fand auch eine Differenzierung der Curricula (und Lerngelegenheiten überhaupt) statt, in solche, die je nach Voraussetzungen der Schüler/-innen verstärkt auf Minimumkompetenzen (‚basic skills‘) oder auf komplexere ‚higher-order skills‘ ausgerichtet waren (vgl. Smith/O'Day 1993). Gerade vor diesem Hintergrund ist es denn auch Anspruch und Programm von ‚No Child Left Behind‘, allen Kindern, die der Förderung bedürfen, diese im Rahmen einer Gesamtstrategie zukommen zu lassen sowie sämtlichen Schüler/-innen grundsätzlich dieselben Lernbedingungen bereitzustellen.

Das *alignment* zwischen den verschiedenen Systemebenen und ihren Anforderungen an das Lehren und Lernen, die Abstimmung zwischen den Inhalten und Leistungen, auf die hin unterrichtet wird, und dem, was in Tests gemessen wird, kann in einem an Standards orientierten Bildungswesen als ein Maß für deren Implementation gelten. Es ist damit auch ein Muss für jedes System, das von den Schulen und z.T. auch den Lehrkräften Rechenschaft für erbrachte Leistungen und Qualität einfordert, will es denn fair sein. Je eindeutiger diese Kohärenz ist und je rigider die Accountability-Mechanismen sind, desto eher findet auf der anderen Seite das statt, was kritisch als ‚teaching to the test‘ bezeichnet wird.

Damit ist ein Spannungsfeld angesprochen, für das sich auch Beispiele im europäischen Kontext finden lassen. Dies trifft etwa auf England zu, aber, wenngleich unter den umgekehrten Vorzeichen von Dezentralisierung (Kommunalisierung) und Deregulierung, auch auf Schweden. In Schweden lassen sich mit der Einführung der freien Schulwahl, der Förderung von privaten Schulen und Wettbewerb tendenziell ähnliche Liberalisierungstendenzen feststellen, wie sie in England die Politik seit den 80er Jahren bestimmen. Zugleich scheint die stark verankerte wohlfahrtsstaatliche Maxime der Gleichheit dieser Stoßrichtung doch auch gewisse Grenzen zu setzen.

## 2.2 Schweden

Die tiefgreifende Verwaltungsreform und Umstrukturierung des Bildungswesens hatte im Schweden der 80er und 90er Jahre ihren Ausgangspunkt in einer über Jahrzehnte vorherrschenden zentralistischen, über Erlasse und Regulative funktionierenden Bildungssteuerung. Schweden kennt zwar keine Standards im eigentlichen Sinn, jedoch ein nationales Curriculum, dessen Dichte rein umfangmäßig zwischen 1962 und 1993 beträchtlich abgenommen hat.<sup>3</sup> Daneben existieren, allerdings ebenfalls äußerst allgemein gehaltene, Syllabi für die einzelnen Fächer; diese unterscheiden sich deutlich vom englischen National Curriculum mit seinen ‚programmes of study‘ für jedes Fach und differenziert nach sogenannten Schlüsselstufen und Niveaus. Inzwischen spielen landesweite Tests in Schweden ebenfalls eine wichtige Rolle für die Überprüfung der Zielerreichung, und die Aufgaben sind von erheblichem Einfluss darauf, wie Lehrkräfte die Lehrpläne interpretieren (vgl. Erickson/Lander 2007).<sup>4</sup> Verpflichtend sind sie in den Hauptfächern in der 9. Klasse und am Ende der Pflichtkurse in den Jahrgangsstufen 10-12. Zusätzliche Tests können freiwillig bzw. je nach Vorgaben des kommunalen Schulträgers bereits früher und in anderen Fächern durchgeführt werden. Zeitgleich mit der Kommunalisierung hat Schweden zudem ein mehrere Ebenen umfassendes System der Qualitätssicherung hervorgebracht. Dazu gehört auf der Ebene der Gemeinden die Verpflichtung, der nationalen Schulbehörde (Skolverket) zielorientierte *Schulpläne* vorzulegen, in denen sie aufzeigen müssen, wie die nationalen Ziele erreicht werden sollen. Neben deren Implementation sind die Gemeinden auch dafür verantwortlich, in einem jährlich veröffentlichten *Qualitätsbericht* die erreichten Fortschritte sowie allfällige Schwächen aufzuzeigen (vgl. Kotthoff 2003, S. 257). Auf der Ebene der Einzelschule wiederum muss jede Schule einen eigenen *Arbeitsplan* entwickeln und jährlich in Form eines schulischen *Qualitätsberichts* Rechenschaft gegenüber der Kommune ablegen; letzterer hat sich explizit auf die kommunalen und nationalen Ziele zu beziehen und Instrumente der innerschulischen Evaluation zu beschreiben (vgl. ebd., S. 282). Es ist letztlich die Einzelschule, die die Hauptverantwortung für die Evaluation der Qualität der Schule und des Unterrichts tragen soll.

Das schwedische Qualitätssicherungs- und Steuerungssystem funktioniert über die Setzung von Zielen und die Überprüfung der Resultate ausgehend von diesen Zielen, wobei die Kohärenz zwischen den Ebenen über die explizite Bezugnahme der Steuerungsdokumente und Evaluationsinstrumente auf die Zielformulierungen der jeweils übergeordneten Ebene(n) hergestellt wird. Die Architektur ist bestechend, verlangt aber, da die Vorgaben zunächst recht vage formuliert wurden, von den Akteuren beträchtliche Übersetzungsleistungen. Die damit offen gehaltenen Interpretationsspielräume für

3 Anzahl der Seiten im Lehrplan für die Grundschule im Vergleich: Formulierung übergeordneter Ziele 1962: 88 S., 1994: 12 S.; Zeitbudgets 1962: 12 S., 1994: 1 S.; fächerspezifische Lernziele 1962: 310 S., 1994: 30 S. (vgl. Ekholm 2006).

4 Die nationalen Tests am Ende der 5. und 9. Klasse üben gemäß einer Untersuchung mehr Einfluss auf den Unterricht aus als das nationale Curriculum (vgl. Kuiper u.a. 2006).

die eigene inhaltliche Umsetzungsarbeit kommen zwar kommunalen und einzelschulischen Autonomiebedürfnissen entgegen, es zeigt sich aber auch, dass sie nur dann produktiv genutzt werden können, wenn die Akteure über die notwendigen Ressourcen verfügen und v.a. Gelegenheit haben, die erforderlichen Kapazitäten auszubilden. Verschiedene Studien verweisen darauf, dass hier die großen Herausforderungen im Zuge der Reform lagen und immer noch liegen. Die abstrakte Ausführung etwa der nationalen Vorgaben für die Erstellung des kommunalen Schulplans und für die kommunalen Evaluationen zu Beginn führte dazu, dass Skolverket die von den Kommunen eingereichten Dokumente in vielen Fällen als mangelhaft beurteilen musste. 1999, zwei Jahre nach deren verpflichtender Einführung, hatten lediglich 45% der Kommunen Qualitätsberichte vorgelegt (vgl. Slenning 2004). Erst daraufhin hat Skolverket, auch im Sinne der Vergleichbarkeit, verbindliche Vorgaben für die Erstellung der Berichte erlassen. 2001 lieferten dann 207 der 289 Kommunen den Bericht ab, jedoch kamen lediglich 107 davon den gesetzlichen Richtlinien nach (vgl. Daun 2004, S. 83). Ähnlich liegen die Verhältnisse bezüglich des schulischen Qualitätsberichts, wo ebenfalls mit der Entwicklung nationaler Richtlinien erst allmählich eine gewisse Standardisierung erreicht worden ist, wobei die Evaluationsverfahren der einzelnen Schulen untereinander noch immer erheblich variieren.

Slenning (vgl. 2004) bezeichnet das schwedische System als ‚loosely coupled‘, was, positiv formuliert, Entscheidungs- und Handlungsspielräume für die lokale Entwicklung eröffnet – „ (...) a development that might not occur in a more tightly regulated and assessed environment“ (ebd., S. 54). Entscheidend dabei ist das Vertrauen in die Professionalität der Lehrkräfte. Dennoch mag dieses Arrangement Ambiguitäten mit sich bringen, die nicht unbedingt als produktiv bezeichnet werden können. Dies gilt etwa, wenn nationale und lokale Curricula oder kommunale und lokale Schulpläne nicht miteinander korrespondieren; die Existenz von Steuerungsdokumenten verschiedener Provenienz mit unterschiedlichen Standards erschweren einer Lehrperson unter Umständen die Unterrichtsarbeit sowie die Bewertung der Schülerleistungen. Insgesamt bleibt doch die Gefahr, dass „(...) [t]he sheer complexity of a system that includes loosely coupled sets of standards and comprehensive syllabi, a strong local planning mandate, (...) requires a level of competence that may not exist among all actors within the system“ (ebd., S. 60f.).

### *2.3 England*

Die Spannung zwischen zentralen Kontroll- und Regulationsmechanismen einerseits, professionellen Handlungs- und Entwicklungsmöglichkeiten der Lehrerschaft andererseits lässt sich auch am Beispiel Englands nachvollziehen. Dort wurden Ende der 80er/Anfang der 90er Jahre ein nationales Curriculum, ein ausgebautes System von National Curriculum Assessments und ein Inspektionswesen mit ausgeprägter Kontrollfunktion eingeführt, während die Lehrkräfte vorher über relativ weitgehende pädagogische Autonomie im Rahmen der dezentralen Aufsicht, verkörpert in den Local Educa-

tion Authorities (LEAs), verfügt hatten. Die zentrale Regulationsmacht ist in einzigartiger Weise verkörpert im detaillierten National Curriculum mit seinen differenzierten Niveauvorgaben für jedes Fach. Im Gegensatz zu den Niederlanden und Schweden hat die Inspektionsagentur Ofsted weitgehende Kompetenzen bezüglich der Einleitung von Maßnahmen auch sanktionierender Art. Ausgesprochen weit entwickelt ist im englischen Schulwesen das *alignment* zwischen Standards und Tests, aber auch zu den Kriterien der externen und schulinternen Evaluation, was eine Voraussetzung bildet für die ausgeprägte Accountability bei gleichzeitiger Forcierung eines ‚Quasi-Marktes‘ im Bildungsbereich. Zwar wurde den Schulen, parallel zu einer Schwächung der LEAs, mehr Selbstständigkeit und Eigenverantwortung zugesprochen, dennoch spüren die Lehrkräfte aber in erster Linie die Einschränkung ihrer traditionellen Lehrfreiheit. Besonders in dieser Hinsicht kann der Wandel des englischen Bildungswesens nicht nur dem schwedischen, sondern auch dem niederländischen gegenübergestellt werden.

Bereits 1993/94 mussten das National Curriculum von 1988 und die Assessments einer Überprüfung und daran anschließend einer Revision unterzogen werden, mit ausgelöst durch Drohungen von Lehrerseite, die nationalen Assessments zu boykottieren. Zu entsprechenden Aufrufen kam es immer wieder auch im Hinblick auf die Publikation von Schulranglisten. Daraufhin wurde eine Reduktion des National Curriculum bezüglich seines Umfangs vorgenommen, zudem sollten 20% der Unterrichtszeit ins Ermessen der Einzelschule gestellt werden; auch wurden die aufwändigen Testverfahren auf die Kernfächer (Englisch, Mathematik, Naturwissenschaften) beschränkt.

New Labour bestätigte bei ihrem Regierungsantritt (1997) die Bildungspolitik der Ära Thatcher weitgehend und übernahm dabei die wichtigsten Elemente, nämlich die erweiterte Selbstständigkeit und Eigenverantwortung der Schulen, die Forderung nach Rechenschaftslegung und damit die Durchführung regelmäßiger landesweiter National Curriculum Assessments sowie die Sorge für ein effizientes Schulinspektionswesen; dies alles wiederum im Rahmen eines Kampfes um höhere Leistungsergebnisse (‚höhere Standards‘, wie man es in England ausdrückt). Insbesondere die bereits von der konservativen Regierung eingeleitete Stärkung der Konsumentenmacht und der Elternrechte durch umfassende und systematische Information der Öffentlichkeit über die Qualität der Schulen im Rahmen eines forcierten Wettbewerbs fanden ihre Verwirklichung unter New Labour. Das Anliegen eines Qualitätszuwachses der Schulen, und zwar mit besonderem Fokus auf die Diversifikation des Angebots zugunsten der Elternwahl, steht denn auch weiterhin im Zentrum des Strategieplanes (vgl. Department for Education and Skills 2004).

Interessant sind dabei aber gewisse Akzentverschiebungen im Steuerungskonzept seit den vergangenen zehn Jahren; diese können durchaus im Sinne von Korrekturmaßnahmen am ‚pressure‘-Ansatz, der viele Lehrkräfte in Opposition zur Politik gebracht hat, gedeutet werden. Dazu gehört die neuerliche Ausbalancierung von Kontrolle und Unterstützung der Schulen; das Angebot an Support und die Bemühungen um soziale Inklusion sind unter New Labour insgesamt gestiegen. Waren die Einzelschulen in der Vergangenheit zunehmend zu Objekten externer Evaluation geworden, kommt der Selbstevaluation inzwischen vermehrt Bedeutung zu; die Schule ist nun auch die In-

stanz, der die führende Rolle in der Qualitätsarbeit zugesprochen wird. Das Inspektionsaufkommen erlebt dagegen eine Reduktion, indem leistungsstarke Schulen seltener besucht werden; dies auch im Sinne des staatlichen Willens, bürokratischen Aufwand abzubauen (vgl. Department for Education and Skills 2004). Ebenfalls wurde die Funktion der LEAs erneut gestärkt, wenngleich deren Rolle eine neue Definition fand und sie ihre ehemalige Kontrollfunktion nicht wiedererlangt haben. Drittens zeichnet sich v.a. seit der zweiten Legislaturperiode eine versöhnlichere Haltung des Bildungsministeriums gegenüber schwachen Schulen in sozioökonomisch benachteiligten Gegenden und eine insgesamt positivere Darstellung der Leistungen der Lehrerschaft in der Öffentlichkeit ab (vgl. Kotthoff 2003, S. 104). Dazu gehören auch bis in die unmittelbare Gegenwart Bemühungen um immer weiter verfeinerte Methoden der Messung von Leistungen bzw. Leistungszuwächsen, um zusätzlichen Einbezug von Kontextdaten (Contextual Value Added, CVA) und die Herstellung fairer Vergleichsmöglichkeiten (vgl. Department for Education and Skills 2006).

### **3. Standards, Tests und Accountability im Spannungsfeld zwischen Politik und Profession**

Mit den wachsenden Steuerungsansprüchen der zentralen Bildungsbehörden ging in England seit den 80er Jahren das Erstarken markt- und betriebswirtschaftlicher Prinzipien einher. Die wichtigsten Akteure auf Seiten der Behörden sind deutlich ökonomisch ausgerichtet, während die Pädagogik inzwischen eine eher untergeordnete Position einnimmt. Interessenkonflikte gegenüber pädagogischen Ansprüchen der unterrichtenden Profession zeichnen sich dabei v.a. dort ab, wo es um Fragen des Testens geht. Eine ähnliche Konfliktlage zwischen Politik und Profession prägt die Diskussionen in den USA, gerade auch im Zusammenhang mit der Implementation von Standards und Accountability-Systemen. Zwei sich gegenüberstehenden Lagern gelingt es dabei, so der Eindruck, ihre Thesen mit je ‚eigener‘ Forschung zu untermauern, was oftmals ein Bild widersprüchlicher Befunde hinsichtlich des Erfolgs staatlich initiiert Reformen ergibt.

Ravitch (vgl. 2002) spricht diesbezüglich von zwei Paradigmen, dem *professional education paradigm* und dem *policymaker paradigm*. Während das zweite voraussetzt, dass das Schulsystem ebenso wie andere Organisationen verantwortlich für seine Leistungen gehalten und dabei über Mechanismen der Belohnungen und Sanktionen angetrieben werden muss, setzt das erste Paradigma auf ‚Lernen‘. Implementationsprozesse werden hier im Kontext von Theorien des professionellen und organisationalen Lernens gefasst, wobei die konstruktivistische Perspektive Prozesse des individuellen und kollektiven ‚sense-making‘ besonders hervorhebt (vgl. Spillane/Reiser/Reimer 2002). Demgegenüber gingen frühere Ansätze aus dem Feld der Innovations- und Change-Theorien bezüglich der Implementation von Reformen eher von einem ‚Produkt‘ aus, das es von der oberen Systemebene an die ‚Endnutzer/-innen‘ an der Basis zu bringen gilt. Dabei wird von Widerständen bei den ‚Abnehmern/-innen‘ ausgegangen, die unter einen behavioristischen Blickwinkel gestellt werden. Inzwischen hat sich aber im Bereich der

Theorien der policy-Implementation der Fokus zunehmend auf die Dynamiken und Prozesse verlagert, welche politische Maßnahmen und Verordnungen auf dem Weg durch die verschiedenen Ebenen und Instanzen des Systems prägen. Besonders die Rolle der Distrikte als lokale Kontexte der adaptierenden und reinterprezierenden Umsetzung staatlicher Reforminitiativen wurde ab Mitte der 90er Jahre Gegenstand vermehrter Aufmerksamkeit (vgl. z.B. Spillane 2004). Man bemerkt in den USA insgesamt seit den 90er Jahren eine gesteigerte Sensibilität gegenüber der Komplexität von Implementationsprozessen im pädagogischen Bereich, und es hat sich zunehmend ein Bewusstsein der Bedeutung der Mikroebene, also des Ortes, wo die Transformation der Reformideen in veränderte Lehr-/Lernformen stattfindet, sowie personaler Faktoren etabliert. Dieser neue bzw. zusätzliche Fokus wirkt sich nicht nur auf der Ebene von Konzepten aus, die Implementationsprozesse theoretisch zu fassen versuchen, sondern es hat sich auch gezeigt, dass Evaluationen von in Umsetzung begriffenen Reformen, die unter diesem Blickwinkel durchgeführt werden, ganz andere, d.h. pessimistischere Ergebnisse bezüglich Graden der Implementation ergeben. Knapp (vgl. 1995, 1997, 2002) hat dies sehr eindrücklich mit der Gegenüberstellung zweier Untersuchungen mit unterschiedlichem Evaluationsansatz zum selben Phänomen, nämlich der Implementation der kalifornischen Mathematik-Frameworks (eigentlich: Standards)<sup>5</sup> von 1985, aufgezeigt. Die von Knapp sogenannte ‚outside-in‘-Perspektive<sup>6</sup>, welche Marsh/Odden (vgl. 1991) angesetzt hatten, konnte die Reform als Erfolg präsentieren: Offenbar verwendeten die Lehrpersonen die neuen Frameworks in ihrem Unterricht. Zur Frage der methodischen und konzeptionellen Umsetzung der anspruchsvollen und innovativen, an den NCTM-Standards orientierten ‚frameworks‘ müssen jedoch ergänzend die Einzelstudien in Guthrie (vgl. 1990), die zum Umgang und den Erfahrungen von fünf Lehrkräften durchgeführt wurden, herangezogen werden. Diese beruhen auf intensiven Unterrichtsbeobachtungen (‚inside-out‘-Perspektive) und Befragungen der Lehrer/-innen. Sie zeigen eine komplexere Geschichte betreffend intendierter Effekte und tatsächlicher Wirkungen. In den geschilderten Fällen nahmen die Lehrkräfte einzelne Aspekte der Innovation auf, während sie in Vielem beim bereits Bekannten blieben. Dies ist auf ein begrenztes Verständnis dessen zurückzuführen, was die neuen ‚frameworks‘ von ihnen an Veränderungen ihres Unterrichts verlangten. Damit lässt sich auch erklären, warum trotz

5 Die Begriffe ‚curriculum frameworks‘ und ‚content standards‘ werden häufig synonym verwendet für Spezifikationen von staatlicher Seite, was Schüler/-innen können und wissen müssen. In Kalifornien stehen ‚frameworks‘ für ‚standards‘.

6 Knapp (vgl. 2002) rechnet zur outside-in-Perspektive insbesondere Ansätze, die Theorien der policy-Implementation verpflichtet sind, und systemische Strategien mit Fokus auf das alignment von Reformen. Was in diesen Designs weitgehend unberücksichtigt bleibe, sei ein spezifisches Verständnis von der Wirkung dieser Politiken auf das individuelle und kollektive Denken und Handeln von Lehrkräften. Diese Fragen bilden den Ausgangspunkt der inside-out-Perspektive, welche auf Forschungen zu Unterricht, Lehrerlernen und zur professionellen Entwicklung von Lehrkräften rekurriert und dabei Bezug nimmt auf kognitions- und sozialpsychologische Theorien, z.T. unter Einschluss einer soziokulturellen Perspektive. Im Zentrum steht nicht die politische Umgebung, sondern die/der individuelle Praktiker/-in und die Entwicklung der professionellen Identität und Kompetenz.

*alignment* zwischen ‚curriculum frameworks‘ und Unterrichtsmaterial die Wirkung der Reform dennoch begrenzt blieb.

Ein bedeutender, wenn nicht *der* Kristallisationspunkt für Auseinandersetzungen zwischen den beiden genannten Lagern bzw. Paradigmen sind Formen, Funktionen und Umgangsweisen mit Tests und Assessments<sup>7</sup>. In den USA entzündet sich die Diskussion v.a. an den High-Stakes-Tests. Zur Debatte stehen dabei die Qualität verwendeter Testsysteme, d.h. die Aussagekraft resultierender Ergebnisse, allfällige negative Auswirkungen und insbesondere die Fairness einer daran geknüpften Accountability von Schulen und Lehrpersonen. Die Frage der Auswirkungen von Tests ist im Zusammenhang mit der Implementation von Standards von eminenter Bedeutung, wie die Befunde einer Review von Lauer u.a. (vgl. 2005) zeigen. Demnach sind es v.a. die Tests und weniger ausgeprägt Standards und Curricula, die Entscheidungen der Lehrpersonen bezüglich Inhalten und Methoden des Unterrichts anleiten.

Systemische Reformansätze sehen in Assessments ein wirksames Mittel zur Herbeiführung von Schulreformen, da sie offensichtlich das Verhalten von Lehrpersonen an sich stark beeinflussen. Die Bandbreite der verwendeten Tests/Assessments ist in den USA vielfältig und reicht von (relativ kostengünstigen) Multiple-Choice-Verfahren bis zu ‚performance-based assessments‘, von standardisierten Large-Scale-Tests bis zu aufwändigen sogenannten ‚authentic assessments‘. Ein Streitpunkt in den Auseinandersetzungen zwischen den Vertreter/-innen beider genannter Paradigmen bezieht sich nun auf die Frage, welche Tests mit welchen Funktionen und Auswirkungen auf das Lehrer- und Schülerverhalten vertretbar sind; ob Assessments in erster Linie zur externen Kontrolle der Lehrkräfte dienen oder als integraler Bestandteil des Unterrichtshandelns ihr Lernen im Hinblick auf dessen Optimierung befördern sollen. Nur im letzteren Fall jedoch, so Vertreter/-innen des Lern-Paradigmas, kommt Assessments eine positive Bedeutung zu. Gemäß Darling-Hammond (vgl. 1994) haben testbasierte Entscheidungsprozesse etwa bezüglich der Selektion von Schüler/-innen zu einem Unterricht beigetragen, der die Vermittlung von ‚lower-order cognitive skills‘ ins Zentrum stellt. Die Selektionsmaßnahmen wiederum haben die Ungleichverteilung der Lerngelegenheiten für Schüler/-innen befördert. Den Hintergrund für den Aufsatz von Darling-Hammond zur Frage der Rolle von Assessments bot damals die Einführung neu entwickelter performance-basierter, ‚authentischer‘ Assessments, die den Anspruch erheben, das Equity-Postulat gegenüber den traditionelleren (Multiple-Choice-)Verfahren der 80er Jahre besser einzulösen. Darling-Hammond betont jedoch vor diesem Hintergrund, dass die zentralen Kriterien für die Beurteilung der Güte von Tests weniger in der Präzision von Selektionsentscheidungen oder der ‚Objektivität‘ im Hinblick auf die Legitimation von Strafen und Belohnungen liegen, sondern im Informationsgehalt der Tests für Lehrpersonen und ihrer Einbettung in Prozesse der Unterrichtsentwicklung. Die Frage hat sich

7 Glaser (vgl. 1990) nimmt folgende Unterscheidung zwischen Tests und Assessments vor: Testing zielt auf Selektion von Schüler/-innen und damit auf die Voraussage von Lernerfolgen vor der Platzierung z.B. in einer Klasse oder Schulform; Assessments eruieren outcomes und messen somit die Lernergebnisse.

schlussendlich immer darauf zu richten, ob Assessments Curriculum und Unterricht positiv beeinflussen; dies ist der Fall, wenn Assessments der Lehrperson Informationen über das Lernen der Schüler/-innen und die Effektivität des eigenen Unterrichts und damit Gelegenheiten des professionellen Lernens bieten. In Zusammenhang damit steht die Forderung, Lehrer/-innen müssten selbst in die Durchführung, Auswertung und Ergebnisinterpretation von Assessments einbezogen werden (vgl. Murnane/Levy 1996). Die notwendigen Kompetenzen für den Umgang mit Assessments im Sinne von Informations- und Diagnoseinstrumenten lassen sich allerdings nicht in separaten Kursen vermitteln, sondern müssen Bestandteil der unterrichtsmethodischen Ausbildung sein (vgl. Shepard 2000). Shepard geht bei ihrer Argumentation von (sozial-)konstruktivistischen Prämissen aus, die sie behavioristischen Annahmen diametral gegenüberstellt. Von Lehrer/-innen durchgeführte Assessments, Feedbacks von Kolleg/-innen und Ergebnisse von Schüler-Selbst-Assessments sind in dieser Sichtweise zentrale Elemente eines sozialen Prozesses, der die Entwicklung von intellektuellen Kompetenzen, die Konstruktion von Wissen und die Entwicklung der Identität der Schüler/-innen vermittelt.

Weder in den Niederlanden noch in Schweden stehen die nationalen Tests bzw. Abschlussprüfungen dermaßen in der Kritik wie in den USA. Das hat u.a. damit zu tun, dass die angeführten Postulate teilweise eingelöst sind. In beiden Ländern existieren diagnostische Tests, die von den Lehrkräften freiwillig und in Eigenregie durchgeführt werden können, es kommen verschiedene Arten von Assessments zum Einsatz,<sup>8</sup> die Lehrpersonen nehmen die Korrektur und Bewertung der Tests selbst vor (Schweden) bzw. die Schulen stellen im Rahmen der nationalen Abschlussprüfungen einen Teil der Aufgaben selbst (Niederlande).

#### **4. Capacity building oder: Von der Notwendigkeit von Inputs unter Bedingungen der Output-Orientierung**

Die von den Verfechter/-innen des Lern-Paradigmas präferierte Funktion von Assessments setzt das Vorhandensein von *capacity* voraus und zielt selbst wiederum – vermittelt über Prozesse des professionellen Lernens – auf den Aufbau von Kapazität (vgl. McLaughlin/Shepard 1995). Mit dem Begriff *capacity* bzw. ‚capacity building‘ ist auf ein Konzept verwiesen, das seit den 90er Jahren eine Schlüsselstellung in den US-amerikanischen Diskussionen um die Implementation von standardbasierten Reformen einnimmt. Verschiedene Untersuchungen hatten gezeigt, dass der Verlauf entsprechender Prozesse maßgeblich von den vor Ort, in den Distrikten und Schulen, vorhandenen Kapazitäten beeinflusst wird. Man kann sagen, dass sich in dem Begriff die Einsicht ab-

8 Dies fordern auch verschiedene amerikanische Wissenschaftsorganisationen: „Any decision about a student’s continued education, such as retention, tracking, or graduation, should not be based on the results of a single test, but should include other relevant and valid information“ (American Psychological Association 2001, S. 2).

bildet, dass bessere Leistungen im Schulsystem (Output) nicht mit der Setzung höherer Standards und der Einführung rigider Accountability-Mechanismen allein zu gewinnen sind, sondern dass es nach wie vor eines qualitativ hoch stehenden Inputs und gewisser Kontextbedingungen bedarf. *Capacity* steht sowohl für die notwendigen individuellen und organisationalen Ressourcen auf der Input-Seite wie auch für die Fähigkeiten und Möglichkeiten, diese in optimaler Weise im Unterricht ein- und umzusetzen (vgl. Newmann/King/Rigdon 1997; Ingram/Colby 1998).

Verschiedene Konzepte von *capacity* wurden inzwischen definiert: Während die einen das Gewicht v.a. auf die Kompetenzen und das Lernen der Lehrkräfte, etwa im Rahmen von Lerngemeinschaften (*learning communities*), legen (vgl. Corcoran/Goertz 1995), beziehen sich andere stärker auf Unterrichtsbedingungen, die motivationale Bereitschaften und das Vorhandensein qualitativ hoch stehender Unterrichtsmaterialien und -werkzeuge (*tools*) (vgl. Cohen/Ball 1996). Goertz/Floden/O'Day (1995) beziehen *capacity* auf die das Klassenzimmer umgebenden Kontexte, Strukturen und Organisationen (Schule, Distrikt, Staat) und attestieren diesen eine herausragende Bedeutung für die erfolgreiche Implementation standardbasierter Reformen. Massell (vgl. 1998) schließlich hat ausgehend von diesen Arbeiten Bereiche von *capacity* abgeleitet, die sowohl auf die Ebene des Unterrichts wie auf die organisationalen Bedingungen auf den unterschiedlichen Systemebenen abheben.

Die *capacity*-Frage erhielt in den USA gerade im Zuge der Durchführung systemischer Standards-Reformen Auftrieb. Die Output-Orientierung des Systems zusammen mit der Einführung bzw. Verlagerung von Verantwortung für die Leistungsergebnisse hatte in verschiedenen Staaten eine Art Vakuum auf der Input-Seite zur Folge (vgl. Massell 1998). Neue staatliche Standards und ‚*curriculum frameworks*‘ wurden nämlich oftmals eher allgemein formuliert, um den Distrikten und Schulen die detaillierte Ausführung selbst zu überlassen. Dazu verlangten die Distriktadministrationen und Lehrkräfte aber mehr Anleitung und Support, als die Staaten bereitstellten; diese sahen sich in der Folge aufgefordert, eine aktivere Rolle zu übernehmen und die Schulen und Distrikte etwa bei der Entwicklung standardkompatibler Curriculummaterialien, Unterrichtsprogramme und Aufgabenbeispiele zu unterstützen.<sup>9</sup> Eine weitere Einsicht betrifft die Ausrichtung formaler Unterstützungs- und Weiterbildungsangebote im Zusammenhang mit der Implementation von Standards. Diese erwiesen sich oftmals als zu wenig spezifisch und zu kurzfristig angelegt, um effektiv zu sein. Eine Studie aus New Jersey zeigt etwa, dass viele Lehrpersonen selbst nach intensiver, an den neuen Standards orientierter Weiterbildung Schwierigkeiten hatten, eine Lektion mit Aktivitäten zu planen, die auf einen selber ausgewählten ‚*content standard*‘ hinführen (vgl. O'Shea 2003). O'Shea schließt, dass das *alignment* von Weiterbildung mit Standards und Curriculum nicht ausreiche, wenn nicht Unterrichtsentwicklung bis auf die Ebene der kon-

9 Traditionell verfügen in den USA weder die Lehrer/-innen noch das Distriktpersonal vor Ort über eine besondere Expertise, um Curricula und Unterrichtsprogramme selber zu entwickeln, was sie in der Gestaltung des Unterrichts in eine gewisse Abhängigkeit von kommerziellen Lehrmittel- und Testanbietern gebracht hat.

kreten Lektionsplanung mit Fokus auf Lernergebnisse, die diese Standards einlösen, der Gegenstand ist: „(...) teachers need to learn how to plan a standards-based lesson that focuses on the achievement of outcomes related to the standards rather than simply conducting an activity aligned with standards. Furthermore, teachers need to learn how to evaluate student work samples in relation to the performance outcomes included in the standards-based lessons“ (O’Shea 2003, S. 16).

Gerade in Systemen, die auf eine Tradition relativ ausgeprägter Regulierung und Verwaltung zurückblicken, können die Kapazitäten, die unter Bedingungen vermehrter Autonomie und Verantwortung notwendig sind, nicht einfach vorausgesetzt bzw. in kürzesten Fristen ‚hergestellt‘ werden. Am Beispiel von Schweden, wo ein recht radikaler Wandel in Richtung Ergebnissteuerung und Dezentralisierung stattgefunden hat, hat sich gezeigt, dass Lehrpersonen die ihnen neu zukommenden Handlungsspielräume nicht ohne weiteres nutzen bzw. nutzen können; trotz veränderter Rahmenbedingungen erwiesen sich Arbeitsformen und die Organisation des Unterrichts als stabil. Zu rechnen ist auch mit Akzeptanzproblemen, wie sie in Schweden einerseits als Skepsis der Öffentlichkeit und besonders der Lehrerschaft gegenüber den Kompetenzen und der Arbeitsqualität der Kommunen bestanden; außerdem begegneten viele Lehrer/-innen der Rollendifferenzierung der Schulleitung mit Misstrauen (vgl. Kotthoff 2003; Daun 2004). Diese Probleme sind im Übrigen nicht spezifischer Art, sondern lassen sich ähnlich auch in Bezug auf England und die Arbeit der LEAs aufführen (vgl. Kotthoff 2003).

In Schweden lässt sich in jüngerer Vergangenheit eine leichte Tendenz zurück in Richtung zentraler Steuerung feststellen. „Allerdings haben inzwischen die Politiker auf Landesebene immer mehr ihr Vertrauen in kommunale Steuerungssysteme verloren“ – konstatieren Erickson/Lander (2007, S. 34). So fordere zum Beispiel der Staat auf der einen Seite die Schulen auf, eigene Konzepte und Pläne für bestimmte Bereiche zu entwickeln, gleichzeitig erlasse er aber ziemlich genaue Vorgaben, wie solche Konzepte (formal) aussehen sollen (vgl. ebd.). Dies betrifft etwa auch die von den Schulen zu erstellenden Qualitätsberichte, die zunehmend standardisiert werden – auch um Vergleiche zwischen den Schulen zu ermöglichen. Ein weiteres Zeichen setzt die neue Schulin-spektion innerhalb von Skolverket, die unter dem Leitsatz ‚Inspecting for Improving‘ ab 2003 regelmäßig sämtliche Schulen überprüft. Schließlich beabsichtigt die neue konservativ-liberale Regierung – nicht untypisch – die Anzahl der verbindlichen nationalen Tests und der getesteten Fächer zu erhöhen.

Demgegenüber scheinen die Niederlande, welche traditionell über ein ausgesprochenes Maß an Autonomie im Schulbereich verfügen, weiterhin auf dem Deregulierungs- und Dezentralisierungskurs zu bleiben. Zwei Drittel des Unterrichts sind hier durch allgemein gehaltene Kernziele, deren Zahl in der Vergangenheit stetig reduziert wurde, national ‚definiert‘. Umso größer ist die Rolle, die damit den schuleigenen Curricula zukommt; Projekte, die die Schulen bei deren Entwicklung unterstützen, gewinnen damit an Bedeutung (vgl. Nieveen u.a. 2006). Lehrlinien und Zwischenziele konkretisieren seit neuem zwar den Weg zur Erreichung der nationalen Ziele, dennoch kommt eine Studie zur Curriculumpolitik in neun Ländern (vgl. Kuiper u.a. 2006) zum Schluss, dass die Niederlande bezüglich Lehrplanautonomie weiterhin den Titel ‚the uncrowned

king“ (ebd., S. 75) beanspruchen dürfen. Schulen und Lehrpersonen seien weit davon entfernt, sich präskriptive Lehrpläne herbeizusehnen (ebd., S. 76). Viel eher wünschen sie zur Unterstützung der eigenen Lehrplanarbeit die Verbreitung prototypischer Beispiele schuleigener Lehrpläne im Sinne von ‚best practice‘.

## 5. Schluss

Die Vergangenheit lehrt, dass Reformprojekte in der Praxis oftmals nicht den erwarteten Erfolg gezeigt haben oder vom System ohne große Effekte absorbiert wurden, weil sie unvollständig implementiert worden sind, weniger also aufgrund von mangelhafter Konzeption. Dabei mag der ursprünglich aus den Ingenieurwissenschaften stammende und dann von der Computertechnologie übernommene Begriff der ‚Implementation‘ suggerieren, dass es sich hierbei um einen eindimensionalen Vorgang des ‚Einführens‘ einer Innovation handelt. Dass dem nicht so ist, dass es sich semantisch eher um Prozesse der ‚Umsetzung‘ handelt, die Lernen und Entwicklung voraussetzen, haben die aufgezeigten Erfahrungen hinlänglich gezeigt. Demnach ist das Ziel der Implementation von Bildungsstandards – die Etablierung einer neuen Steuerungspolitik verbunden mit der Hoffnung auf eine Hebung der schulischen Leistungen – ohne den Bezug auf Schul- und Unterrichtsentwicklung nicht zu erreichen; es wäre schließlich trügerisch zu glauben, bessere Ergebnisse in einem System lediglich durch die Änderung der Steuerungslogik herbeiführen zu können.

Der Grad der Implementation nationaler Bildungsstandards wird sich nicht zuletzt daran messen lassen müssen, inwieweit die mit ihnen verfolgten Zielsetzungen im Hinblick auf die Gestaltung von Lehr- und Lernprozessen von den anderen Systemkomponenten unterstützt und getragen werden. Neben der Frage der Tests ist damit besonders auch diejenige nach der vorhandenen *capacity* und den zur Verfügung stehenden Unterstützungssystemen von Bedeutung. Denn: Im Zuge einer ‚Umsteuerung‘ dürfen neu zugestandene Entscheidungsspielräume auf der lokalen Ebene nicht einseitig mit einem Mehr an Verantwortung und Rechenschaftspflicht ‚vergolten‘ werden; die Innovation ‚Bildungsstandards‘ dürfte auf diese Weise für die unterrichtende Profession kaum attraktiv sein und ‚Sinn‘ machen. Hier liegt zugleich ein sensibler Punkt im Sinne der Austarierung von regulativen Vorgaben, Lern- und Unterstützungsangeboten sowie professioneller Handlungsfreiräume auf dem Weg zur Erreichung vorgegebener Ziele.

Damit sind Prozesse angesprochen, die einen langen Atem brauchen und sich auch auf die Tradition eines Bildungswesens, seine je spezifischen kulturellen Voraussetzungen einstellen müssen. Dies wiederum macht deutlich, dass das ‚Lern-Paradigma‘ nicht gegen das ‚policy-Paradigma‘ auszuspielen ist. Vielmehr muss die Politik mit konsistenten Signalen und Impulsen längerfristig den für die intendierte Entwicklung günstigen Kontext bereitstellen – und gleichzeitig bereit sein, aufgrund von Feedbacks aus der Praxis und der begleitenden Forschung zu lernen.

## Literatur

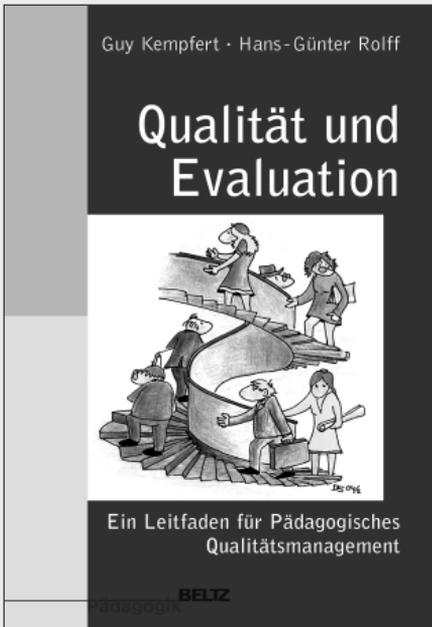
- American Psychological Association (2001): How should student learning and achievement be measured? URL: <http://www.apa.org/releases/testing.html> (Stand 12.08.08).
- Blank, R.K./Porter, A./Smithson, S. (2001): New tools for analyzing teaching, curriculum and standards in mathematics and science. Results from Survey of Enacted Curriculum Project. Final Report. Washington: Council of Chief State School Officers.
- Cohen, D.K./Ball, D.L. (1996): Capacity and building capacity for instruction. Paper presented at the annual meeting of the American Education Research Association. New York.
- Corcoran, T./Goertz, M.E. (1995): Instructional capacity and high performance schools. In: *Educational Researcher* 24, H. 9, S. 27–31.
- Covington, M.V. (1996): The myth of intensification. In: *Educational Researcher* 25, H. 8, S. 24–27.
- Darling-Hammond, L. (1994): Performance-based assessment and educational equity. In: *Harvard Educational Review* 64, H. 1, S. 5–30.
- Daun, H. (2004): What makes Swedish education what it is? Some facts and reflections. In: Fitzner, T. (Hrsg.): *Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform*. Bad Boll: Evangelische Akademie, S. 62–102.
- Department for Education and Skills (2004): Five year strategy for children and learners. URL: <http://www.dcsf.gov.uk/publications/5yearstrategy/docs/DfES5Yearstrategy.pdf> (Stand 12.08.08).
- Department for Education and Skills (2006): Publication of 2006 test and examination results in the school and college achievement and attainment tables. URL: <http://www.dcsf.gov.uk/performance-tables/Final-Decisions-on-Changes-to-the-Content-of-the-2006-Achievement-and-Attainment-Tables.pdf> (Stand 12.08.08).
- Ekholm, M. (2006): Curriculum in Sweden: Use and discussion. In: Letschert, J.F.M. (Hrsg.): *Curriculum development re-invented. Proceedings of the invitational conference on the occasion of 30 years SLO 1975-2005 Leiden, the Netherlands, December 7-9, 2005*. Enschede: SLO, S. 88–96.
- Erickson, G./Lander, R. (2007): Der Kitt, der ein wachsendes System zusammenhält? Nationale Tests als Kern der Qualitätssicherung in Schweden. In: *Pädagogik* 59, H. 3, S. 32–35.
- Glaser, R. (1990): *Testing and assessment: O tempora! O mores!* Pittsburgh: University of Pittsburgh, Learning, Research and Development Center.
- Goertz, M.E./Floden, R.E./O'Day, J.A. (1995): *Studies of systemic reform*. New Brunswick: Rutgers University, Consortium for Policy Research in Education.
- Guthrie, J.W. (Hrsg.) (1990): *Educational Evaluation and Policy Analysis* 12, H. 3.
- Ingram, D./Colby, J. (1998): Taking standards beyond the classroom door: A process for implementation. Paper presented at the annual meeting of the American Education Research Association. San Diego.
- Knapp, M.S. (1995): Education policy and the improvement of teaching: Two accounts of the early implementation of the California Mathematics Framework. Paper presented at the annual meeting of the American Educational Research Association. San Francisco.
- Knapp, M.S. (1997): Between systemic reform and the mathematics and science classroom: The dynamics of innovation, implementation, and professional learning. In: *Review of Educational Research* 67, S. 227–266.
- Knapp, M.S. (2002): *Understanding how policy meets practice: Two takes on local responses to a state reform initiative*. Seattle: University of Washington, Center for the Study of Teaching and Policy.
- Kotthoff, H.-G. (2003): *Bessere Schule durch Evaluation? Internationale Erfahrungen*. Studien zur International und Interkulturell Vergleichenden Erziehungswissenschaft, Bd. 1. Münster: Waxmann.
- Kuiper, W./Akker, J. v.d./Hooghoff, H./Letschert, J. (2006): Curriculum policy and school practice in an European comparative perspective. In: Letschert, J. (Hrsg.): *Curriculum development re-invented*. Enschede: SLO, S. 56–77.

- Lauer, P.A./Snow, D./Martin-Glenn, M./Van Buhler, R.J./Stoutemyer, K./Snow-Renner, R. (2005): The influence of standards on K-12 teaching and student learning. A research synthesis. Mid-continent Research for Education and Learning: URL: [http://www.mcrel.org/pdf/synthesis/5052\\_RS-InfluenceofStandards.pdf](http://www.mcrel.org/pdf/synthesis/5052_RS-InfluenceofStandards.pdf) (Stand 12.08.08).
- Lusi, S. (1997): The role of state departments of education in complex school reform. New York: Teachers College Press.
- Marsh, D.D./Odden, A.R. (1991): Implementation of the California Mathematics and Science Frameworks. In: Odden, A.R. (Hrsg.): Educational policy implementation. Albany: State University of New York Press, S. 219–240.
- Massell, D. (1998): State strategies for building capacity in education: Progress and continuing challenges. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania, Graduate School of Education.
- McLaughlin, M.W./Shepard, L.A. (1995): Improving education through standards-based reform: A report of the National Academy of Education Panel on Standards-Based Educational Reform. Stanford: National Academy of Education.
- Murnane, R.J./Levy, F. (1996): Teaching the new basic skills: Principles for educating children to thrive in a changing economy. New York: Free Press.
- National Council of Teachers of Mathematics (2000): Principles and standards for school mathematics. Reston: NCTM.
- Newmann, F.M./King, M.B./Rigdon, M. (1997): Accountability and school performance: Implications from restructuring schools. In: Harvard Educational Review 67, S. 41–74.
- Nieveen, N./Handelzalts, A./Akker v.d. J./Hooghoff, H./Krüger, J./Homminga, S./Berkien, H. (2006): Collaborative curriculum renewal as propelling force for school and teacher development. In: Letschert, J. (Hrsg.): Curriculum development re-invented. Enschede: SLO, S. 78–87.
- Oelkers, J./ Reusser, K. (2008): Qualität entwickeln, Standards sichern, mit Differenz umgehen. Berlin: BMBF.
- O’Shea, M.R. (2003): Implementing state academic standards in the classroom. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education. New Orleans.
- Porter, A. (2002): Measuring the content of instruction: Uses in research and practice. In: Educational Researcher 31, H. 7, S. 3–14.
- Ravitch, D. (2002): Testing and accountability, historically considered. In: Evers, M.W./Walberg, H.J. (Hrsg.): School accountability. Stanford: Hoover Press 2002, S. 9–21.
- Shepard, L.A. (2000): The role of assessment in a learning culture. In: Educational Researcher 29, H. 7, S. 4–14.
- Slenning, K. (2004): Managing educational standards in a decentralised organisation: Critical points of transition. In: Fitzner, T. (Hrsg.): Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform. Bad Boll: Evangelische Akademie, S. 54–61.
- Smith, M.S./O’Day, J.A. (1990): Systemic school reform. In: Fuhrman, S.H./ Malen, B. (Hrsg.): The politics of curriculum and testing: The 1990 yearbook of the politics of education association. London: Taylor and Francis.
- Smith, M.S./O’Day, J.A. (1993): Systemic reform and educational opportunity. In: Fuhrman, S.H. (Hrsg.): Designing coherent education policy. San Francisco: Jossey-Bass, S. 250–312.
- Spillane, J.P. (2004): Standards deviation. How schools misunderstand education policy. Cambridge: Harvard University Press.
- Spillane, J.P./Reiser, B.J./Reimer, T. (2002): Policy implementation and cognition: Reframing and refocusing implementation research. In: Review of Education Research 72, S. 387–431.

*Anschrift der Autoren:*

Dr. des. Esther Berner/Prof. Dr. Jürgen Oelkers/Prof. Dr. Kurt Reusser, Pädagogisches Institut, Universität Zürich, Freiestrasse 36, CH-8032 Zürich.

# Pädagogisches Qualitätsmanagement



Guy Kempfert/Hans-Günter Rolff  
**Qualität und Evaluation**

Ein Leitfaden für Pädagogisches  
Qualitätsmanagement.

4. Auflage 2005.

286 Seiten. Broschiert.

EUR 24,90 D.

ISBN 978-3-407-25360-6

Dieses Arbeitsbuch zeigt, wie jede einzelne Schule pädagogische Qualitätsentwicklung in Gang setzen und vervollkommen kann. Dabei steht der Unterricht im Mittelpunkt. Hinzu kommt die Verbesserung der Zusammenarbeit in allen Gremien

von der Fachkonferenz bis zur Schulleitung. Dass all dies Einfluss auf die Schulkultur hat, versteht sich (fast) von selbst. Die zahlreichen Checklisten, Fragebögen und Kopiervorlagen erleichtern den sofortigen Einstieg.

»Das praxisnahe Buch ist klar und übersichtlich aufgebaut und enthält neben den leicht nachvollziehbaren Beschreibungen von Methoden und Instrumenten eine Vielzahl von Beispielen, die auch als Kopiervorlagen zu verwenden und in der eigenen Praxis leicht einsetzbar sind.«

*Friedrich Jahresheft 2001*

»Dieses praxisorientierte Arbeitsbuch greift die aktuelle Debatte um Schulqualität auf und unterbreitet eine Fülle konkreter Vorschläge, wie die Qualität von Unterricht, Lernleistungen und Abschlüssen gesichert werden kann.«  
*Förderschulmagazin*

»Kollegien, die anfangen wollen, die Qualität ihrer Schule zu verbessern, finden dazu in diesem Buch viele Anknüpfungspunkte und Hilfen.«  
*Neue Deutsche Schule*

**BELTZ**

Beltz Verlag · Weinheim und Basel · Weitere Infos: [www.beltz.de](http://www.beltz.de)

## Erziehungswissenschaft im internationalen Kontext



### **Handwörterbuch Erziehungswissenschaft**

2009. Ca. 944 Seiten. Gebunden.  
ISBN 978-3-407-83159-0

Eine aktuelle Darstellung der relevanten Begriffe der Erziehungswissenschaft, bestens geeignet zur Einarbeitung und Auffrischung von Grundlagenwissen. Führende Experten aus dem In- und Ausland stellen die wesentlichen Begriffe dar, maßgeblich ist jeweils der internationale Forschungsstand.

Dieses Wörterbuch geht erstmals über die nationale Perspektive hinaus und berücksichtigt auch die jüngsten

Entwicklungen in der Erziehungswissenschaft:

- das internationale Zusammenwachsen der Bildungsforschung,
- den üblich gewordenen Vergleich nationaler Bildungssysteme,
- die Berücksichtigung der Themen und Problemlagen der Systementwicklung.

Die über 60 Stichwörter erschließen das Spektrum der modernen Erziehungswissenschaft im internationalen Kontext. Zu diesem Zweck wurden die Autorinnen und Autoren allein nach ihrer in Fachzeitschriften ausgewiesenen Kompetenz ausgewählt; sie gehören keiner bestimmten Richtung oder Schule an. So ist ein Nachschlagewerk entstanden, das die Problemlagen, Befunde und Fragestellungen der verschiedenen Bereiche der Erziehungswissenschaft darstellt und zugleich einen raschen Überblick ermöglicht.

Herausgegeben von  
Sabine Andresen, Rita Casale,  
Thomas Gabriel, Rebekka Horlacher,  
Sabina Larcher Klee und  
Jürgen Oelkers.

### **Sparen Sie mehr als 30 Euro!**

Subskriptionspreis bis 17.03.2009:  
EUR 98,- D, danach EUR 139,- D.

# **BELTZ**

Beltz Verlag · Weinheim und Basel · Weitere Infos: [www.beltz.de](http://www.beltz.de)

## Aus dem Inhalt

**Administration** (Ulrich Otto) · **Begabung** (Roland H. Grabner/Elsbeth Stern) · **Beratung** (Volker Kraft) · **Beruf und Berufsbildung** (Katrin Kraus) · **Bildung** (Sabine Andresen) · **Bildungsfinanzierung** (Heiner Barz) · **Care** (Nel Noddings) · **Citizenship** (Anne Sliwka) · **Curriculum und Lehrmittel** (Rudolf Künzli) · **Curriculumentwicklung** (William F. Pinar) · **Demografische Entwicklung** (Michael Winkler) · **Demokratie** (Terri S. Wilson/David T. Hansen) · **Differenz** (Paul Mecheril/Melanie Plöber) · **E-Learning** (Damian Miller) · **Entwicklung** (Tobias Krettenauer) · **Erwachsenenbildung und Weiterbildung** (Carola Iller) · **Erziehung** (Jürgen Oelkers) · **Europa und europäischer Bildungsraum** (Georg Krücken/Christine Musselin) · **Evaluation und Qualitätssicherung** (Sandra Mittag/Hans-Dieter Daniel) · **Exzellenz** (Michael Hartmann) · **Familie** (Sabine Walper/Eva-Verena Wendt) · **Frühkindliche Erziehung** (Cornelia Wustmann) · **Gender** (Bettina Dausien/Christine Thon) · **Generationen** (Jutta Ecarius) · **Gesellschaft** (Dirk Tänzler) · **Gewalt** (Corinna Seith) · **Globalisierung** (Edgar Forster) · **Historische Methoden** (Rebekka Horlacher) · **Interkulturalität und interkulturelle Bildung** (Cristina Allemann-Ghionda) ·

**Intervention und Prävention** (Erik J. Knorth/Jana Knot-Dickscheit/Johan Strijker) · **Jugend** (Wolfgang Schröer) · **Kindheit** (Helga Kelle) · **Körperlernen** (Birgit Althans) · **Kompetenz** (Katharina Maag Merki) · **Kreativität** (R. Keith Sawyer) · **Kultur** (Rita Casale) · **Lebensalter und Lebenslauf** (Simone Scherger) · **Lehrerbildung** (Sigrid Blömeke) · **Leistung und Leistungsmessung** (Wilfried Bos/Andreas Voss/Martin Goy) · **Lernen** (Nicole Becker) · **Literalität** (Ewald Kiel) · **Ökonomie** (Stefan Wolter) · **Ökonomisierung** (Frank-Olaf Radtke) · **Organisation** (Sabina Larcher Klee) · **Parenting** (Bettina Grubenmann) · **Problemlösen** (Gert Biesta) · **Professionalisierung und Professionalität** (Catrin Heite/Fabian Kessler) · **Qualitative Methoden** (Barbara Frieberthshäuser) · **Quantitative Methoden** (Martin Wellenreuther) · **Recht und Verrechtlichung** (Stefan Weyers) · **Religion** (Hans-Georg Soeffner/Christine Matter) · **Schule** (Hans Döbert) · **Soziale Arbeit** (Thomas Gabriel) · **Sozialisation** (Michael-Sebastian Honig) · **Sozialpolitik** (Holger Ziegler) · **Standards** (Josef Abs) · **Steuerung** (Botho von Kopp) · **Struktur der Erziehungswissenschaft** (Heinz-Elmar Tenorth) · **Test** (Urs Moser) · **Unterricht** (Kurt Reusser) · **Wissen** (Thomas Höhne)

**BELTZ**

Beltz Verlag · Weinheim und Basel · Weitere Infos: [www.beltz.de](http://www.beltz.de)

## Besser unterrichten durch kollegiales Feedback



Guy Kempfert  
**Kollegiale Unterrichtsbesuche**  
Besser und leichter unterrichten  
durch Kollegen-Feedback.  
Mit Werkstatt-Teil, auch zum  
Download im Internet.  
2008. 157 Seiten. Broschiert.  
EUR 29,90 D.  
ISBN 978-3-407-25459-7

Seit Jahrhunderten sind Lehrer/innen Einzelkämpfer, Unterricht findet hinter verschlossenen Türen statt. Dabei sind kollegiale Unterrichtsbesuche mit anschließendem Feedback eine einfache und wirksame Methode, leichter und besser zu unterrichten. Das Buch zeigt unterschiedliche Methoden der kollegialen Unterrichtshospitalation und bietet praxiserprobte Beispiele und Anleitungen. Schulleitungen finden praktische Hilfestellungen für ein Gesamtkonzept, von dem nicht nur die Unterrichtsqualität profitiert, sondern auch die Teamentwicklung und Feedback-Kultur. Mit Werkstatt-Teil und einem neu entwickelten Kompetenzraster, das Indikatoren für alle unterrichtsrelevanten Kompetenzen bereitstellt – auch zum Download im Internet (Code im Buch).

Der erste deutsche Titel zum Thema.

# **BELTZ**

Beltz Verlag · Weinheim und Basel · Weitere Infos: [www.beltz.de](http://www.beltz.de)

# Schulen evaluieren sich selbst



Dietlinde Granzer/Peter Wendt/  
Regine Berger

## **Selbstevaluation in Schulen**

Theorie, Praxis und Instrumente.

2008. 160 Seiten. Broschiert.

EUR 39,90 D.

ISBN 978-3-407-25482-5

Selbstevaluation ist eine überzeugende Antwort auf die Kritik an Schule: Sie dokumentiert, was in der Schule nachweislich geleistet wird – und hilft so, die Qualität der schulischen Arbeit zu steigern.

Deswegen wird Selbstevaluation in Schulen – bis vor kurzem kaum denkbar – heute immer selbstverständlicher.

Dieses Buch gibt dazu eine Starthilfe. Es zeigt:

- theoretische Zugänge zum Thema
- die ersten Schritte zur Selbstevaluation
- die Rolle der Schulleitung
- praktische Anwendungen
- konkrete Beispiele
- Tipps zur professionellen Erstellung von Fragebögen

# **BELTZ**

Beltz Verlag · Weinheim und Basel · Weitere Infos: [www.beltz.de](http://www.beltz.de)

## Der erste umfassende Überblick



Franz Schultheis/Pasqualina  
Perrig-Chiello/Stephan Egger

### **Kindheit und Jugend in der Schweiz**

2008. 220 Seiten. Broschiert.

EUR 29,90 D.

ISBN 978-3-407-25498-6

Der Bericht bietet einen umfassenden Überblick zur Lage von Kindheit und Jugend in der Schweiz. Die Bestandsaufnahme des Nationalen Forschungsprogramms, für die mehr als 3000 Kinder und Jugendliche befragt wurden, hat eine intensive Diskussion in der Schweiz angestoßen.

# **BELTZ**

Beltz Verlag · Weinheim und Basel · Weitere Infos: [www.beltz.de](http://www.beltz.de)

Das Buch liefert nicht nur aktuelle Befunde zu den konkreten Lebensverhältnissen: von der Krippe bis zum Berufsbeginn, von Gameboy bis Sportverein. »Kindheit und Jugend in der Schweiz« beschreibt die Sozialisation im familiären und institutionellen Kontext – und beleuchtet dabei auch die Schattenseiten wie Kinderarmut, Bildungsversagen und die Schicksale eines früh defizitären Lebens.

»Wer Bernhard Buebs Bestseller ›Lob der Disziplin‹ im Regal stehen hat, darf das Büchlein getrost entsorgen: Das Nationale Forschungsprogramm ›Kindheit, Jugend und Generationenbeziehungen im gesellschaftlichen Wandel‹ (NFP 52) weist den deutschen Pädagogen und dessen strenge Erziehungsmethoden in die Schranken.

Der vorliegende Bericht erfüllt die Forderung der Uno, repräsentative Daten über die Lebensumstände von Heranwachsenden zu erheben. Zudem liefert er den politischen Entscheidungsträgern Grundlagen zur Weiterentwicklung der Schweizer Kinder- und Jugendpolitik.«

*Die Südostschweiz*