

Hermann Josef Abs/Peter Döbrich/Alexander Wicker

Benchmarking im Bildungswesen

Benchmarking scheint heute allgegenwärtig¹. Ob jedoch der zahlreichen Verwendung des Begriffs auch die zahlreiche Anwendung des entsprechenden Verfahrens entspricht, oder ob der Begriff auch gewählt wird, weil sich mit ihm die Aura des Modernen, des Ökonomischen, verbindet, ist eine andere Frage. Die Inflation des Begriffs deutet darauf hin, dass Benchmarking als Antwort auf drängende Gegenwartsprobleme aufgefasst wird, etwa jenes, in Zeiten knapper Budgets verbesserte Systemleistungen erzeugen zu müssen. Um in einer solchen Situation die alleinige Verwendung des modischen Begriffs von der Anwendung des Steuerungsinstruments unterscheiden zu können, erscheint es sinnvoll, sich zunächst die Frage nach der ursprünglichen Idee des Benchmarking und dem entsprechenden Vorgehen zu stellen (1.). Danach sollen aktuelle Übertragungen ins Bildungssystem aufgezeigt werden (2.), um diese Transferprozesse zu bewerten (3.). Abschließend wird die Frage nach der Steuerungskraft des Ansatzes Benchmarking aufgeworfen (4.).

1 Ursprung und Idee des Benchmarking

Benchmarking ist ein Begriff aus dem betriebswirtschaftlichen Controlling. Das Controlling als zentrale Managementfunktion kann unterschiedlich weit gesteckte Aufgaben haben, vom Aufbereiten von Führungsinformationen zu Kennziffern des Unternehmens über das Aufzeigen von Chancen und Risiken verschiedener Handlungsoptionen bis zur Unternehmenszielformulierung. Es kann somit in einem engen Verständnis lediglich zur Informationsbereitstellung genutzt werden und damit eine eher operative Führungsunterstützung sein; es kann aber auch auf dem Wege der Mitentscheidung strategische Bedeutung erlangen.

Die Entwicklung des Benchmarking ist eng mit der *Xerox Corporation* verknüpft, einem Hersteller von Kopiergeräten aus dem US-Bundesstaat Connecticut. Der Marktanteil von *Xerox* war in den Sechziger- und Siebzigerjahren von ca. 80 % auf ca. 40 % dramatisch geschrumpft, weshalb 1979 das firmenweite Programm „*Competitive Benchmarking*“ aufgesetzt wurde. Der japanische Mitbewerber *Canon* konnte zu diesem Zeitpunkt ein vergleichbares Produkt zu einem Preis anbieten, der den Herstellungskosten bei *Xerox* entsprach. Zunächst wandte man ein Verfahren an, das „*Reverse Product Engineering*“ genannt wird. Das Konkurrenzprodukt wurde gekauft, in seine Einzelteile zerlegt und die Komponenten mit einander verglichen. Dieser sehr detaillierte, aber rein produktbasierte Vergleich lieferte kein aussagekräftiges Ergebnis, weshalb man bei *Xerox* dazu überging, die einzelnen Wertschöpfungsketten, also Prozesse, näher zu betrachten. Das Ergebnis waren Probleme vor allem im Logistik- und Vertriebsbereich. So wurde außerhalb des Unternehmens nach Lösungen gesucht, die man schließlich bei einem Nichtkonkurrenten fand, dem Sportartikelversand *L.L. Bean*. Mit diesem Unternehmen vereinbarte man

¹ So wird z.B. PISA als Benchmarking dargestellt (Prenzel, M./Baumert, J./Blum, W. et al., PISA 2003 Ergebnisse des zweiten internationalen Vergleichs. Zusammenfassung. http://pisa.ipn.uni-kiel.de/Ergebnisse_PISA_2003.pdf S. 3.). Weiterhin sind Benchmarks ein zentrales Element der EU-Lissabon Strategie (s.u.). Aber auch das Institut der deutschen Wirtschaft nennt seinen Vergleich von Bildungssystemen Benchmarking (Klös, H.-P./Weiß, R. (Hrsg.), *Bildungsbenchmarking Deutschland*. Was macht die Effizienz eines Bildungssystems aus? Köln 2003). Schließlich erscheint der Begriff auch als ein zentrales Konzept im Bericht der österreichischen Zukunftskommission (Haider, G. Eder, F./Specht, W./Spiel, C./Wimmer, M., Abschlussbericht der Zukunftskommission, Wien 2005. <http://www.bmbwk.gv.at/schulen/zukunftskommission.xml>).

intensiven Datenaustausch und gegenseitige Besuche, die gewonnenen Erkenntnisse konnten dazu beitragen, als erfolgreich identifizierte Prozesse bei *Xerox* zu implementieren.²

1.1 Definition von Benchmarking

Die Etymologie erschließt ein erstes Verständnis der Bedeutung von Benchmarking: „Abgeleitet von dem Begriff Benchmark, der in der Landvermessung eine Markierung für Höhen- oder Richtungsvergleiche bedeutet, wird der Leistungsvergleich von Organisationen allgemein als Benchmarking bezeichnet.“³ Verlässt man jedoch die Anschaulichkeit des Vergleichs und versucht den Begriff konzeptuell zu fassen, tun sich Schwierigkeiten auf. *Böhnert* und *Straub*⁴ zitieren zusammen an die 20 Definitionen, die jeweils unterschiedliche Aspekte hervorheben.

Mit der Vielzahl von Definitionen kann man entweder spielerisch umgehen, wie *Michael J. Spendolini* das tut, wenn er ein ganzes „Benchmarking-Definitions-Menü“ anbietet⁵, aus dem sich der Leser die für das eigene Projekt passende Definition aus Satzbausteinen zusammenfügen kann. Man kann aber auch systematisierend vorgehen, um das Gemeinsame einer Mehrheit von Autoren als abstrahiertes Ergebnis zu ermitteln. Diesen Weg wählt *Böhnert*⁶ und schlägt eine aus der Zusammenschau gewonnene Definition für „Benchmark“ und „Benchmarking“ vor, die einerseits weit genug ist, um den Begriff nicht für ein isoliert betrachtetes, eng umschriebenes Vorgehen zu reservieren, und andererseits präzise genug um die völlige Beliebigkeit der Begriffsverwendung noch erkennen zu können. Demnach wird „Benchmark [...] eine Maßzahl genannt, die es ermöglicht, die Leistung kritischer Erfolgsfaktoren abzubilden und zwischen zwei Organisationseinheiten (Entitäten) zu vergleichen. Die höchste Ausprägung dieser Maßzahl wird als Richtwert zur Steigerung der eigenen Leistungsfähigkeit benutzt.“ Benchmarking, führt er weiter aus, „steht für den kontinuierlichen, systematischen Prozess, mittels Messung, Vergleich und Analyse geeigneter Benchmarks Strategien, Prozesse/Funktionen, Methoden/Verfahren oder Produkte/Dienstleistungen einer Organisationseinheit zum Zwecke der Sicherung oder Steigerung des Unternehmenserfolges zu verbessern.“⁷

In dieser wie in den meisten Definition wird Benchmarking als *Prozess* verstanden, das heißt als Abfolge von Entscheidungen, die jeweils einen neuen Entscheidungsraum eröffnen. Einen weiteren essentiellen Bestandteil bildet der *Vergleich*, dem sich ein an einem Benchmarking-Prozess interessiertes Unternehmen stellen muss. Verglichen werden jeweils *operationalisierte Messgrößen* (Benchmarks), die einen bestimmten Bereich des Systems abbilden. Ziel ist dabei die Identifikation von *erfolgreichen Funktionszusammenhängen* (Successful Practices⁸). Bisweilen wird auch von Best Practices gesprochen.⁹ Es gibt keine Einschränkung bei der Auswahl

² Nach *Camp, R.C.*, Benchmarking, München/Wien 1994; *Straub, R.*, Benchmarking. Eine Darstellung des Benchmarking als modernes Instrument zur Leistungsverbesserung, Dissertation, Zürich 1997, S. 44 ff.

³ *Camp, R.C.*, Benchmarking, München/Wien 1994, S. XI.

⁴ Vgl. *Straub, R.*, Benchmarking. Eine Darstellung des Benchmarking als modernes Instrument zur Leistungsverbesserung, Dissertation, Zürich 1997, S. 47 f.

⁵ *Spendolini, M.J.*, The Benchmarking Book, New York 1992, S. 10.

⁶ *Böhnert, A.-A.*, Benchmarking. Charakteristik eines aktuellen Managementinstruments, Hamburg 1999, S. 7 ff.

⁷ *Böhnert, A.-A.*, Benchmarking. Charakteristik eines aktuellen Managementinstruments, Hamburg 1999, S. 14.

⁸ Vgl. *Bodmer, C.*, Die Methode Benchmarking, Kap. 1.2, in: Fahmi, F./Völker, R./Bodmer, C., Erfolgreiches Benchmarking in Forschung und Entwicklung, Beschaffung und Logistik, München/Wien 2002, S. 3–25, hier S. 6.

⁹ Die Rede von best practice impliziert jedoch fälschlicherweise, dass eine bestimmte Praxis unabhängig von ihrem kulturellen Hintergrund und ihren Kontextbedingungen (weltweit) als beste bezeichnet werden kann. Während eine solche Denkfigur im Bereich der Betriebswirtschaft schon problematisch ist, weil auch dort z. B. unterschiedliche Ressourcen an verschiedenen Standorten berücksichtigt werden müssen, erscheint sie im Bereich der Bildungspolitik, wo auch die Frage nach den Nutznießern einer bestimmten Praxis Gegenstand von Politik ist, in ihrer Wertgebundenheit offensichtlich.

des Benchmarking-Objekts, d.h. des zu untersuchenden Teilbereichs. Sämtliche die Wertschöpfungskette direkt oder indirekt beeinflussende Aspekte können einem Benchmarking unterworfen werden.

1.2. Arten des Benchmarking

Ähnlich zahlreich wie die Definitionen sind die Arten des Benchmarking.¹⁰ Es finden sich Systematisierungen nach Benchmarking-Objekten (*Was* soll einem Benchmarking unterzogen werden?), nach Benchmarking-Zielen (*Weshalb* soll ein Benchmarking durchgeführt werden?), oder nach Vergleichsentitäten (*Womit* soll verglichen werden?). Weil jeder Funktionsbereich zum Gegenstand von Benchmarking werden kann und das Ziel letztlich immer in einer Verbesserung von Qualität besteht, folgen die beiden ersten Systematisierungsansätze klassischen Diskursen darüber, was Funktionsbereiche eines Unternehmens sind und worin Qualität besteht.¹¹ Stärker spezifisch ist eine Systematisierung, die sich aus den unterschiedlichen Modi des Vergleichs herleitet, wie sie im Rahmen eines Benchmarking zur Anwendung kommen können.

- **Internes Benchmarking** vergleicht eine Organisationseinheit des eigenen Unternehmens mit einer anderen Organisationseinheit desselben Unternehmens, möglicherweise verschiedene Unternehmen desselben Konzerns
- **Externes Benchmarking** betrachtet dagegen die eigene Organisationseinheit und vergleicht sie mit einer anderen Organisationseinheit eines anderen Unternehmens

Innerhalb des externen Benchmarkings wird wiederum zwischen folgenden vier Unterarten unterschieden:

- **Konkurrenzbezogenes Benchmarking** als Vergleichsentität wird eine Organisationseinheit eines Konkurrenzunternehmens gewählt, wobei dies entweder verdeckt durch die Auswertung öffentlich zugänglicher Informationen oder offen durch Gewinnung des Vergleichsunternehmens als Partner im Benchmarking-Prozess geschehen kann. Offenes konkurrenzbezogenes Benchmarking hat die Stärke, besonders aussagekräftig zu sein, aber gleichzeitig die Schwäche, dass sich der Prozess der Partnerfindung als äußerst schwierig erweisen kann. Konkurrenzunternehmen sind naturgemäß eher zurückhaltend, was die Bereitstellung u.U. sensibler Daten angeht. Aus diesem Grund hat sich ein Markt für Clearinghouse-Stellen und Unternehmensberatungen entwickelt, die als „neutrale Dritte“ zwischen Benchmarking-Interessenten dergestalt vermitteln können, dass sie die Informationen beider Seiten sammeln, anonymisieren und so aufbereitet der jeweils anderen Seite zur Verfügung stellen können, ohne dabei ein einschlägiges Eigeninteresse zu verfolgen. Weiterhin können sie Datenbanken vorhalten, die die entsprechenden Informationen eines oder mehrerer Unternehmen sammeln und anderen wieder zugänglich machen können.
- **Branchenbezogenes Benchmarking** hier stammt die Vergleichsentität zwar aus der gleichen Branche, aber stellt keinen direkten Konkurrenten um das gleiche Marktsegment dar.

¹⁰ Vgl. Kempf, S./Siebert, G., Klassifizierendes Benchmarking – ein neuer Ansatz, Kap 9, in: Mertins, K./Siebert, G./Kempf, S. (Hrsg.), Benchmarking. Praxis in deutschen Unternehmen, Heidelberg u.a. 1995, S. 125–141.; Böhner, A.-A., Benchmarking. Charakteristik eines aktuellen Managementinstruments, Hamburg 1999, S. 15 ff.; Leibfried, K.H.J./McNair, C., Benchmarking. Von der Konkurrenz lernen, die Konkurrenz überholen, Freiburg i.Br. 1993, S. 45 ff.; Straub, R., Benchmarking. Eine Darstellung des Benchmarking als modernes Instrument zur Leistungsverbesserung, Dissertation, Zürich 1997, S. 53 ff.

¹¹ Vgl. Harvey, L./Green, D., Qualität definieren. Fünf unterschiedliche Ansätze, in: Zeitschrift für Pädagogik 41, Beiheft 2002, S. 17–39.

Es dürfte hier zwar einfacher fallen, *überhaupt* einen Benchmarking-Partner zu finden. Gleichzeitig steigt aber der Schwierigkeitsgrad, den *richtigen* Benchmarking-Partner mit der entsprechenden Successful Practice zu identifizieren, was wiederum das Verbesserungspotenzial geringer erscheinen lässt.

- **Generisches Benchmarking** überwindet die Grenzen der eigenen Branche, es kann mit jedem anderen Unternehmen durchgeführt werden, das auf dem zu untersuchenden Gebiet Spitzenlösungen implementiert hat. Eine generische Benchmarking-Studie ist einerseits in der Durchführung einfacher, weil man außerhalb der eigenen Branche noch leichter auf die Bereitschaft zum Austausch trifft, aber andererseits auch schwieriger, weil die Identifikation eines geeigneten Partners aufwendiger ist. Von allen Typen gilt dieser als der mit dem in der Regel höchsten Innovationspotenzial und den größten Transferschwierigkeiten.
- **Globales Benchmarking** gilt gleichsam als die Königsklasse der Disziplin, weil als Vergleichsentität weltweit nach der Best Practice gesucht wird. Diese letzte Art des Benchmarkings überlagert sich mit den Vorgenannten. Zugleich muss die implizierte Behauptung, dass es immer *eine* beste Praxis gebe, und diese in das anders strukturierte Zielunternehmen übertragbar sei, hinterfragt werden.¹²

Die Reihenfolge der Vergleichsmodi wurde hier so gewählt, dass im Verlauf der Abfolge einerseits das mögliche Innovationspotenzial und andererseits aber auch die Kosten für die Informationsbeschaffung und Transfer in das eigene System ansteigen. Ob diese Kosten jeweils in einem günstigen Verhältnis zum erwarteten Innovationsertrag stehen, ist in der Betriebswirtschaft eine Berechnungsaufgabe des Controllings. In der Bildungspolitik wird man zunächst danach fragen, was der angezielte Ertrag eines Bildungssystems sein soll, um dann entsprechende Gegenstände und Vergleichsentitäten auszuwählen.

1.3 Vorgehen beim Benchmarking

Beschreibungen des Vorgehens beim Benchmarking sind meist nach dem *PDCA-Modell* (**P**lan **D**o **C**heck **A**ct) nach *William Edwards Deming*¹³ strukturierbar. Dieser aus dem Bereich des Qualitätsmanagements bekannte kreisförmige Prozess beschreibt das Grundprinzip der kontinuierlichen Verbesserung, bei dem in der ersten Phase Ziele definiert und Abläufe beschrieben werden („Plan“), die wiederum danach ausgeführt („Do“) und überwacht werden („Check“, ursprünglich auch „Study“), bevor die Ergebnisse dieser Überwachung im Idealfall als Regelverfahren übernommen werden („Act“).

Vor diesem Hintergrund verwundert es auch nicht, wenn Böhnert, der sich die Arbeit gemacht hat, 14 verschiedene Phasenmodelle darzustellen, am Ende seines Versuchs, ein Basismodell herauszufiltern, eine Erweiterung zu Deming vorschlägt. Es werden vor allem „Phasen der Kommunikation“ ergänzt, die er zwischen die anderen Phasen (die bei ihm „Planung“, „Datenerhebung“, „Analyse“, „Implementierung“ und „Kontrolle“ heißen) schaltet, um Ergebnisse zu sichern und den Prozess immer wieder neu zu justieren, eventuell sogar eine weitere Schleife durchlaufen zu lassen.¹⁴

¹² Vgl. *Patton, M.Q.*, Evaluation, knowledge management, best practice and high quality lessons learned, in: *American Journal of Evaluation* 22, 2002, S. 329–336.

¹³ *Deming* (1900–1993), Ingenieur und Mathematiker, der seit den Fünfziger- und Sechzigerjahren zunächst in Japan, später in den Vereinigten Staaten eine neue Unternehmensphilosophie begründete, die in Abkehr vom Taylorismus die Bedürfnisse des Menschen, der Hersteller und der Kunden ins Zentrum stellt.

¹⁴ *Böhnert, A.-A.*, Benchmarking. Charakteristik eines aktuellen Managementinstruments, Hamburg 1999, S. 129 ff.

2 Übertragungen von Benchmarking auf das Bildungssystem

Der Bildungssektor ist in wesentlichen Teilen bis heute als „Non-profit-Bereich“ organisiert. Daraus folgt, dass seine Funktionserfüllung zunächst nicht in der Erwirtschaftung eines monetären Profits gemessen wird, sondern in der Wahrnehmung eines gesellschaftlichen Auftrags. Aus diesem Grund könnte die Übertragung einer Methode aus der Betriebswirtschaftslehre auf das Bildungswesen als unangemessen empfunden werden. Eine solche Sichtweise würde jedoch zwei Gegebenheiten übersehen. Erstens können Agenten im Non-profit-Bereich als unternehmerisch betrachtet werden, insofern sie danach streben, bestimmte Leistungsziele zu erreichen, zweitens lassen sich für die Wahrnehmung eines gesellschaftlichen Auftrags Indikatoren analog zum betriebswirtschaftlichen Handeln erarbeiten.¹⁵ In dieser Denkweise ist der Ertrag des Bildungswesens umso größer, je umfassender der gesellschaftliche Auftrag mit den eingesetzten Ressourcen erreicht wird. Weiterhin ist nicht nur eine Betrachtung unter Effektivitätsgesichtspunkten, sondern auch unter Effizienzgesichtspunkten möglich. Dennoch wäre es voreilig, ohne weiteres von einer Übertragbarkeit betriebswirtschaftlichen Benchmarkings auszugehen. Es gibt erleichternde und erschwerende Bedingungen. Erschwerend wirkt die bislang erst anfänglich erfolgte Indikatorisierung von Leistungsprozessen und Leistungszielen im Bildungsbereich. Dies kann dazu führen, dass man sich zum Teil auf Kennziffern bezieht, nur weil sie leicht zu erfassen sind, dabei aber ihren eingeschränkten Aussagewert über Prozesse und Zielerreichung übersieht. So führte z.B. der EU-Indikator „Anzahl der Computer pro Schüler“, der für das Ziel einer verstärkten informationstechnischen Bildung steht, im EU-Beitrittsland Zypern dazu, dass zwar mindestens ein Computer für jedes Klassenzimmer angeschafft wurde, aber die Nutzung dieser Geräte wegen unzureichender Lehrerweiterbildung nahezu vollständig unterblieb.¹⁶ Eine weitere Schwierigkeit im Vergleich zum betriebswirtschaftlichen Handeln besteht darin, dass die Leistungsziele des Bildungssystems mehrdimensional sind, es sollen z.B. nicht nur Qualifikationen für den Fortbestand der Gesellschaft bei einer hinreichend großen Anzahl von Heranwachsenden erzeugt werden, sondern es sollen auch Integrationsleistungen für alle Heranwachsenden im Hinblick auf das Zusammenleben in einem Staat erzielt werden.¹⁷ Darüber hinaus fehlt es insbesondere auf der Meso- und Mikroebene des Bildungssystems weithin an Ressourcen (Kompetenzen und Budgets) um Benchmarkingprozesse durchzuführen.¹⁸

Erleichternd für Benchmarkingprozesse im Bildungssektor kann hingegen die bislang relativ geringe Konkurrenz unter Bildungseinrichtungen wirken. So haben beispielsweise Ulrich G. Herrmann und Sascha Koch u.a. mit unterschiedlichen, voneinander unabhängigen Bildungseinrichtungen eine sogenannte „Benchmarkingkreise“ durchgeführt, worunter sie den geplanten Austausch von „guten Ideen und Lösungen“ durch die Leiter verschiedener Institutionen verstehen.¹⁹ Dieses Verständnis von Benchmarking als wechselseitigem voneinander Lernen lebt

¹⁵ Van Akeren, I./Hovestadt, G., Indikatorisierung der ‚Forum Bildung‘-Empfehlungen. Ein exemplarischer Versuch unter Berücksichtigung der bildungsbezogenen Indikatorenforschung und -entwicklung, Berlin (Bundesministerium für Bildung und Forschung) 2003. www.forum-bildung.de/files/Indikatorisierung.pdf.

¹⁶ Hadjithoma, C., ICT policy and the construction of ‚ICT in teaching and learning‘ in Cypriot primary schools with the scope of European integration. Paper presented at the YURE-Conference 22.–23.08.2005, Nicosia, Cyprus.

¹⁷ Vgl. klassische Theorien der Schule z.B. Tippelt, R., Bildung und sozialer Wandel. Eine Untersuchung von Modernisierungsprozessen am Beispiel der Bundesrepublik seit 1950, Weinheim 1990.

¹⁸ Aus der theoriegeleiteten Perspektive des Change Managements fehlt es an Wandlungsfähigkeit, wahrgenommenem Handlungsbedarf und zum Teil auch an Wandlungsbereitschaft, vgl. Krüger, W., Excellence in Change. Wege zur strategischen Erneuerung, Wiesbaden 2002; Böttcher, W./Brohm, M., Die Methodik des Change Managements und die aktuelle Schulreform. Über das gebrochene Verhältnis von Chancen und Realität, in: Die Deutsche Schule 96, 2004, S. 268–278.

¹⁹ Herrmann, U.G./Koch, S., Leitfaden zum EFQM-gestützten Qualitätsmanagement und Benchmarking in schulischen Bildungseinrichtungen. Institut für Pädagogik – Lehrstuhl für Berufs- und Wirtschaftspädagogik, Ruhr-Universität Bochum 2002.

davon, dass der öffentliche Raum einen Schutzraum bietet, indem der Fortbestand der Einrichtung nicht davon abhängt, bessere Ergebnisse als vergleichbare Einrichtungen zu erzielen.²⁰ Weiterhin sind erleichternde Bedingungen, der in hohem Maße analoge Organisationsbedarf und die große Zahl an möglichen Benchmarkingpartnern. In Zukunft könnte Benchmarking zusätzlich einen Schub von der sich erweiternden Autonomie von Meso- und Mikroebene im Bildungswesen erfahren. Diesen Weg zeichnet beispielhaft das Gutachten der Zukunftskommission des österreichischen Bundesministeriums für Bildung, Wissenschaft und Kunst. Die aus Wissenschaftlern gebildete Kommission hat aus den aktuellen Entwicklungen den Schluss gezogen, dass die Interpretation von Benchmarks eine neue Anforderung an praktisch alle Lehrer und Lehrerinnen und mittelbar auch an die Lehrerbildung ist.²¹ Dabei ist die Einsicht leitend, dass die notwendigen Maßnahmen zur Erfüllung von Benchmarks eine innere Entwicklung in jeder Schule erforderlich machen. Weil nur auf der Ebene jeder Schule letztlich eine Verbesserung der Gesamtleistung entstehen könne, müsse Schulentwicklung auch mit einer datengestützten (Selbst-)evaluation verbunden sein. Reformen der Lehrerbildung sollen auf Prozesse des Benchmarking im Schulbereich bezogen werden.

Auf der Makroebene politischer Steuerung stellt sich die Situation jedoch derzeit schon weiter fortgeschritten dar. Durch die kontinuierliche Erfassung von Vergleichsdaten im Rahmen der OECD²² ist eine Berichterstattung entstanden, die Kosten und Ergebnisse von Bildung auf internationaler Ebene in relativ differenzierter Weise sichtbar macht. Dieses wird ergänzt durch die PISA-Studien, die die erreichten Leistungen der Schüler/-innen am Ende der Pflichtschulzeit in den Fokus der Aufmerksamkeit rücken. Dadurch entsteht eine Wettbewerbssituation zwischen den beteiligten Nationen, und bei PISA auch innerhalb der Nationen zwischen ihren Teilstaaten, die sich aufgrund laufender Verbesserung und Aktualisierung der Datenlage durchaus mit dem Wettbewerb zwischen unterschiedlichen „Dienstleistern“ vergleichen lässt.

Auf der Ebene der Europäischen Union wird das Konkurrenzmotiv in Abgrenzung zu den nicht EU-Staaten explizit betont, um einen Raum von Kooperation zu schaffen, innerhalb dessen die einzelnen Staaten Benchmarkingprozesse eingehen. So formulierten die EU-Staats- und Regierungschefs im Rahmen des Lissabon Gipfels 2000 das Ziel „the most competitive and dynamic knowledge based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion“²³ zu werden. Dieser Zielsetzung folgend haben sich die Bildungsminister im Rahmen der umfassenden „Lissabon-Strategie“ auf gemeinsame Ziele für Verbesserungen bis zum Jahre 2010 verständigt. Diese wurden von der Europäischen Kommission als Benchmarks ausformuliert.²⁴ So lautet beispielsweise ein Benchmark: „Bis 2010 ist der Prozentsatz der 15-Jährigen, die im Bereich von Lesekompetenz, mathematischer Grundbildung und naturwissenschaftlicher Grundbildung schlechte Leistungen erzielen, in jedem Mitgliedstaat im Vergleich zum Jahr 2000 mindestens zu halbieren.“ Als schlechte Leistung gilt dabei Level 1 oder darunter in PISA. Diese Benchmarks wurden vielfach kritisiert;

²⁰ Es mag eingewandt werden, dass sich bisweilen mehrere Schulen an einem Standort in der Konkurrenz um zurückgehende Schülerzahlen befinden. Die Praxis zeigt jedoch, dass in einer solchen Situation nicht indiktorisierbare Leistungen von Schulen als Entscheidungsgrundlage für Schulschließungen dienen, sondern die Verfügbarkeit von Gebäuden, die relative örtliche Lage, die verbleibende Amtsdauer der Schulleiter, der Wert des Schulgrundstücks oder ähnliche Kriterien, die durch besondere Leistungen der Einzelschule nicht zu beeinflussen sind.

²¹ Haider, G./Eder, F./Specht, W./Spiel, C., Reformkonzept der österreichischen Zukunftskommission. Zukunft: Schule – Strategien und Maßnahmen zur Qualitätsentwicklung, Wien 2003.

²² OECD (Hrsg.), Education at a glance. OECD indicators 2003.

²³ Conclusions of the Lisbon European Council, 2000. Paragraph 5. (http://www.bologna-berlin2003.de/pdf/presidency_conclusions_Lissabon.pdf).

²⁴ Europäische Kommission, Mitteilung der Kommission vom 20. November 2002 – Europäische Benchmarks für die allgemeine und berufliche Bildung, Follow-up der Tagung des Europäischen Rates von Lissabon 2002 [KOM(2002). 629 endg.] <http://europa.eu.int/scadplus/leg/de/cha/c11064.htm>.

Bildungsforscher verwiesen darauf, dass die benannten Ziele im angegebenen Zeitraum praktisch uneinlösbar seien.²⁵ So verwundert es nicht, dass sich in einem Ratsbeschluss der EU vom 5. und 6. Mai 2003 eine inhaltliche Korrektur der Benchmarks findet; hier heißt es nun: „Bis 2010 soll der Prozentsatz der 15-Jährigen, die im Bereich Lesekompetenz schlechte Leistungen erzielen, um mindestens 20 % im Vergleich zum Jahr 2000 abgesenkt werden.“²⁶ Man hat sich also politisch dafür entschieden, nicht den Zeitraum zur Zielerreichung auszudehnen, sondern das Ziel bescheidener zu formulieren. Die Steuerung über Benchmarks, die von allen anerkannt werden, bleibt dabei zentraler Bestandteil einer Politik, die auch als „Open Method of Coordination“ (OMC) bezeichnet wird.

Ziel dieses Politikansatzes ist es, einerseits das Prinzip der Subsidiarität, nach dem kleinere Einheiten ihre Angelegenheiten möglichst selbständig regeln sollten, und andererseits das Bedürfnis nach Konvergenz im Rahmen des gemeinsamen Marktes zu versöhnen. Um dieses Ziel zu erreichen, sollen mittels OMC die verschiedenen Ebenen der Entscheidung miteinander verbunden werden. OMC versucht dabei einen Mittelweg zu gehen zwischen der Integration verschiedener Politiken in eine gemeinsame Politik und der bloßen Koordination unterschiedlicher Politiken.²⁷ Dies soll erreicht werden, indem einerseits gemeinsame Ziele durch die Formulierung von Benchmarks im Sinne von verbindlichen Kennziffern hinsichtlich bestimmter Output-Indikatoren konkretisiert werden. Andererseits bleiben die weiteren Schritte des Benchmarkings in der Verantwortung der Beteiligten. Die EU-Kommission unterstützt den Prozess durch die fortschreitende Indikatorisierung des Bildungswesens, durch Berichte zum Sachstand und die Suche nach so genannten Best Practice Beispielen, aber die Verantwortung für Entscheidungen in der Bildungspolitik verbleibt auf der Ebene der Mitgliedsländer.

Bei dieser Methode wird das Verständnis von Benchmarking insofern ausgeweitet, als die Benchmarks hier nicht von einem real existierenden Partner genommen, sondern fiktiv gesetzt werden, und ihre Geltung beweisen müssen, ohne dass auf einen Kontext verwiesen wird, indem entsprechende Wandlungsprozesse schon einmal gelungen sind. Hinzu kommt, dass die Formulierung von Benchmarks noch keine Benchmarkingprozesse in den Mitgliedsländern garantiert.

Gleichwohl können in diese Politik auch Elemente des ursprünglichen Konzepts von Benchmarking – im Sinne eines Vergleichs mit besser operierenden Systemen – integriert werden. So widmete sich beispielsweise eine EU-Arbeitsgruppe aus Experten der Lehrerbildung der Identifikation und Darstellung von vorbildlichen Ansätzen in der Lehrerbildung.²⁸ Auch wird die Formulierung und Nutzung von Benchmarks auf EU-Ebene selbst kritisch reflektiert.²⁹ Die „Standing Group on Indicators and Benchmarks (SGIB)“ des Directorate-General for Education and Culture der europäischen Kommission hat schon 2003 eine Liste mit Indikatoren veröffentlicht. Im Bereich der Lehrerbildung ist ein Benchmarking beispielsweise hinsichtlich der Stunden in Weiterbildung im Verhältnis zur Arbeitszeit von Lehrkräften und eines zur Teil-

²⁵ Klieme, E./Avenarius, H./Blum, W. et al., Zur Entwicklung nationaler Bildungsstandards – eine Expertise. Bundesministerium für Bildung und Forschung (BMBF), Berlin 2003.

²⁶ Council of the European Union, Council Conclusions on Reference Levels of European Average Performance in Education and Training (Benchmarks). Brussels 2003 (Council of the European Union 8486/03 EDUC 83) <http://europa.eu.int/comm/education/doc/official/keydoc/2003/benchmark.pdf>.

²⁷ De la Porte, C./Poche, P., Social benchmarking, policy making and new governance in the EU, Journal of European Social Policy 11, 2001, S. 291–307.

²⁸ European Commission Directorate-General for Education and Culture, Implementation of „Education & Training 2010“ work programme. Working Group „Improving education of teachers and trainers“, Progress Report, Brussels 2003. http://europa.eu.int/comm/education/policies/2010/doc/opening-up-learning-and-citizenship_en.pdf.

²⁹ Commission of the European Communities 2004.

nahme an den Mobilitätsprogrammen (z.B. Sokrates) vorgesehen. Zunächst entstanden jedoch nur Indikatoren, die sich auf den Lehrerbedarf als demographisches Problem und als Rekrutierungsproblem beziehen. Auch innerhalb der EU-Kommission steht außer Zweifel, dass damit die Komplexität des Gegenstands noch nicht ausreichend abgebildet wird. Die oben erwähnte EU-Arbeitsgruppe schlug deshalb 2004 vor, zunächst weitere Indikatoren zu den Themen Zufriedenheit mit der Lehrererstausbildung und Investitionen in die Lehrerbildung zu bilden.³⁰ Insgesamt wird es für die Qualität der Benchmarkingprozesse in Zukunft entscheidend sein, inwieweit es erstens gelingt, die Arbeit der Experten für Lehrerbildung und die Arbeit der Experten für Indikatorisierung zusammenzuführen um angemessene Benchmarks zu formulieren.³¹ Zweitens wird als ein kritischer Erfolgsfaktor erkannt, ob diese Indikatoren auch zuverlässig und regelmäßig erfasst werden können.³² Gelingt es nicht in diesen beiden Punkten voranzukommen, droht Benchmarking zu einer Art symbolischer Politik beizutragen, bei der nicht mehr an übergreifenden Aufgabenstellungen gearbeitet wird, sondern nur noch an der Darstellung von Zahlen, deren Bedeutung für das Bezugssystem zweifelhaft ist (vgl. das Beispiel der ICT-Einführung in Zypern).

In Entsprechung zur Politik auf europäischer Ebene lassen sich auch in Deutschland Reformmaßnahmen beobachten, die in einem engen Verhältnis zu Benchmarking stehen. Dabei gehen die Bundesländer in der Umsetzung sowohl inhaltlich als auch zeitlich teils gemeinsame, teils unterschiedliche Wege. Zu diesen Veränderungen im Bildungsbereich gehören u.a.:

- die Einführung von fachspezifischen Bildungsstandards
- die Konzeption von zentralen Abschlussprüfungen für alle Bildungsabschlüsse
- eine kontinuierliche und verbesserte nationale Bildungsberichterstattung
- die kontinuierliche Beteiligung Deutschlands an internationalen und nationalen Schülerleistungsmessungen
- eine stärkere Selbstbestimmung der einzelnen Schule über ihre konkreten Ziele und Maßnahmen, verbunden mit einer Evaluation ihrer Arbeit

Aber nicht nur Benchmarking im Sinne einer Quantifizierung und eines Vergleichs von Bildungsergebnissen wird angebahnt, sondern es findet auch ein Benchmarking im ursprünglichen Sinn einer Orientierung an den erfolgreichsten Bildungssystemen statt. So förderte das deutsche Bundesbildungsministerium (BMBF) eine Vergleichsstudie, in deren Rahmen die Bildungssysteme von in PISA erfolgreichen westlichen Industriestaaten auf mögliche Verbesserungsansätze für das deutsche Bildungssystem untersucht wurden. Das Besondere bei dieser Studie im Vergleich zu den Arbeiten auf EU-Ebene ist neben einer anderen Auswahl an Vergleichsländern, dass in ihr weniger nach aktuellen Reformen und vorbildlichen Detaillösungen einzelner Länder gesucht wurde, und stattdessen die Einordnung der Teilfunktionen in die jeweiligen Bildungs-

³⁰ Vgl. *European Commission Directorate-General for Education and Culture*, Implementation of „Education and Training 2010“ Work Programm. Working Group A „Improving the education of teachers and trainers“, Progress Report September 2004, S. 16.

³¹ *European Commission Directorate-General for Education and Culture*, Implementation of „Education & Training 2010“ work programme. Working Group „Improving education of teachers and trainers“, Progress Report, Brussels 2003, S. 5–6, Annex IV. http://europa.eu.int/comm/education/policies/2010/doc/opening-up-learning-and-citizenship_en.pdf.

³² Vgl. *European Commission Directorate-General for Education and Culture*, Implementation of „Education and Training 2010“ Work Programm. Working Group A „Improving the education of teachers and trainers“, Progress Report September 2004, S. 55; Annex 3.

systeme und die mögliche Bedeutung für Reformen in Deutschland stärker berücksichtigt werden konnte.³³ So sehr diese Studie allgemein begrüßt wurde³⁴, so liegen ihre Grenzen jedoch darin, dass Sie nur auf Daten zurückgreifen kann, die auch schon an anderer Stelle zur Verfügung standen; ihr Verdienst besteht gleichwohl darin, dass sie den Daten breitere, auf das Herkunftssystem bezogene und interpretatorische Kontexte zurückgibt, als sie in der öffentlichen Diskussion regelmäßig zur Verfügung stehen.

3 Bewertung der Methode Benchmarking im Bildungsbereich

Benchmarking entstand in einem Kontext, in dem die Macht zu steuern, die Notwendigkeit zur Veränderung und die Motivation, diese Veränderung auch durchzuführen, als klar bestimmbar und gegeben betrachtet werden durften. In der Übertragung auf die Handlungsebenen des Bildungsbereichs werden diese Punkte fraglich. Auf der Mikro- und Mesebene der Bildungseinrichtungen und Bildungsverwaltung lassen sich die Mitarbeiter grundsätzlich nicht mit den Angestellten eines Wirtschaftsunternehmens vergleichen, weil sie sich als Professionelle³⁵ betrachten, die aufgrund ihres Status ein höheres Maß an Autonomie mitbringen. Ihre beamtenrechtliche Stellung reduziert die persönliche Notwendigkeit zu Veränderungen, und die Steuerungswirkung von Maßnahmen übergeordneter Instanzen muss vergleichsweise geringer als in Wirtschaftsunternehmen eingeschätzt werden.

Bei Auswahl und Formulierung von Benchmarks sind mehrere Kriterien zu beachten. Es ist danach zu fragen, inwieweit ausgewählte Benchmarks die Zieldimensionen einer Institution abdecken, bzw. welche Zieldimensionen (noch) nicht durch Benchmarks erfasst werden. Weiterhin ist danach zu fragen, inwieweit ein Benchmark geeignet ist, einen Fortschritt im Hinblick auf die Erreichung von Zielen abzubilden. So verweist die Kennziffer „durchschnittliche Weiterbildungsstunden pro 1000 Arbeitsstunden“ im Detail auf ein anders Ziel als die Kennziffer „Lehrer ohne Weiterbildung pro Jahr“. Die Formulierung von Benchmarks bedeutet nun, die Daten, die zu den Kennziffern erhoben wurden, zu vergleichen und Zielgrößen für den eigenen Verantwortungsbereich zu bestimmen. Dabei kann sich der Vergleich auf dieselbe Institution zu einem anderen Zeitpunkt beziehen, indem z.B. als Benchmark eine Steigerung um 10 % innerhalb einer bestimmten Frist formuliert wird. Der Vergleich kann sich zweitens auf andere Institutionen beziehen, wobei ein Benchmark z.B. der Mittelwert des oberen Drittels der Vergleichsinstitutionen sein kann. Drittens kann sich der Vergleich auf ein objektiv gesetztes Kriterium beziehen, z.B. die Verpflichtung für alle Lehrkräfte, mindestens eine Weiterbildung pro Jahr zu besuchen. Als Steuerungspolitik könnte dann festgelegt werden, die Abweichung von diesem Kriterium (Vollzugsdefizit) unter einer Toleranzschwelle von 5 % zu halten.

Wenn Benchmarks ein Versuch sind, die Verbesserungsziele einer Institution an quantifizierbare Kennziffern zu binden, ist damit noch nicht gesagt, wie groß der Unterschiedswert zum Ausgangszustand gewählt werden sollte. Einerseits geht es darum, nicht durch überfordernde Benchmarks das Instrument zu entwerten und Frustration in den Institutionen zu erzeugen. Deshalb müssen Rahmenbedingungen berücksichtigt werden, die für die Leistungsfähigkeit einer In-

³³ Döbert, H./Klieme, E./Sroka, W. (Hrsg.), *Conditions of School Performance in Seven Countries. A Quest for Understanding the International Variation of PISA Results*. Münster 2004; Döbert, H./van Ackeren, I./Bos, W. et al., *Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten*, Berlin 2003 (Bundesministerium für Bildung und Forschung). http://www.dipf.de/publikationen/pisa_vergleichsstudie.pdf.

³⁴ Die OECD betrachtete die Studie als beispielhafte Fortsetzung von PISA und fertigte eine englische Übersetzung für ein internationales Publikum. <http://www.pisa.oecd.org/dataoecd/30/24/33858946.pdf>.

³⁵ Combe, A./Helsper, W. (Hrsg.), *Pädagogische Professionalität. Untersuchungen zum Typus pädagogischen Handelns*, 4. Aufl. Frankfurt 2002.

stitution verantwortlich sind. Andererseits sollen Benchmarks das Management bzw. Politiker zu einer entsprechenden Allokation von Ressourcen herausfordern. Benchmarks können also nicht sinnvoll festgesetzt werden, ohne den Zusammenhang der Ressourcen und weiteren Qualitätsentwicklungsmaßnahmen in einer Institution zu berücksichtigen. Benchmarks sind Zielpräzisierungen, sie geben nicht den Weg vor und liefern auch keine Ausrüstung für diejenigen, die den Weg beschreiten sollen. Deshalb ist es sinnvoll, von real existierenden Partnern zu lernen; ihr Beispiel zeigt, was unter konkreten Bedingungen möglich ist. Bei der ursprünglichen Konzeption von Benchmarking, in der die Prozesse und Strukturen eines vorbildlichen Unternehmens oder Bildungssystems für die eigene Entwicklung genutzt werden, sind die sich zeigenden „Successful Practice“ stärker präskriptiv. Dieses ursprüngliche Verständnis ist insofern anspruchsvoller, als es von den Verantwortlichen in einem Unternehmen, einer Schule oder einem Bildungssystem verlangt, dass sie über die notwendige Umgestaltungsmacht verfügen. Wahrscheinlich verblasst die ursprüngliche Bedeutung von Benchmarking im politischen Rahmen auch deshalb, weil die Akteure zum Teil nicht über die Zuständigkeit oder politischen Durchsetzungsmöglichkeiten verfügen, entsprechende Wandlungsprozesse einzuleiten. Ein weiteres Problem besteht darin, die beiden unterschiedlichen Benchmarkingprozesse (Indikatorisierung und vergleichsorientierte Verbesserung) aufeinander zu beziehen. So ist „Successful Practice“ oft nur unzureichend indikatorisiert, und vorhandene Indikatoren bilden die Komplexität der Bildungsprozesse nur unzureichend ab.

4 Benchmarking als Instrument der Steuerung

Der Wandel in den politischen Steuerungsoptionen des Bildungswesens wird derzeit oft als Wandel von der Inputsteuerung zur Outputsteuerung begriffen. Gemeint ist damit, dass weniger durch zentrale Vorgaben und zentrale Maßnahmen der Budgetverwaltung und des Personalwesens gesteuert werden soll, sondern mehr durch Kontrolle der Ergebnisse von Bildungsinstitutionen und Bildungssystemen. Wenn der Begriff Benchmarking, in diesem Kontext verwendet wird, so bezieht er sich auf die Feststellung von Kennzahlen, die als relevante Indikatoren für die Ergebnisse des Bildungssystems gelten können. Die Leistungsprozesse, die zu diesen Ergebnissen hinführen, werden weniger in den Blick genommen. Hier herrscht gleichsam die „Open Method of Coordination“. Es wird dem einzelnen Staat, dem Land, der Schule, dem Fachkollegium bzw. der einzelnen Lehrkraft überlassen, wie sie die auf oberer Ebene festgesetzten Leistungsziele erreicht. Je nach dem wird man diesen Freiraum als wohlverstandene Autonomie oder als Alleingelassensein mit den tatsächlichen Problemen vor Ort interpretieren. Ein vollständiges Benchmarking, in dessen Verlauf auch die Leistungsprozesse in den Blick kommen, wird nicht konsequent in Gang gesetzt, stattdessen lässt sich die Politik als Steuerung über das Setzen von Zielen (management by objectives) beschreiben. Schon das Beispiel Xerox zeigte jedoch, dass es im Management nicht ausreichend ist, auf das Produkt zu schauen.

Indikatorisierte Information zeigt auf dem Stand der gegenwärtigen Theoriebildung nur, dass etwas zu tun ist, nicht aber, was genau zu tun ist. Wenn Indikatoren pädagogisches und politisches Entscheidungshandeln ersetzen sollen, indem aus ihnen unmittelbar Entscheidungen abgeleitet werden, so setzt dies empirisch geprüfte Modelle mit zuverlässigen verallgemeinerbaren Wirkungsketten voraus, zu denen alle Indikatoren erfasst werden. Liegen solche modellbezogenen Informationen vor, so müssen in ihrer Interpretation immer die Grenzen des Modells mitbedacht werden. Solche Grenzen können dort liegen, wo in einem Modell nur ein Teil der Bildungsziele abgebildet wird, wo Interaktionseffekte zwischen Bedingungen nicht in den Blick

genommen werden, oder wo den Handelnden Zweckrationalität im Hinblick auf Systemziele statt kulturell gebundene Wertrationalität unterstellt wird.³⁶

Solange solch umfassende Modelle nicht vorliegen – und es darf bezweifelt werden, ob sie jemals vorliegen werden – kann Benchmarking das Entscheidungshandeln auf den unterschiedlichen Ebenen des Systems nicht ersetzen, sondern lediglich Hinweise geben, wo unser Alltagsverständnis des Gegebenen zu eng ist, wo mehr Aufmerksamkeit notwendig ist und welche Ideen anderenorts als hilfreich erachtet werden. Wird Benchmarking zum Automatismus, der die beschriebenen Einschränkungen missachtet, mutiert es zu einem blinden „educational borrowing“³⁷, das zwar von politischer Aktivität zeugt, aber nicht mehr den Anspruch auf rationales, evidenzbasiertes Handeln erheben kann.

Zu beachten ist, dass mit der Formulierung und Erfassung von Benchmarks, oder auch mit der Identifikation einer vorbildlichen Praxis allein, noch keine Verbesserung der Qualität im Bildungswesen erreicht ist. Auf dieser Grundlage müssen erst noch weitere Maßnahmen der Qualitätsentwicklung aufbauen. Es muss ein Einvernehmen über die Prozessgestaltung der weiteren Entwicklung hergestellt werden, die Verantwortlichkeiten der unterschiedlichen Ebenen in diesem Prozess müssen geklärt werden, erst dann können Veränderungen in der Praxis im Sinne eines kontinuierlichen Benchmarkings zielführend reflektiert werden.

Verf.: Dr. Hermann Josef Abs, Dr. Peter Döbrich, Alexander Wicker, Deutsches Institut für Internationale Pädagogische Forschung, Schloss Str. 29, 60486 Frankfurt am Main

³⁶ Vgl. Bryk, A.S./Hermanson, K.L., Educational indicator systems: observations on their structure, interpretation and use, in: Review of Research in Education 19, 1993, S.451–484.

³⁷ Steiner-Khamsi, G., Re-framing educational borrowing as policy strategy, in: Caruso, M./Tenorth, H.-E. (Hrsg.), Internationalisierung – Internationalisation. Semantik und Bildungssystem in vergleichender Perspektive – Comparing Educational Systems and Semantics, Frankfurt 2002, S. 57–89.