

K o n z e p t e d e r Q u a l i t ä t
i n d e r E r w a c h s e n e n b i l d u n g



A u f s ä t z e u n d P r o t o k o l l i m R a h m e n d e r
W e r k s t a t t a m B u n d e s i n s t i t u t f ü r
E r w a c h s e n e n b i l d u n g S t . W o l f g a n g
v o m 2 . b i s 3 . O k t o b e r 2 0 0 0

Materialien zur Erwachsenenbildung Nr. 2/2001

***Konzepte der Qualität in der Erwachsenenbildung
Bericht zur Tagung am BIfEB vom 2. bis 3. Oktober 2000***

© 2001

Bundesministerium für Bildung, Wissenschaft und Kultur
Abteilung Erwachsenenbildung V/E/18

1014 Wien, Wipplingerstraße 20
MR Mag. Ernst Koller

und

Bundesinstitut für Erwachsenenbildung St. Wolfgang

5350 Strobl, Bürglstein 1-7
OR Dr. Christian Kloyber

Dokumentation, Lektorat und Layout: Dr. Christian Kloyber
Umschlaggestaltung: Robert Radelmacher

Vorwort

Das Bewusstsein über die Bedeutung und Notwendigkeit von Erwachsenenbildung und Weiterbildung ist vor dem Hintergrund der immer rascher vor sich gehenden technologischen, wirtschaftlichen und gesellschaftlichen Veränderungen in letzter Zeit deutlich gestiegen. In diesem Zusammenhang hat das Bildungsministerium gemeinsam mit der Konferenz der Erwachsenenbildung Österreichs ein EB-Aktionsprogramm entworfen, in dem u. a. auch die Entwicklung von theoretischen Grundlagen und die Publikation von Ergebnissen der Beschäftigung mit solchen Grundlagen enthalten sind.

Leider existiert seit der Einstellung der Zeitschrift Erwachsenenbildung in Österreich kein gesamtösterreichisches Publikationsorgan zur Erwachsenenbildung mehr – wenn auch verschiedene Zeitschriften der einzelnen Verbände sowie der Förderungsstellen für Erwachsenenbildung bzw. von Landes-Arbeitsgemeinschaften neben verbandsinternen bzw. regionalen Informationen sehr wohl auch wichtige theoretische Auseinandersetzungen beinhalten - und auch sonstige Publikationen fehlten in letzter Zeit weitgehend.

Aus diesem Grund soll nunmehr eine neu begonnene Reihe von Materialien zur Erwachsenenbildung herausgegeben werden. Sie wird in unregelmäßiger Folge erscheinen und für die Erwachsenenbildung interessante Materialien, wie Tagungsberichte, Studien, Analysen, methodisch-didaktische Arbeiten und sonstige für alle in der EB Tätigen brauchbare Texte und „graue Literatur“ öffentlich zugänglich machen.

Die Materialien werden zwar sicherlich nicht eine regelmäßig erscheinende Zeitschrift ersetzen können und sind auch keineswegs als Konkurrenz zu den verschiedenen Zeitschriften und sonstigen Veröffentlichungen der Verbände oder der Förderungsstellen gedacht. Sie sollen aber mit dazu beitragen, die Auseinandersetzung mit (theoretischen) Grundlagen für die Bildungsarbeit mit Erwachsenen und vor allem den Diskurs über eine zeitgemäße Erwachsenenbildung voranzutreiben.

Mit dieser Ausgabe wird eine zweite Nummer zum Thema Qualitätsentwicklung vorgelegt, die eine Art Fortsetzung zur Nummer 1/2000 der „Materialien“ darstellt. Im Gegensatz zur Nummer 1/2000, die eher praktische Aspekte der Qualitätsentwicklung der einzelnen EB-Verbände und von Landesarbeitsgemeinschaften vorstellte, standen diesmal verschiedene Ansätze und Instrumente der Qualitätsentwicklung aus eher grundsätzlicher und theoretischer Sicht im Vordergrund. Damit soll ein weiterer Beitrag zur Auseinandersetzung mit dem Thema Qualitätsentwicklung geleistet werden; weitere Ausgaben zum Thema werden folgen.

ERNST KOLLER (BMBWK, Abt. V/E/18 - Erwachsenenbildung)

Inhaltsverzeichnis

	Seite
<i>Vorwort</i>	3
<i>Inhaltsverzeichnis</i>	4
<i>Einleitung</i>	5
<i>Evaluation und Qualitätsmanagement. Systemische Verfahren zur Entwicklung von Qualität im Bildungswesen (WOLFGANG BEYWL)</i>	7
<i>Programmevaluation in pädagogischen Praxisfeldern: begriffliche und konzeptionelle Grundlagen (WOLFGANG BEYWL)</i>	18
<i>Konzepte der Qualität in der Erwachsenenbildung Protokoll und kritische Betrachtung (ELISABETH WAPPELSHAMMER)</i>	35

Einleitung

Konzepte der Qualität in der Erwachsenenbildung

Dokumentation der Informationsveranstaltung am Bundesinstitut für Erwachsenenbildung St. Wolfgang vom 2. bis 3. Oktober 2000



Seit nunmehr 6 Jahren veranstaltet das Bundesinstitut für Erwachsenenbildung (BIfEB) eine modulare Reihe von Seminaren und Workshops zum Thema der Qualitätsentwicklung und der Evaluation in der Erwachsenenbildung. Über 150 Mitarbeiter und Mitarbeiterinnen aus Institutionen der österreichischen Erwachsenenbildung haben an diesen Veranstaltungen teilgenommen, aber auch Kollegen und Kolleginnen aus benachbarten Staaten, wie Ungarn, Kroatien, Slowenien, Liechtenstein, Deutschland und Italien (Südtirol).

In diesen Jahren hat sich die Zusammenarbeit mit Prof. Klaus Künzel und Wolfgang Beywl von der Universität zu Köln bewährt und erfolgreich entwickelt. Das

Angebot dieser Veranstaltungen reicht von Einführungen in Bereiche der Qualitätsentwicklung bis zur praktischen Erprobung einzelner Methoden und Instrumenten der Evaluation. Damit wird ein Prinzip verfolgt, theoretische und praxisorientierte Einführung und eine exemplarische Darstellung von Methoden, Techniken und Instrumente miteinander zu verbinden. Bedarf und Interesse der Institutionen der Erwachsenenbildung und konkrete Vorstellungen über die Weiterentwicklung des Programms dieser Reihe geben dem Veranstalter und den Experten die Entscheidungsmöglichkeit das Angebot jährlich neu zu gestalten. Ein wichtiges Instrument dieser Bedarfserhebung ist die Form dieser Informationsveranstaltung.

Am 2. und 3. Oktober 2000 fand daher eine Informationsveranstaltung zum Thema „Konzepte der Qualität in der Erwachsenenbildung“ statt, die von Wolfgang Beywl (Geschäftsführer der interdisziplinären Arbeitsstelle für Evaluation Pädagogischer Dienstleistungen an der Erziehungswissenschaftlichen Fakultät der Universität zu Köln) geleitet wurde.

Zur Steuerung und Entwicklung von Qualität in der Erwachsenenbildung sind in den vergangenen Jahren verschiedene Ansätze und Instrumente angewandt worden: Zertifizierung nach DIN EN ISO 9000, Total Quality Management, House of Quality, European Foundation Quality Award, Balanced Score-Card, Bildungscontrolling, Monitoring, Gütesiegelgemeinschaften, externe und interne Evaluation (summativ oder formativ ausgerichtet), Selbstevaluation ... In der Veranstaltung

werden diese Konzepte und ihre jeweilige Reichweite diskutiert, die Erfahrungen der Teilnehmenden systematisiert. Grundannahme ist, dass kleine oder große Anbieter in der Eb, Anbieter standardisierter Lehrgänge oder Anbieter von Unikaten, Anbieter politischer / kultureller Bildung oder Anbieter von Schul/ Berufsabschlüssen jeweils einen unterschiedlichen Mix von Qualitätsansätzen benötigen, der konkreter zu bestimmen ist. Als Ziel wurde vereinbart, dass die Teilnehmenden sich einen breiten Überblick über verschiedene Ansätze der Bestimmung und Steuerung von Qualität in der Erwachsenenbildung schaffen können. Sie lernen disziplinäre Hintergründe kennen und sind in der Lage, Vorteile und Nachteile eines Ansatzes für verschiedene thematische, konzeptionelle und organisatorische Konstellationen in der Erwachsenenbildung zu benennen. Sie sind in ihrer Aufgabe fachlich gestärkt, strategische Entscheidungen zu Qualität und ihrer systematischen Bearbeitung zu treffen.

Das Programm

dieser Veranstaltung umfasste folgende Themen:

- ✪ Die fünf Meta-Dimensionen der Qualität und typische Ansätze zu ihrer Steuerung
- ✪ Alternative Gegenstände: Programm – Organisation – Personal
Perspektiven: Kontext – Konzept – Struktur– Input – Prozess – Ergebnis (outputs und outcomes)
- ✪ Evaluation und Qualitätsmanagement – ein systematischer Vergleich komplementärer Ansätze
- ✪ Referat mit Diskussion: Welcher Qualitätsansatz für welche Zwecke und Konstellation in der EB? Identifikation
(a) typischer Zwecksetzungen von „Qualitätsbemühungen“
(b) typischer Konstellationen – Prüfung von Stärken und Schwächen der Qualitätsansätze
- ✪ Erarbeitung in Arbeitsgruppen – Schlussfolgerungen im Plenum: Orientierungslinien für eine differenzierte Qualitätspolitik in der EB

Sie war auch als Fortsetzung der Tagung „Qualitätsentwicklung in der österreichischen Erwachsenenbildung“ gedacht, die im Jänner 2000 stattfand und in der ersten Nummer 1/2000 der „Materialien zur Erwachsenenbildung“ dokumentiert ist.

CHRISTIAN KLOYBER (*Bundesinstitut für Erwachsenenbildung St. Wolfgang*)

Evaluation und Qualitätsmanagement

Systematische Verfahren zur Entwicklung von Qualität im Bildungswesen

Wolfgang Beywl, UNIVation e.V.

Die Bezeichnung „Qualitätsentwicklung“ breitet sich gegenwärtig auch im Bildungsbereich aus. Sie umfasst zunehmend die Bezeichnungen „Qualitätssicherung“ oder „Qualitätsmanagement“ und auch „Evaluation“ / „Selbstevaluation“. Dieser Beitrag klärt Gemeinsamkeiten, Unterschiede und gegenseitige Ergänzung der beiden Verfahrensweisen.

Zu Beginn der neueren „Qualitätsdebatte“ im Bereich von Bildung und Erziehung wurde oft die Bezeichnung „Qualitätssicherung“ genutzt. Es kann vermutet werden, dass dieses Wort sich auch deshalb schnell verbreitete, da es im wörtlichen Sinne Sicherheit verspricht: Die bereits erreichte Qualität wird durch geeignete Maßnahmen auf Dauer gestellt. Verlässliches und als bewährt anerkanntes wird in einer „Qualitätssicherungsinitiative“ noch einmal festgezurr, damit es auf dem Bildungsschiff, das in schwere See geraten ist, nicht ins Rutschen gerät und über Bord geht.

Ein prominentes Beispiel hierfür sind die „Materialien zur Qualitätssicherung in der Kinder- und Jugendhilfe“, von denen das Deutsche Bundesministerium für Familie, Senioren, Frauen und Jugend seit 1996 ca. 30 Ausgaben veröffentlicht hat – hier war „Qualitätssicherung“ noch der Oberbegriff für verschiedenste Verfahren der Steuerung und Überprüfung von Qualität. Andere deutschsprachige Beispiele sind das Konzept der Landesregierung in Bozen für die dortige öffentlich geförderte Erwachsenenbildung oder ein bereits Mitte der 90er Jahre nach DIN ISO 9000 zertifiziertes Erwachsenenbildungsinstitut in Österreich, in der Steiermark (Erwachsenenbildungsinstitution der Landwirtschaftskammer Steiermark).

„Qualitätssicherung“ ist zunächst eine im guten Sinne konservative Strategie. Etwas Vorhandenes – nämlich (gute) Qualität – soll bewahrt werden. Dies stärkt das Selbstbewusstsein der bildungspolitischen Akteure, dass sie schon immer etwas Gutes geleistet haben.

„Qualitätsmanagement“ wurde zeitlich leicht versetzt durch die Bildungsszene entdeckt und teilweise offener und partizipativer wahrgenommen als etwa die scheinbar einengenden Vorgaben der DIN EN ISO 9000ff und daher leichter akzeptiert, manchesmal in Euphorie umgesetzt. Auch schien hier eine Nähe zum „Bildungsmanagement“ gegeben, einer spätestens seit Mitte der 90er Jahre etablierten Teildisziplin im Schnittfeld von Erziehungswissenschaft (pädagogischer Prozess) und Betriebswirtschaftslehre (ökonomische Funktion). Das vielen Sympathische hier ist der Wortteil Management, der – endlich Lösung, Steuerungsfähigkeit, zielgeführtes Handeln – in einer vielleicht lange als schwer veränderbaren institutionell geprägten Bildungslandschaft versprach.

Wegen des „guten Klangs“ der beiden Wörter wurden diese einerseits samt der damit verbundenen Konzepte z.T. wohlwollend in das Vokabular der pädagogischen Literatur übernommen. Marianne Meinhold empfiehlt in Ihrem Buch „Qualitätssicherung und Qualitätsmanagement in der Sozialen Arbeit“ (1996, S. 21): „Mit einer gewissen Sicherheit kann man davon ausgehen, dass in der Sozialen Arbeit ein Hilfe- oder Pflegebedürftiger als „Kunde“ respektvoller umworben wird als der Klient im alten Sinne. Aus diesen Gründen ist die Forderung nach einer größeren Kundenorientierung auch in der Sozialen Arbeit sinnvoll“.

Zu Verwirrung und bisweilen offener Ablehnung führte andererseits immer wieder, dass sowohl „Qualitätssicherung“ als auch „Qualitätsmanagement“ als Begriffe und Konzepte durch die technisch bzw. ökonomisch geprägte Qualitätswissenschaften definiert und mit den dort verankerten – der Humandienstleistungen oft fremden – Instrumentarien verbunden sind. Vehemente Kritiker sind z.B. Joachim Merchel (1998) oder Burkhard Müller (1996), die die fremddisziplinären Qualitätskonzepte ebenso global rezipieren wie viele ihrer Befürworter, sie jedoch als für Soziale Arbeit nicht oder wenig leistungsfähig ablehnen.

In neuerer Zeit wird zunehmend der allgemeinere und nicht vorgeprägte Begriff der „Qualitätsentwicklung“ genutzt. Diese Benennung soll kennzeichnen, dass soziale Dienste und Einrichtungen speziell angepasste, zusätzliche oder auch ganz andere Qualitätsverfahren benötigen. Dabei schließt Qualitätsentwicklung das Ziel von systematischen Qualitätsbemühungen (i.S. von Förderung und Entfaltung von Qualität, „mehr“ Qualität) ein, als auch verschiedenste eingesetzte Verfahrensweisen zur Steuerung, Regelung und Überprüfung von Qualität.

Ein systematischer Überblick über alle diese Verfahren müsste u.a. die DIN EN ISO 9000ff und ihre Adaptionen, die Qualitätspreise (i.B. European Foundation for Quality Management - EFQM) und seine Adaptionen (inklusive der Assessment-Verfahren), die unterschiedlichen Spielarten des (Total) Quality Management, ggfs. auch Konzepte wie *Balanced Scorecard*, Ansätze des Fach – oder pädagogischen Controlling, die Neue Steuerung / das „Kommunale Management“ und schließlich die Ansätze von Evaluation und Selbstevaluation einbeziehen.

Beispielhaft soll ein solcher Vergleich für die beiden prominenten Verfahrens-Großfamilien des Qualitätsmanagements einerseits und der Evaluation andererseits thesenartig vorgenommen werden. Im Vordergrund stehen die Gemeinsamkeiten und die gegenseitige Unterstützung, die beide Verfahrensweisen füreinander leisten können. An anderer Stelle habe ich begründet, dass Evaluation für bestimmte Zwecksetzungen die Methode der Wahl ist und durch Qualitätsmanagement nicht substituiert werden kann¹. Umgekehrt gilt dies für andere Aufgabenstellungen ebenso, z.B. dafür, dass sozialrechtlich präzise definierte Abläufe und Mindeststandards mit hoher Sicherheit tatsächlich eingehalten werden.

Wer hohe Qualität – wie immer zu konkretisieren – anstrebt, muss sich beider Verfahren bedienen, so lautet die hier vertretene zentrale These. Die folgenden Ausführungen sollen diese begründen und dazu beitragen, deren Verhältnis zueinander zu klären und zu einem Klima pragmatischen Zusammenwirkens von Vertretern des Qualitätsmanagement sowie der Evaluation beitragen.

Auf eine vorangestellte Definition beider Begriffe wird verzichtet – die folgenden thesenartigen Ausführungen sollen diese begriffliche Klärung durch einen Vergleich befördern.

Ähnlichkeiten zwischen Evaluation und Qualitätsmanagement*

1. Beide Verfahren kommen ursprünglich aus den USA. Evaluation, geboren in den 20er Jahren, verbreitete sich in den 60er Jahren („Sputnik-Schock, Wiederentdeckung der Armut“) und reifte Mitte der 70er Jahre zu einer eigenständigen sozialwissenschaftlichen Methode. Das moderne Qualitätsmanagement wurde Anfang der 50er Jahre von Amerikanern erfunden (Deming / Juran), zunächst nach Japan exportiert und schließlich in den 70er und frühen 80er Jahren durch die Automobilindustrie nach Amerika re-importiert. Beide Methoden tragen charakteristische Züge ihrer kulturellen Herkunft. Es ist dies insbesondere das pragmatische amerikanische Grundverständnis, dass die Anwendung rationaler Verfahren und Techniken zu Verbesserungen führt.

2. Beide Verfahrensweisen sind „Symbole der Modernisierung“. Kaum ein Trägervorstand / Einrichtungsleiter / Behördenleiter kann sich heute dem Anspruch verweigern, systematische Verfahren der Qualitätsentwicklung einzuführen. Die Gefahr besteht allerdings auch darin, dass Qualitätsentwicklung (Qualitätsmanagement oder Evaluation) sich schließlich als bloße Rituale erweisen könnten.

3. Studien weisen nach, dass es häufig misslingt, Qualität unter Nutzung systematischer Verfahren zu entwickeln. Dies gilt für beide Vorgehensweisen: Das Qualitätsmanagement bleibt in ein perfektes Handbuch verpackt in den Regalen stehen – wird aufgrund von Überkomplexität nicht umgesetzt. Aufwendig erzeugte Evaluationsergebnisse verstauben nicht selten auf Fensterbänken. Sie werden weder für Entscheidungen noch systematische Verbesserung genutzt. Beides sind also keine perfekten Sozialtechnologien, sie geben insbesondere keine Umsetzungsgarantie (kann / sollte es diese für Humandienstleistungen überhaupt geben?). Vielmehr sind es zuweilen risikoreiche Experimente, von denen einige fehlschlagen – und hoffentlich viele gelingen.

4. Beide Verfahrensweisen konzentrieren sich – inzwischen – stark auf Kunden, Klienten, Nutzer oder – umfassender – Beteiligte & Betroffene (stakeholder). Ursprünglich waren die Verfahren stark an den Perspektiven von Produzenten / Fachkräften / Auftraggebern / Entscheidern orientiert: Evaluatorinnen verstanden sich lange Zeit als unabhängige, neutrale Expertinnen, die wissenschaftliche Experimente durchführen, um kausale Zusammenhänge öffentlicher Programme aufzudecken und ihre Effektivität objektiv (i.S. von durch wissenschaftliche Theorien begründet) zu messen. Ebenso war das Qualitätswesen lange durch Expertenvorstellungen von dem dominiert, was Qualität ausmacht: Das beste Urteil wurde Ingenieuren, Werkstoffkundlern und anderen Produktionsexperten zugetraut.

Dies hat sich drastisch geändert. Beide Systeme wollen ihren Nutzern optimale Informationen und Verfahrensweisen zur Verfügung stellen. In beiden Systemen verstärkt sich der Trend, Kunden, Beteiligte & Betroffene usw... aktiv in den Prozess der Qualitätsentwicklung, insbesondere auch in die Festlegung von Zielen und Bewertungskriterien einzubeziehen (wenn auch Form und Reichweite derartiger Partizipation weiterhin umstritten ist).

5. Eng damit korrespondierend ist Beteiligung von Mitarbeitern/-innen inzwischen ein Merkmal beider Verfahrensweisen. Dies gilt für das Qualitätsmanagement allerdings in dieser Klarheit nur für solche sozialen Dienste und Einrichtungen, deren Leitungskräfte Partizipation als hohen professionellen Wert vertreten (in aller Regel nicht für medizinische Einrichtungen oder etwa den Strafvollzug). Damit spielt die Initiierung und Steuerung von Aushandlungsprozessen in beiden Verfahren eine wichtige Rolle.

6. Beide Verfahren geraten leicht in Konflikt mit dem Anspruch fachlich-professioneller Autonomie. Es überrascht nicht, dass diese Spannungen am stärksten in hochprofessionalisierten öffentlichen Dienstleistungsfeldern sind, ganz besonders im Bereich Gesundheit und Erziehung: Ärzte, Lehrer und Universitätsprofessoren sehen es besonders ungern, wenn „Experten“ anderer Professionen „fremde“ Vorstellungen über Qualität ins Spiel bringen, auch und gerade wenn es die der Patienten, Schüler und Studierenden sind.

Unterschiede zwischen Evaluation und Qualitätsmanagement²

1. Die Verfahren kommen aus verschiedenen makroökonomischen Sektoren: Qualitätsmanagement kommt aus dem privaten, industriellen Sektor. Es hat sich von dort auf den privaten Dienstleistungssektor und schließlich auf den öffentlichen Sektor ausgeweitet. Evaluation hingegen hat ihre Ursprünge im öffentlichen Sektor und trägt viele Kennzeichen und Denkweisen aus Wohlfahrtsverbänden, öffentlichen Dienstleistungseinrichtungen u.ä. Organisationen. Evaluation betont deutlich den Unterschied zwischen Effizienz und Effektivität und stellt die Frage der Wirksamkeit von Prozessen und Programmen in den Vordergrund. Qualitätsmanagement beschäftigt sich originär mit Effizienz, da diese mit dem herausragenden Ziel privatwirtschaftlichen Wirtschaftens (Gewinnerzielung) eng korrespondiert.

2. Die Einführung beider Verfahren wird zwar eher von externen Beratern angeleitet als durch Organisationsinterne. Dies gilt beim Qualitätsmanagement lediglich für die Einführungsphase. Anspruch ist hier, dass es sich auf Dauer selbständig im Unternehmen / in der Organisation etabliert und von internen Kräften selbständig getragen wird. Auch gibt es zunehmend Versuche, interne Evaluationen durchzuführen oder Selbstevaluationen von Basis-Fachkräften einzuleiten. Allerdings gibt es aus der Evaluationsgeschichte zahlreiche Erfahrungen über die schwierig zu bewältigenden Rollenkonflikte, die sich aus dieser Doppelzuständigkeit – Programmverantwortlicher und Evaluatorin – ergeben. Evaluation wird daher sinnvollerweise nicht als rein internes Verfahren etabliert werden.

Evaluation beschäftigt sich seit Jahrzehnten mit der Frage, wie die unterschiedlichen internen und besonders auch externen Hierarchien und Einflussgruppen in Untersuchungs- und Bewertungsprozesse angemessen einbezogen werden können (ohne dass es hierzu Konsens gäbe). Qualitätsmanagement leistet ebenfalls diesen Aushandlungsprozess im Inneren, hat dabei weiterhin Schwierigkeiten, insbesondere die für soziale Dienstleistungen sehr relevante Beteiligtegruppe „Politiker“ angemessen einzubeziehen. Entweder behandelt es sie – aus dem industriellen Qualitätsmanagement übertragen – als eine Art oberste Leitungsebene oder blendet sie – da auf die Leitung der Organisation fokussiert – ganz aus. Auch

konzentriert Qualitätsmanagement das Aushandeln faktisch – allen selbst postulierten Zyklen (bei Deming, im aktualisierten EFQM- Modell usw. zum Trotz) – auf die Einführungsphase, während Evaluation dies aus ihrer Logik heraus zu Beginn jedes zyklischen Evaluationsdurchlaufs erneut leisten muss.

3. Qualitätsmanagement ist umfassend angelegt („total“). Dieses Merkmal wird besonders bei der Ist- Analyse, einem der ersten Schritte im QM- Prozess deutlich. Alle Einheiten der Organisation, alle Segmente der Dienstleistungserstellung, alle Schlüsselprozesse, alle Schnittstellen zwischen den Prozessen sind danach einzuschätzen, in welchem Maße „Qualität“ erreicht ist, wie weit fortgeschritten die Einführung von QM ist. Raster hierfür sind z.B. die 20 Elemente der DIN EN ISO 9000, das SERV-AS Schema (gerade in sozialen Einrichtungen verbreitet), die sog. Selbsteinschätzungsinstrumente nach EFQM (mit den zugrunde liegenden neun „Kriterien“).

Evaluation hingegen ist (abgesehen von Versuchen einer „comprehensive evaluation“) auf bewusst gewählte Ausschnitte des zu beschreibenden und bewertenden Gegenstandes ausgerichtet („fokal“). Die Entscheidung über den zu evaluierenden Ausschnitt wird festgehalten in mit Prioritäten versehenen Evaluationsfragestellungen. Diese Entscheidungen sind Ergebnis eines mehr oder minder partizipativen Aushandlungsprozesses zwischen den relevanten Beteiligten & Betroffenen (stakeholder), die ihre „Einsätze“ („stakes“) ins Spiel bringen, wenn es um die Prioritätensetzungen für die Evaluation geht.

4. Evaluationen beschränken sich in der Regel auf einen der möglichen Zuschnitte von Gegenstandsbereichen – Organisation, Personal oder Programm.

◇ Organisationsevaluationen werden häufig im Zusammenhang von Akkreditierungsverfahren durchgeführt. Ein freiwilliger Zusammenschluss von Humandienstleistungsanbietern einer bestimmten Branche – z.B. Erziehungsheime – verabschiedet Kriterien und Standards für Organisationsqualität. Um durch den Zusammenschluss anerkannt (akkreditiert) zu werden und damit nach außen hin werben zu dürfen, unterwerfen sich seine Mitgliedsorganisationen regelmäßigen Audits, die bei Erfolg zur erstmaligen oder zur Wiederholungsakkreditation führen.

◇ Personalevaluationen werden im Zusammenhang mit Mitarbeiterbeurteilungen, Zielvereinbarungen im Rahmen der Personalentwicklung oder im Rahmen der Karriereplanung durchgeführt. Sie unterliegen in Deutschland den Regelungen des Betriebsverfassungsgesetzes (und analogen Bestimmungen bezüglich von Personal- und Mitarbeitervertretungen). In den USA gibt es ein spezielles Regelwerk zur „Evaluation von Personal“, das von der amerikanischen Evaluationsgesellschaft und wichtigen Berufs- und Fachgesellschaften gemeinsam verabschiedet ist.

◇ Programmevaluationen richten sich auf programmförmig beschriebene und durchgeführte „Bündel von Maßnahmen“, die aus einer Folge von Aktivitäten/ Interaktionen bestehen, auf der Basis von ausgewiesenen Ressourcen durchgeführt werden und darauf gerichtet sind, bestimmte, bei bezeichneten Zielgruppen zur erreichende Ergebnisse auszulösen. In der Jugendhilfe zählen dazu pädagogische Projekte, Maßnahmen, Veranstaltungen und Kurse usw. , die auch

als „Angebote“, „Hilfformen“ oder „Produkte“ bezeichnet werden. Gegenstand der Evaluation ist hier sowohl das Konzept des Programms (Programm als Plan) als auch die Durchführung des Programms und seine Resultate (Programm in Aktion).

Im Unterschied zum Qualitätsmanagement, das in allen drei Gegenstandsbereichen gleichzeitig ansetzt und darauf abzielt, Verknüpfungen und Schnittstellen zwischen Organisation, Personal und Programm zu identifizieren und zu optimieren, insistiert Evaluation auf der Trennung und ggf. sequentiellen Bearbeitung je eines Gegenstandsbereiches (ggfs. durch entsprechende Teilevaluationen).

5. Die Periodizität und das damit verbundene Modell der Einwirkung der beiden Verfahren ist unterschiedlich ausgeprägt: Qualitätsmanagement ist ein kontinuierlich durchgeführtes Vorhaben. Ziel ist es, dauerhaft zur Stabilisierung der Prozesse beizutragen, sie wenn möglich zu standardisieren, und Qualitätsinstrumente als Regelbestandteil in die Erstellung von Dienstleistungen einzubauen. Es folgt der Annahme, dass zunehmende Sicherheit des Leistungsprozesses erreicht wird („Kumulation“). Evaluation hingegen ist üblicherweise eine periodische Aktivität, die für eine bestimmte Zeitphase durchgeführt wird. Nach einer gewissen Zeitspanne beginnt Evaluation mit abgewandelten oder ganz anderen Fragestellungen erneut, um ausschlaggebende Anstöße für Entscheidung oder Verbesserung zu geben („Katalyse“). Evaluation sollte unterschieden werden vom Dokumentationswesen, vom Berichtswesen oder anderen Formen des kontinuierlichen Monitoring³ (wie auch des Controllings, auf das in diesem Beitrag nicht eingegangen wird).

6. Qualitätsmanagement steht den verschiedenen Ansätzen der Managementlehre, z.B. „Führen durch Ziele“, Leistungsverbesserungs-Management (performance management) usw... nahe. Es ist Teil des Managements. Evaluation hingegen ist keine originäre Managementaufgabe, sondern eine Beratungs- oder Stabsfunktion. Hingegen ist Qualitätsmanagement (zunehmend) unverzichtbare Aufgabe der obersten Leitung.

„Unabhängigkeit“ der Evaluatorin ist folglich ein wichtiges Thema (ganz unterschiedlich beantwortet in den verschiedenen Schulen der Evaluation), während es als Thema für das Qualitätsmanagement eine untergeordnete Rolle spielt. An dieser Stelle würde es sich lohnen, den Unterschied zwischen distanzierten, objektivistischen sowie kontaktintensiven / partizipativen Evaluationsansätzen herauszuarbeiten und zu klären, ob es im Qualitätsmanagement eine ähnliche Bandbreite gibt.

7. Das „Audit“, die Überprüfung der festgelegten Verfahrensweisen auf Umsetzung durch einen internen oder externen Auditor, teils verbunden mit einem Zertifizierungs- oder Akkreditierungsverfahren, ist oft tragender Bestandteil des Qualitätsmanagements. Das Audit setzt voraus, dass Verfahrensregelungen, Qualitätsziele und Kriterien verbindlich (organisationsweit) eingeführt sind, die durch leicht beherrschbare Beobachtung und Befragung überprüfbar sind. Diese Auditoren-/Prüferrolle kommt in der Evaluation nicht vor. Die Evaluatorin ist nicht nur für den datenmäßigen Soll- Ist- Abgleich zuständig, sondern auch für die

Aushandlung von Zielen, Abgrenzung von Gegenständen der Beschreibung und Bewertung, Planung und Durchführung speziell darauf zugeschnittener Datenerhebungen, Erstellung von auf die Nutzergruppen zugeschnittenen Berichten. Die Rolle der Evaluatorin ist im Qualitätsmanagement auf verschiedene Teilrollen aufgeteilt (Qualitätsleiter, Qualitätsbeauftragter, Auditor usw.) und umfasst zusätzliche Elemente.

Gegenseitige Unterstützungsfunktionen von Evaluation und Qualitätsmanagement

Funktionen von Evaluation für Qualitätsmanagement

Evaluationen – besonders in ihren offenen, ungerichteten, responsiven Formen – können sich mit der gesamten Komplexität humaner Dienstleistungserstellung beschäftigen. Sie können dabei sehr unterschiedliche Schwerpunktsetzungen bilden, z.B. stärker das professionelle Handeln in den Blick nehmen, die bei den Zielgruppen ausgelöst oder nicht ausgelöst Outcomes oder auch (nicht erwünschten) nicht intendierten Wirkungen, die Eignung der strukturellen, organisationalen oder finanziellen Rahmenbedingungen für reibungslose Prozesse / befriedigende Ergebnisse usw... Evaluation ist hochgradig flexibel an die jeweils in einem System aktuellen Informations-, Verbesserungs- und Entscheidungsbedarf anpassbar.

Evaluation beginnt typischerweise mit einem Aushandlungsprozess, der viele relevante Beteiligte einbezieht. Bis dieser – mit dem Ende der Phase der Evaluationsplanung – bis zu einem nächsten Evaluationszyklus abgeschlossen ist, also bis die konkret zu folgenden Evaluationsfragestellungen feststehen, bleibt die Komplexität des nur vorläufig abgegrenzten Evaluationsgegenstandes erhalten. Aus dieser zunächst komplexitätserhaltenden Vorgehensweise ergeben sich nützliche Funktionen von Evaluation für das Qualitätsmanagement:

a) Evaluation kann dazu beitragen, dass die Wertmaßstäbe geklärt werden, auf deren Basis die Beteiligten und Betroffenen das Programm beschreiben und bewerten. Werte werden konkretisiert und als schriftlich formulierte Texte den relevanten Beteiligten und Betroffenen zugänglich gemacht. Dadurch gelingt es, den Bereich des vorhandenen Wertkonsenses zwischen verschiedenen Beteiligten sichtbar zu machen und ggf. zu erweitern, aber auch herauszuarbeiten, welche kurzfristig nicht minderbaren Wertkonflikte bestehen. Dies gilt nicht nur für organisationsweite Wertkonflikte, z.B. die zwischen verschiedenen Hierarchieebenen, sondern auch solche, die zwischen Fachkräften der Dienstleistungsbasis bestehen (die typischerweise in pädagogischen und therapeutischen Settings stark ausgeprägt sind, gerade auch zwischen den Angehörigen verschiedener Professionen). Für eine erfolgsorientierte Entwicklungsstrategie klärt Evaluation, bezüglich welcher Bewertungselemente und Evaluationsfragestellungen Daten und Informationen bereit gestellt werden sollen, für die erwartet werden kann, dass sie von einer großen Zahl von Beteiligten rezipiert und auch akzeptiert werden. Andererseits werden Bereiche mit starken Wertkonflikten identifiziert,

in denen – bei gegebenen Macht- und Einflußverhältnissen – erst über längere Dauer systematische Qualitätsentwicklung erfolgreich sein kann.

b) Evaluation unterstützt die Programmverantwortlichen und andere Beteiligte, die auf das Programm Einfluss haben, die Programmziele zu klären, d.h. deren Erreichbarkeit und realistische Terminierung zu überprüfen. Insbesondere kann Evaluation dazu beitragen, operationale, messbare Zielbeschreibungen zu erstellen. Diese Leistungsfähigkeit hat sie – im Unterschied zum Qualitätsmanagement – insbesondere für die Kernprozesse der Humandienstleistung, also diejenigen Leistungen, die im unmittelbaren Zusammenwirken von Erziehendem und Erzogenem, Bildendem und Gebildetem, Beratendem und Beratenem entstehen. Hiermit trägt sie zur fachlichen Qualifizierung von Mitarbeitern und Mitarbeiterinnen bei.

c) Evaluation kann plausible Belege dafür bereitstellen, welche Interventionen zu welchen Zielerreichungen bzw. Ergebnissen führen. Sie unterstützt damit unmittelbar das fachliche Bemühen um die Optimierung der Zielerreichung. Sie trägt dazu bei, dass die „Black Box“ ausgeleuchtet wird (die Kernprozesse), in denen die eigentliche Erstellung der Humandienstleistungen (in Koproduktion zwischen Fachkräften und Klienten) stattfindet. Dies ist mit Methoden des Qualitätsmanagements lediglich in Randbereichen systematisch beschreibbar und bewertbar.

d) Evaluation kann schließlich dazu dienen, gegebene oder in Entwicklung befindliche Qualitätsmanagement- Systeme zu beschreiben und zu bewerten. Sie bringt dabei über die Perspektiven von Kunden (im engeren Sinne) und Leitung sowie Mitarbeitern/innen andere Perspektiven ein, z.B. die von Fachpolitikern (z.B. aus Jugendhilfeausschüssen) oder von externen Helfern (Psychotherapeuten, Arbeitstherapeuten usw.).

e) Evaluation leistet auf ihrem Hintergrund der empirischen Sozialforschung die Entwicklung und Validierung von Datenerhebungsinstrumenten (Befragung, Beobachtung, Inhaltsanalyse und insbesondere auch interventionsnahe Datenerhebungsinstrumente) und ergänzt somit die klassischen Instrumente des Qualitätsmanagements. Evaluation hat in den vergangenen Jahrzehnten das Methodeninventar der empirischen Sozialforschung an die Erfordernisse humandienstleistender Einrichtungen angepasst. Sie leitet dazu an, praktikable, ressourcenschonende Erhebungsinstrumente zu entwickeln, ohne bei den Gütekriterien für diese Instrumente (insbesondere Validität und Reliabilität) nicht hinnehmbare Kompromisse einzugehen. Sie verfügt über eine den Erfordernissen der Humandienstleistungen angepasste Methodenlehre, die sowohl methodischem Dilletantismus wie unpraktikabler methodenzentrierter Perfektion entgegen tritt. Evaluatorinnen können somit im Rahmen des Qualitätsmanagement das Wissen und die Erfahrung zum von der Datenmenge her Bewältigbaren und von der Datenqualität her Unverzichtbaren fördern.

Funktionen des Qualitätsmanagements für Evaluationen

Qualitätsmanagement leistet unmittelbar eine Klärung und Stabilisierung der Aufbau- und Ablauforganisation von sozialen Diensten und Einrich-

tungen. Es klärt die Verkettung von unterschiedlichen Teilprozessen, identifiziert Schnittstellen und expliziert (gerade interne) Kunden- Lieferanten- Beziehungen. Dadurch entsteht ein umfassendes, gegliedertes und detailliertes Gesamtbild aller Prozesse, typischerweise symbolisiert durch Flussdiagramme. Die Verschriftlichung zentraler Leitziele, Schlüsselprozesse und Verfahrensweisen schafft eine verlässliche Basis für Evaluationen. Dies gilt insbesondere für summative Evaluationen, die Stabilität und Reife des zu beschreibenden und bewertenden Programms und der diesen tragenden Organisation zur Voraussetzung haben. Dies betrifft:

a) Die Klärung von Leitbildern und von Leitziele der Organisation im Rahmen eines Leitbildprozesses, in den strategische zentrale Qualitätsziele ebenso wie Kundenanforderungen systematisch eingearbeitet sind. Wenn Werte geklärt und Leitziele bekannt gegeben sind, sind verlässliche Grundlagen für die Zusammenarbeit über die Abteilungen und Hierarchieebenen hinweg geschaffen: Die Leitungsebene hat sich zu Interessen bekannt und Richtungen vorgegeben; Mitarbeiter können jetzt für sich erarbeiten, ob ihre Handlungsziele (die sie an der pädagogischen Basis verfolgen und bekannt geben sollen) mit denen der Organisation, von der obersten Leitung verfolgten übereinstimmen.

b) Klärung von Zuständigkeiten (von Leitungspersonen und Gremien), von Entscheidungsabläufen, des Protokoll- und Dokumentationswesens. Damit ist innerorganisatorisch vorgeklärt, wer an der Planung von Evaluation zu beteiligen ist und an wen unverzichtbar in welcher zeitlichen Abfolge Bericht zu erstatten ist. Oft wird erstmals mit dem QM eine „Organisationsspinne“ erarbeitet und veröffentlicht. Die Orientierung von externen aber auch internen Evaluatoreninnen im Geflecht der sachlichen und Entscheidungszuständigkeiten wird damit erleichtert.

c) Eine umfassende Ist- Analyse zu Beginn des QM-Prozesses schafft einen Überblick über die gesamte Organisation und fixiert – oft erstmals – für alle Organisationsmitglieder schriftlich, welche Elemente zusammengenommen das Ganze der Organisation und ihrer Leistungen ausmachen. Wenn eine organisationsweite Selbsteinschätzung (z.B. nach EFQM) durchgeführt wird, gibt es erste Klärungen zu Stärken und Schwächen, an welche die Evaluationsplanung – als eine unter mehreren Quellen – anschließen kann.

d) Basisdaten zur Organisation, zu Input-, Prozess- und Outputdaten werden bereitgestellt (z.B. Zahl und Merkmale der in das Dienstleistungsprogramm eintretenden Kunden / Klienten, monetäre und nicht- monetäre Daten zu Umfang, zeitlicher Abfolge und örtlicher Planung von Leistungsprozessen, Statistiken zu Teilnehmerzahlen, Leistungsstunden, Aufnahmen, Austritten).

e) Qualitätsmanagement schafft Klarheit über die Ablauforganisation. Seine besondere Leistung besteht darin, die aus Sicht des Qualitätsmanagement zentral wertschöpfenden (sich wiederholenden) Schlüsselprozesse zu identifizieren und abzugrenzen und damit eine Beschreibung des Kontextes von zu evaluierenden (Teil-)Programmen zu leisten. Gleichzeitig werden die Schnittstellen identifiziert, die besonders fehleranfälligen Orte der „Übergabe“ von Prozessver-

verantwortlichkeit und Informationen. Die Basis für team-, bereichs-, abteilungs- und einrichtungsübergreifendes systematisches Zusammenwirken ist damit gelegt.

f) Qualitätsmanagement trägt (neben dem Controlling) dazu bei, die finanziellen Rahmenbedingungen, in denen Humandienstleistungen produziert werden, transparent zu machen. Es verdeutlicht damit die ökonomischen Rahmenbedingungen, in denen die zu evaluierenden Programme durchgeführt werden. Damit markiert es auch die von der Evaluation zu beachtenden Grenzen und Anforderungen von sozialen Einrichtungen und Diensten als Betriebe bzw. Unternehmen.

Fazit

Bei Ähnlichkeiten werden beide Verfahrensweisen dauerhaft unterschiedliche Leistungspotentiale aufweisen. Oft werden sie erst in ihrer Kombination zu einer optimalen Qualitätsentwicklung der Humandienstleistungen beitragen.

Qualitätsmanagement wird insbesondere in größeren Organisationseinheiten standardmäßig eingesetzt werden. Oberste Leitung kann gleichzeitig Qualitätsleitung sein, denn QM ist ein – auf Zustimmung der Mitarbeiter angewiesener – top- down-Prozess. Interne Qualitätsfachleute werden das Qualitätsmanagement, ist es einmal etabliert, selbständig führen.

Evaluation ist – und wird es bleiben – eine Methode der Wahl, zu der sich eine Organisation (oder ein Teil einer Organisation) entschließt. Die Oberste Leitung wird nicht selbst die Steuerung von Evaluationsprozessen in die Hand nehmen. Hierfür ist ein Minimum an Interessenneutralität erforderlich, die einer externen Evaluatorin oft eher als einer internen zugeschrieben wird. Auch wenn interne Evaluatorinnen Akzeptanz finden, werden sie zumindest sporadisch auf externe Evaluationsberatung zurückgreifen.

Die organisationsinternen Rollen „Qualitätsleiter“, „Qualitätsauditor“ und „Evaluatorin“ sollten getrennt beschrieben und entsprechend sollten Aufgabenprofile konkretisiert werden. Diese Empfehlung macht vielleicht die ernüchternde Botschaft dieses Beitrages deutlich: Es gibt zwar Ausnahmefälle von Humandienstleistungen (oder von Segmenten ihrer Erstellung), in denen QM alleine ausreicht (z.B. „Pflege auf der Intensivstation“) oder in denen Evaluation alleinstehend die Entwicklung von Qualität tragen kann (z.B. im kleinen Gemeinwesenarbeitsteam eines Stadtteilvereins). Meist ist jedoch beides vonnöten, beispielsweise im großen Feld der ambulanten, der stationären und teil- stationären erzieherischen Hilfen. Auf vielen Feldern wird es zu einer – es ist zu hoffen – fruchtbaren Konkurrenz beider Ansätze kommen. In einigen Jahren wird es auf Basis der bis dahin gemachten Erfahrungen eine lohnende Aufgabe sein, optimale Kombinationsmöglichkeiten für verschiedene Arbeitsfelder und institutionellen Kontexte systematisch zu bestimmen.

Fußnoten

- 1 Nutzenfokussierte Evaluation von Humandienstleistungen. Plädoyer für eine sozialwissenschaftliche Rückbesinnung in der Qualitätsdebatte; in: Sozialwissenschaften und Berufspraxis, 2/1999, S. 143-156.
- * Die Abschnitte 1 und 2 basieren wesentlich auf einem Vortrag von Christopher Pollitt, vormals Vorsitzender der UK-Evaluationsgesellschaft, heute Hochschullehrer in Rotterdam. Diesen Vortrag hielt er zum Thema „Quality Management and Evaluation in Europe - Innovation or Legitimation“ auf dem Europäischen Kongreß zu Evaluation und Qualitätsmanagement in Sozialer Arbeit und Gesundheitswesen, Universität Osnabrück, 1. bis 3. 9. 1999 (abgedruckt in Müller-Kohlenberg / Münstermann 2000)
- 2 Es gibt auch Übergänge zwischen beiden Verfahren, z.B. eine stark dokumentierende (monitoring) Evaluation oder die team- / fachkraftgebundene Selbstevaluation, wie sie maßgeblich von Maja Heiner entwickelt wurde. Auch diese nahe am Qualitätsmanagement stehenden Formen der Evaluation weisen eine beschränkte Leistungsbreite auf, innerhalb derer sie besonders effektiv sind. Damit stellen sie die These einer grundlegenden Verschiedenheit von QM und Evaluation m.E. nicht in Frage.
- 3 Mit „dokumentierender“ oder „Monitoring-Evaluation“ existieren Grenzformen, welche eine leichte Unterscheidbarkeit wiederum erschweren - Grenzziehungen sind ihrer Natur nach willkürlich.

Literatur

Beywl, Wolfgang, Nutzenfokussierte Evaluation von Humandienstleistungen. Plädoyer für eine sozialwissenschaftliche Rückbesinnung in der Qualitätsdebatte; in: Sozialwissenschaften und Berufspraxis, 2/1999, S. 143-156.

Joint Committee on Standards for Educational Evaluation; Sanders, James R. (Hrsg.) Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“ Opladen: Leske + Budrich 2000 (2. durchgesehene Auflage)

Meinhold, Marianne, Qualitätssicherung und Qualitätsmanagement in der Sozialen Arbeit, Freiburg: Lambertus 1996

Merchel, Joachim (Hrsg.) Qualität in der Jugendhilfe. Kriterien und Bewertungsmöglichkeiten, Münster: Votum /VVA 1998

Müller, Burkhard, Qualitätsprodukt Jugendhilfe. kritische Thesen und praktische Vorschläge, Freiburg: Lambertus 1996

Müller-Kohlenberg, Hildegard; Münstermann, Klaus (Hrsg.) Bewertung von Humandienstleistungen. Evaluation und Qualitätsmanagement in Sozialer Arbeit und Gesundheitswe-

Beywl, Wolfgang; Dr., Sozial- und Wirtschaftswissenschaftler,
Geschäftsführender Vorstand von UNivation e.V., Institut für Evaluation
und wissenschaftliche Weiterbildung Köln (<http://www.univation.org>;
Vorstandsmitglieder der Deutschen Gesellschaft für Evaluation
(<http://www.degeval.de>)

Programmevaluation in pädagogischen Praxisfeldern: begriffliche und konzeptionelle Grundlagen¹

Wolfgang Beywl

„Man bildet sich zwar insgemein ein, dass Experimente bei der Erziehung nicht nötig wären, und dass man schon aus der Vernunft urteilen könne, ob etwas gut oder nicht gut sein werde. Man irret hierin aber sehr, und die Erfahrung lehrt, dass sich oft bei unsern Versuchen ganz entgegengesetzte Wirkungen zeigen von denen, die man erwartete. Man sieht also, dass, da es auf Experimente ankommt, kein Menschenalter einen völligen Erziehungsplan darstellen kann.“²

„The phrase “reality testing” implies that being “in touch with reality” can’t really be assumed.“³

In beiden Sätzen – im Abstand von 200 Jahren verfasst – geht es um das Bemühen der pädagogischen Profession und Wissenschaft, bei aller Vielgestaltigkeit, Veränderlichkeit und Offenheit ihrer Gegenstände die zugrundeliegenden und durch Bildung und Erziehung beeinflussten empirischen Wirklichkeiten systematisch zu untersuchen, um gute Praxis zu sichern und weiterzuentwickeln: Es geht um die Aufgabe der Evaluation.

Evaluation in Alltagssprache und wissenschaftlicher Terminologie

Evaluation kommt als Bezeichnung und als Konzept aus der angloamerikanischen Kultur und Sprache.⁴ Dies hat sie gemein mit anderen Ansätzen, die seit gut zehn Jahren verstärkt im deutschsprachigen Raum Fuß fassen: (Bildungs-, Sozial-, Wissens-, Performance- und Qualitäts-) Management, Monitoring, Auditing, Self-Assessment, Benchmarking.⁵ Evaluation wurde im Unterschied zu diesen bereits in den späten 60er/frühen 70er Jahren in der deutschsprachigen Sozial- und Erziehungswissenschaft adaptiert⁶ und geriet nach kurzer Blüte - oft in Verbindung mit „Schulversuchen“⁷ - für fast 20 Jahre an den Rand des Blickfeldes praktisch und theoretisch tätiger Pädagogen.⁸

Evaluation wird – vielleicht dem Zeitgeist ihrer Ersteinführung geschuldet – in lateinischer und nicht in englischer Lautgestalt ausgesprochen, wie dies für die seit den späten 80ern „importierten“ Begriffe gilt.⁹ Evaluation zeigt sich damit durch die Lautfolge seiner Aussprache als traditionelles pädagogisches Konzept.

In der englischen Alltagssprache bedeutet *to evaluate* „den Wert oder die Größenordnung von etwas bestimmen oder festlegen.“¹⁰ Synonyme für *Evalua-*

tion sind folglich „(Ein)Schätzung, Festsetzung, Beurteilung, Auswertung, Abwägung, Bewertung.“¹¹

Die wissenschaftliche Wortbedeutung von Evaluation baut auf der alltagssprachlichen auf und fügt Bestimmungen hinzu. Die Definition des zentralen, da von 16 Fach- und Berufsverbänden erarbeiteten und als *American National Standard* anerkannten Referenzwerkes ist denkbar knapp: „Evaluation: Die systematische Untersuchung der Verwendbarkeit oder Güte eines Gegenstandes.“¹² Nicht ganz so knapp formulieren es die von der American Evaluation Association (AEA) verabschiedeten „Leitprinzipien für Evaluatoreninnen“¹³, der mit ca. 2.200 Mitgliedern weltgrößten Berufsorganisation: „Evaluatoreninnen führen systematische, auf Daten gestützte Untersuchungen über das jeweilige Evaluationsobjekt durch. „Sie streben danach, die bestmöglichen Informationen über den Wert dessen, was auch immer bewertet werden soll, herauszuarbeiten und bereitzustellen.“¹⁴

An dieser Stelle sei folgende allgemeine Definition vorgeschlagen: Wissenschaftliche Evaluation bezeichnet die Summe systematischer Untersuchungen, die empirische, d.h. erfahrungsbasierte Informationen bereitstellen über den Wert eines (in der Regel sozialen) Gegenstandes, den Evaluationsgegenstand.

Damit ist die Grenzlinie zu informellen Formen der Bewertung, sei es im Alltagsleben oder in der Berufspraxis, markiert: Weder bei der Inaugenscheinnahme eines Hotelzimmers noch beim Flurgespräch des Fachbereichsleiters mit der Lehrperson handelt es sich um „systematische Evaluation“. Auch das zu Seminarende eingeholte „Blitzlicht“ – Kommentare der Teilnehmenden – oder die Jahresstatistik über die Inanspruchnahme der Weiterbildungsabteilung machen allein noch keine Evaluation; diese Aktivitäten werden dann Elemente einer Evaluation, wenn sie über einen expliziten Evaluationsplan in den Prozess systematischer Informationsgewinnung und –vermittlung eingebunden sind.

Beide prominente Definitionen von Evaluation sind so weit gehalten, da die (z.B. durch die Mitglieder der AEA bearbeiteten) Evaluationsgegenstände aus einer Vielzahl von Feldern kommen und Evaluatoreninnen sich mit unterschiedlichen Typen von Evaluationsgegenständen beschäftigen:

Evaluationsstypen: Evaluation von Politiken (policies), Regierungen, Organisationen, Technologien, Forschungen, Theorien, Programmen, Leistungen, Vorschlägen, Produkten, Personal und von Evaluation selbst.¹⁵

Evaluationsfelder: Evaluation im Bereich Bildung und Erziehung, Gesundheit und Sport, psychosoziale Versorgung, Soziale Arbeit und Dienste, Soziale Sicherung, Einkommen und Steuern, Arbeitsmarkt und Beschäftigung, Wohnungs- und Städtebau, Rechtspflege, Kriminalitätsprävention und Strafvollzug, Entwicklungshilfe, Landwirtschaft, Umwelt, Wirtschaftsförderung, Medien und Software.¹⁶

Bevor ich mich dem Typus Programmevaluation im Evaluationsfeld Bildung und Erziehung ausführlicher zuwende, soll nachfolgend die ebenfalls per Konvention zu ziehende Grenzlinie zur Forschung kenntlich gemacht und die aktuelle Bedeutung der Evaluation im politischen und ökonomischen Umfeld verdeutlicht werden.

Unterscheidung von Evaluation und Forschung

Während in den USA (und anderen Staaten des angelsächsischen Sprachraums) weitgehend Einigkeit herrscht, dass Evaluation eine eigenständige, interdisziplinäre wissenschaftliche Vorgehensweise darstellt¹⁷, steht in Deutschland die Klärung am Anfang. Die Psychologen Jürgen Bortz und Nicola Döring streiten in ihrem Standardwerk „Forschungsmethoden und Evaluation für Sozialwissenschaftler“ der Evaluation Eigenständigkeit ab:

„Die Evaluationsforschung befasst sich als ein Teilgebiet der empirischen Forschung mit der Bewertung von Maßnahmen oder Interventionen. ... Wir teilen damit die Auffassung vieler Evaluationsexperten, die in der Evaluationsforschung ebenfalls keine eigenständige Disziplin sehen, sondern eine Anwendungsvariante wissenschaftlicher Forschungsmethoden auf eine spezielle Gruppe von Fragestellungen.“¹⁸

Drei von fünf der zitierten „Evaluationsexperten“ haben zwischenzeitlich Neuauflagen veröffentlicht und sich darin dem Verständnis einer eigenständigen Evaluation angenähert, wenn nicht angeschlossen: Wottawa und Thierau argumentieren, dass das Rollenverständnis der Evaluators teilweise im Widerspruch stehe zu den Ansprüchen der wissenschaftlichen Forschung. Wissenschaft sei zwar „Basis und Hilfe für die Evaluationsarbeit“, doch sei diese als „besondere Form des wissenschaftlich-empirischen Arbeitens“ anzusehen.¹⁹ Rossi, Freeman, Lipsey halten in der sechsten Auflage ihres (soziologischen) Standardwerkes zwar daran fest, Evaluation sei zweckgerichtete angewandte Forschung, doch konzedieren sie vielfach gegenüber klassischer Forschung Besonderheiten: Evaluatoreninnen wollen zur Lösung praktischer Probleme beitragen, sind selten auf enge Gegenstandsfelder und oft interdisziplinär ausgerichtet; ethische Grenzverletzungen schlagen bei auf praktische Einwirkung angelegter Evaluation unmittelbarer auf Beteiligte und Betroffene durch; die Bewertung von Evaluationen erfolgt weniger durch wissenschaftliche Peers als durch die Auftraggeber und Nutzer.²⁰ Carol Weiss schließlich, die nach 26 Jahren eine völlig überarbeitete, den Erfahrungsschatz eines Evaluatorslebens ausbreitende Zweitaufgabe herausbringt – nicht mehr Evaluation Research sondern Evaluation überschrieben – bezeichnet in Umkehrung zur These von Bortz und Döring evaluation research als spezifische Methode der Evaluation. Als Kern der „research nature of evaluation“ sieht sie die systematische, formelle, strenge Einschätzung.²¹

Maja Heiner fasst Evaluation als einen „Zweig der angewandten Sozialforschung“. In ihren kreativen Argumenten und plastischen Anwendungsfällen einer „Experimentierenden Evaluation“ macht sie jedoch immer wieder deutlich, dass „...diese primär auf eine Optimierung der Praxis (zielt). Erkenntnisinteressen, die nicht direkt damit verknüpft sind, sind nachrangig.“²² Damit trifft sie m.E. den Kern, in dem sich Evaluation von Forschung unterscheidet: Der Evaluation geht es vorrangig um eine durch wissenschaftliche Untersuchungsmethodik vorbereitete Praxiseinwirkung. Hingegen: „Empirisch-wissenschaftliche Forschung will allgemeingültige Erkenntnisse gewinnen.“²³ Auch in ihrer praxisnächsten Variante als Praxisforschung – geht es der Forschung um „den wissenschaftlichen Diskurs und

dessen Wahrheitsansprüche. Theorieentwicklung und – überprüfung gehören zu ihren wesentlichen Aufgaben.“²⁴

Auch durch Evaluation bereitgestelltes Wissen kann konzeptionell (orientierend, aufklärend, theoriesichernd oder -erschließend) genutzt werden. Michael Patton beharrt darauf, dass die knowledge-orientation neben der Beurteilungs-, Entscheidungsorientierung – making overall judgements – und der Gestaltungs-, Verbesserungsorientierung – facilitating improvements – eine dritte Leistungsoption darstellt. Vorrangig soll evaluativ bereitgestelltes Wissen instrumentell nutzbar sein, d.h. unmittelbar als Ressource in der Praxis wirksam.²⁵ Evaluation ist somit Bestandteil der jeweiligen Praxis, ist eingebunden in fachliche/organisationale Handlungsprozesse und soll diese in aller Regel möglichst unmittelbar stützen bzw. politische oder Leitungsentscheidungen grundlegen. Schließlich sollen auch auf Wissensgenerierung angelegte Evaluationen – in einer gegenüber den bei den anderen Leistungsoptionen länger gezogenen Schleife – zur Optimierung der Praxis beitragen: „Such knowledge-generating efforts focus beyond the effectiveness of a particular program to future program designs and policy formulation in general.“²⁶

Fazit: Die beiden wissenschaftlichen Konzepte „Forschung“ und „Evaluation“ berühren sich vielfach, eine Differenzierung macht jedoch Sinn: Denn wenn von Forschung unmittelbare Verwertbarkeit gewonnener Befunde für praktische Vorhaben verlangt wird, wenn von Evaluation verallgemeinerungsfähige, theoriebegründende Erkenntnisse erwartet werden, dann provoziert dies Enttäuschungen. In einer gesellschaftlichen Entwicklungsphase, in der sozialwissenschaftliche Methoden wieder stärker gefragt sind, um Entscheidungen zu orientieren, liegt es im wohlverstandenen Interesse der Humanwissenschaften, die Stärken, aber auch Grenzen ihrer Instrumentarien klar zu benennen. Ich schlage daher erneut²⁷ vor, als idealtypische Unterscheidung²⁸ per Konvention festzulegen: Forschung und Evaluation unterscheiden sich in ihren zentralen Zielsetzungen - Mehrung theoretisch gesicherten, universal gültigen Wissens („Wahrheitskriterium“) vs. Gewinnung in bestimmten Kontexten praktisch nutzbarer Informationen („Brauchbarkeitskriterium“). Liegt der Wert der Forschung in der längerfristig angelegten Akkumulation von verallgemeinerbaren Wissensbeständen, der kritischen Analyse des Bestehenden und der Formulierung grundlegender Alternativen, so kann Evaluation sozialen und pädagogischen Programmen, Interventionen, Politiken ... an ihre Praxis unmittelbar handlungsorientierende Informationen zur Verfügung stellen, um der Öffentlichkeit ein zutreffendes Bild zu verschaffen, Beurteilungen zu ermöglichen, Entscheidungen oder Optimierung vorzubereiten.²⁹

Die Verpflichtung auf instrumentelles Wissen suspendiert Evaluation nicht davon, die Regeln zur Erhebung, Auswertung und Interpretation empirischer Daten einzuhalten. Sie muss „dem aktuellen Stand wissenschaftlicher Techniken und Methoden voll genügen.“³⁰ Allerdings führt die differente Zwecksetzung von Evaluation dazu, dass in allen Untersuchungsphasen Methoden auf die jeweiligen Evaluationsfragestellungen, vorgesehenen Informationsnutzungen und gegebenen Rahmenbedingungen zugeschnitten werden. D.h. die Untersuchungsdesigns werden oft schlanker ausfallen. Die bei quantitativen Erhebungen erreichten Skalenniveaus werden „niedriger“, der Aufwand bei der Erhebung von Daten

(z.B. geringere Stichprobengröße, Gedächtnisprotokolle statt Transskripte) wird begrenzter sein. Umfassende Fragestellungen, die aufwendige Untersuchungen erfordern würden, müssen eingeschränkt und fokussiert werden, um sie mit befriedigender Sorgfalt bearbeiten zu können. Gleichzeitig müssen effiziente Verfahren der Validierung (z.B. Triangulation von Datenquellen/Erhebungsmethoden/Auswertungsverfahren...) zum Einsatz kommen – diesen Beitrag sprengende Themen.³¹

Pointiert möchte ich zusammenfassen: Forschung und Evaluation sind wissenschaftliche Praktiken mit identischen Anforderungen an die Genauigkeit der Informationsgewinnung, aber je eigenem Geltungsanspruch an die Reichweite der bereitgestellten Erkenntnisse und die hierauf abzielenden Untersuchungsaufwendungen. Je klarer dies gesehen und anerkannt wird, desto förderlicher auch für den fruchtbaren Austausch zwischen beiden unverzichtbaren Modi wissenschaftlichen Zugangs zu sozialen Wirklichkeiten.

Evaluation im Wandel von Erziehung, Bildung und sozialer Arbeit

Die Umwälzungen der Rahmenbedingungen für die Arbeit pädagogischer Institutionen sind hinlänglich bekannt und sollen nicht erneut dargestellt werden:³² Neue Steuerung, New Public Management, Controlling in den kommunalen, Landes- und Bundeshaushalten, die Verpflichtung auf Qualitätssicherung, Qualitätsentwicklung, Wirkungsdialo, Schulprogramme in gesetzlichen Grundlagen oder Förderrichtlinien, die voranschreitende Marktorientierung, begonnen bei den Kindertageseinrichtungen über das Elementar- und Sekundarschulwesen, die Hochschulbildung bis hin zur Erwachsenenbildung in öffentlicher und freier Trägerschaft, ebenso in der Jugendarbeit oder bei den erzieherischen Hilfen.

Damit sind auch hier im Verlauf der 90er Jahre Fragen virulent geworden, welche für den amerikanischen Bildungssektor seit einem halben Jahrhundert, für den britischen seit etwa zwei Dezennien auf der Tagesordnung stehen: empirische Belege für Qualität, Bedarfsgerechtigkeit und Wirkungen, Erfolg, Effektivität und Effizienz pädagogischer Leistungen.

An anderer Stelle³³ ist ausführlicher erläutert, dass zu Beginn des zurückliegenden Jahrzehnts auf Seiten der Sozial-/Erziehungswissenschaft das Technologiedefizit in den Humandienstleistungen als unvermeidbar und strukturell nicht minderbar angesehen wurde. Es bestand ein – aus dem herrschenden Selbstverständnis der Wissenschaften hinzunehmender – Mangel an Methoden und Instrumenten zur praxisbezogenen Bearbeitung der vom politischen und ökonomischen System formulierten Anforderungen nach Qualitätsbeurteilung und -steuerung. Eben erst waren mit der „reflexiven Wende“ in der Pädagogik die subjektiven Weltansichten der Teilnehmenden/Adressaten in den Mittelpunkt des Erkenntnisinteresses gerückt. Als empirische Instrumentarien wurden damit hermeneutische, narrative, biographische – kurz – qualitative Forschungsmethoden nicht nur propagiert, sondern breit eingesetzt.³⁴

Während sich die qualitative Forschung intensiv mit der interpersonalen Ebene einer „durch interaktives Handeln konstituierten Welt, die für den Einzelnen, aber auch für Kollektive sinnhaft konstruiert ist“³⁵, beschäftigte, entstand eine Art

„Technologie-Vakuum“, in das aus anderen Disziplinen stammende Methoden „einströmten“.³⁶ Um die mit Organisation, Finanzierung oder Wirksamkeit von Einrichtungen wie Schulen, Weiterbildungsträgern, Universitäten oder Erziehungsheimen verbundenen Bewertungs- und Steuerungsaufgaben zu bearbeiten, bot sich nun das gerade erst in die deutsche Industrie eingeführte betriebswirtschaftliche Bildungscontrolling oder das ingenieurwissenschaftliche Qualitätsmanagement an. Diese schafften mit der organisationsweiten Steuerung und Sicherung von Prozessen Grundbedingungen für das Überleben von zunehmend mit markt-förmigen Außenanforderungen konfrontierten Bildungs- und Sozialeinrichtungen. Sie ermöglichen es historisch gewachsenen pädagogischen Systemen, in neuen internen und externen Vernetzungsstrukturen vielfältige Angebotsformen mit komplexen Leistungsprozessen (neu) zu organisieren.³⁷

Allein die Menge der seit Mitte der 90er Jahre publizierten „Qualitäts“-Literatur belegt die Bedeutung dieser bis dato dem deutschen Erziehungs- und Bildungssystem fremden Verfahrensweisen.³⁸ Die These wäre untersuchenswert, ob Deutschland (mit der parallel geführten „Standortdebatte“) mit seinem starken Ausbau von Controlling- und Qualitätsinstrumentarien bis hin zur Gründung der CERT-QUA³⁹ einen internationalen Sonderweg in der Bearbeitung der Qualitätsfrage in Bildung und Erziehung eingeschlagen hatte.⁴⁰

Evaluation hingegen spielte in den 90er Jahren eine nachgeordnete Rolle: Im Vergleich zu den zahlreichen gesetzlichen Bestimmungen etwa zur Qualitätssicherung und Qualitätsentwicklung in den verschiedenen Büchern des Sozialgesetzbuches (z.B. Kinder- und Jugendhilfegesetz, Arbeitsförderungsgesetz, Pflegeversicherungsgesetz) wird Evaluation bis Ende der 90er auch in den Kernbereichen von Bildung und Erziehung, also den Schul- und Hochschulgesetzen oder Weiterbildungsgesetzen, zwar (z.B. in den Kommentaren) genannt, ist aber nur in der Minderzahl der Ländern kodifiziert.⁴¹ Mit der gesetzlichen Verpflichtung öffentlicher/öffentlich finanziert erziehungs- und Bildungseinrichtungen auf Evaluation – z.B. im Gefolge der verbindlichen Einführung von Schulprogrammen oder von Studiengangsevaluationen – ist in den kommenden Jahren flächendeckend zu rechnen.⁴² Dies erfordert dringlich die konzeptionelle Klärung des wissenschaftlichen und professionellen Kerns von Evaluation unter Einbezug von Fachwissenschaftlern, Evaluatorinnen, Bildungspolitikern und Bildungspraktikern. Ein Rahmen hierfür sei nachfolgend für den im Erziehungs- und Bildungsbereich zentralen Evaluationstypus, die Programmevaluation skizziert.

Programmevaluation - zentraler Evaluationstyp in Erziehung, Bildung, sozialer Arbeit

Evaluation als komplexe Methodologie ist darauf verpflichtet, im Spannungsfeld zwischen Wissenschaft und Praxis nutzbares, d.h. empirisch gesichertes und Komplexität entlang sozialer Relevanzkriterien reduzierendes Wissen zur Verfügung zu stellen. Um im Meer von Wertpositionen, vested interests, Fachsprachen und Detailinformationen nicht verloren zu gehen, benötigen Evaluatorinnen praktikable Orientierungen. Die fachliche Substanz des methodischen Evaluationswissens wird durch vielleicht zwanzig „Navigationshilfen“ konstituiert, die

oft aus Typologien oder Klassifikationen bestehen und eine multiperspektivische Sicht auf Evaluationsgegenstände sowie systematische und nachvollziehbare Entscheidungen über Evaluationsdesigns ermöglichen.⁴³ Als wichtige Typen von Evaluationsgegenständen im Bereich Bildung und Erziehung lassen sich unterscheiden:³³

Politiken oder Politikfelder (policies)⁴⁴, also durch staatliches Handeln, insbesondere Gesetzgebung oder öffentliche Förderpraxis abgesteckte Inhaltsfelder wie Sozial-, Bildungs- oder Gesundheitspolitik.⁴⁵ Diese können auch enger gezogen sein, z.B. wie im Falle des durch das Weiterbildungsgesetz NRW geschaffenen, öffentlich verantworteten Systems der Weiterbildung in kommunaler und freier Trägerschaft oder dem des Kinder- und Jugendplans des Bundes. Kennzeichnend ist, dass grobe Leitziele gesetzt sind, Förderrichtlinien insbesondere Höhe und Abwicklung öffentlicher Finanzierung regeln, aber die Setzung konkreter Ziele der Bildungs- und Erziehungsarbeit den ausführenden Verbänden und Einrichtungen überlassen bleibt, welche die Programme, Maßnahmen, Kursformen usw. inhaltlich verantworten und umsetzen.

Programme⁴⁶: Dies sind Bündel von Maßnahmen, je bestehend aus einer Folge von Aktivitäten, basierend auf einem Set von Ressourcen, gerichtet auf bestimmte bei bezeichneten Zielgruppen zu erreichende Ergebnisse (outcomes). Ein Programm besteht aus einem fixierten (verschriftlichten) Plan oder Entwurf⁴⁷ und dessen Umsetzung in Praxis oder Handeln. Im Programmentwurf sind die Ziele, Vorgehensweisen usw. der Praxis konkret vorweggedacht. Es gibt sehr große Programme auf einer Makroebene, z.B. die AIDS-Präventionsprogramme der Bundeszentrale für gesundheitliche Aufklärung, oder das in diesem Band dargestellte Landesprogramm für Arbeitslosenzentren und Beratungsstellen für Langzeitarbeitslose mit ca. 130 lokalen Standorten. Es gibt Programme auf der Mesoebene einer Organisation oder Organisationsabteilung (ein Schulprogramm, ein Studiengang, das Hilfeangebot „stationäre Unterbringung“ in einer Erziehungshilfeeinrichtung) und Programme auf der Mikroebene (eine Unterrichtsreihe zum Thema „Müll und Recycling“, ein viertägiges Führungskräfte-Training). Programme werden auch als Projekte⁴⁸, Initiativen, Modellvorhaben, Maßnahmen, Hilfeformen, Lehr- oder Studiengänge, sowie Dienstleistungsangebote usw. bezeichnet.⁴⁹

Organisationen: Dies sind auf Dauer gestellte Systeme mit spezifischer – hier pädagogischer – Zweckorientierung, geregelter Arbeitsteilung und beständigen Grenzen⁵⁰, z.B. Kindertagesstätten, Erziehungsheime, Schulen, Hochschulen oder Weiterbildungsseinrichtungen oder deren Sub-Einheiten (z.B. Fachbereiche oder Abteilungen). Für systematische Beschreibung und Bewertung – Evaluation – von häufigem Interesse sind die Aufbau- und oder Ablauforganisation in ihrer Wirkung auf gewünschte Zweckerreichungen der Organisation, z.B. im Vergleich von zentralisierten oder dezentralisierten Entscheidungsabläufen. In Deutschland stark entwickelt ist insbesondere die Schulentwicklungsforschung⁵¹ als (formativer) Ansatz der Organisationsevaluation. Nachgeordnet ist in Deutschland die Entscheidungsvorbereitung von Zulassung/Akkreditation von (Weiter-) Bildungsträgern.⁵²

Personen: Im Bereich von Bildung und Erziehung sind dies insbesondere die Produzenten/-innen sowie Nutzer/-innen von Humandienstleistungen (z.B. Lehrende, Lernende). Für Schüler/-innen oder Studierende oder für Hilfeempfänger ist die Personenevaluation nahe an der Leistungsbeurteilung für Selektionszwecke (Versetzung, Bildungsabschluss, Zertifikat) oder mit Entscheidungen über Gewährung/Vorenthaltung von Leistungen verbunden; für Lehrende oder Fachkräfte liegt sie nahe an der Aufstiegsbeurteilung und ist auch hier unbeliebt. Auch wenn neuere Ansätze die lernunterstützende, auf Verselbständigung zielende Funktion solcher Personenevaluationen betonen und sie in Konzepte der freiwilligen Selbst-Evaluation einpassen⁵³, so werden sie von den Beurteilten (auch wenn sie dies selbst mitsteuern) nicht selten als bedrohlich erlebt. Personenevaluation sollte daher von Programm- oder Organisationsevaluation strikt getrennt durchgeführt werden.⁵⁴

Produkte: Im Bereich von Bildung und Erziehung sind dies Materialien wie z.B. Spielzeug/Spielgeräte, Verbrauchsmaterial für pädagogische Arbeit, audiovisuelle Geräte wie Projektoren, Mobiliar für Beratungs- und Schulungsräume u.v.m. Evaluationen erfolgen in der Regel analog zum vergleichenden Warentest, wobei die Passung auf pädagogische Programme eine der Bewertungsdimensionen darstellt. Komplexe Medien-/Materialienkombinationen wie z.B. Lernsoftware oder andere Selbstlernmedien, für die eine didaktisch-methodische Konzeption vorliegt, betrachte ich als ein individualisierbares Lehr-/Lernprogramm auf der (Sub-)Mikroebene.⁵⁵

Ich gehe davon aus, dass in den kommenden Jahren auch im deutschsprachigen Raum – besonders in den Gegenstandsfeldern Erziehung, Bildung und soziale Arbeit, die Programmevaluation zum dominanten Typus der Evaluation wird. Je „marktnäher“ eine Einrichtung oder Organisation ihre pädagogischen Leistungen bereitstellt, ggf. ausschreibt und verpreist, damit im Preis- und Qualitätsvergleich mit anderen Anbietern steht, desto mehr bietet es sich an, Programme als relativ stabile, gleichzeitig relativ gut anpaßbare und mittelfristig flexible Träger von Qualitätsmerkmalen zum strategischen Ansatzpunkt für die Steuerung und Optimierung pädagogischer Leistungen zu machen. Auch Finanziers, seien es öffentliche Haushalte, Sponsoren, Organisationskunden (z.B. Weiterbildungsleistungen einkaufende Betriebe) oder Endabnehmer werden sich bei Auswahl- und Finanzierungsentscheidungen auf beschriebene Programme und zunehmend auf systematische Evaluationsbefunde dazu beziehen. Die anderen Gegenstandstypen der Evaluation werden oft zur Programmevaluation additiv und unterstützend durchgeführt werden.⁵⁶

Entwicklungslinien der Programmevaluation

In einer über 50 Jahre gewachsenen Evaluationsprofession haben sich unterschiedliche methodologische Ansätze ausgeprägt, über die hier kein systematischer Überblick gegeben werden kann.⁵⁷ Statt dessen seien exemplarisch drei Definitionen aus herausragenden Standardwerken der Programmevaluation wiedergegeben und kurz kommentiert.

„Programmevaluation nutzt sozialwissenschaftliche Forschungsmethoden, um die Wirksamkeit sozialer Interventionsprogramme zu untersuchen; sie ist angepasst an deren politische und organisatorische Umfelder und angelegt, um über soziale Aktionen zu informieren, so dass soziale Bedingungen verbessert werden.“⁵⁸

Rossi und seine Kollegen arbeiten seit Jahrzehnten schwerpunktmäßig im Bereich (großer, oft auf der Ebene der Bundesstaaten angesiedelter) sozialpolitischer Interventionsprogramme (Zielgruppen z.B. Obdachlose, AIDS-Infizierte, Arbeitslose ...). Erziehung und Bildung kommen als Anwendungsbeispiele vor, dann aber meist sozialpädagogisch und kompensatorisch angelegt (Neighborhood Afterschool Program, Teen Mother Parenting Education Program). Gerade da es diesen Programmen um sozialen Ausgleich⁵⁹ geht, sind die Untersuchungsdesigns oft streng wissenschaftlich angelegt (z.B. quasi-experimentelle Designs) und versuchen primär, Wirksamkeit/Zielerreichung (effectiveness) nachzuweisen. Die Evaluatorin tendiert zu einer unabhängigen, neutralen Rolle und versucht sich durch wissenschaftliche Präzision gegen Angriffe mächtiger Interessengruppen zu schützen.

„Evaluation bedeutet die systematische Einschätzung der Durchführung und/oder der Ergebnisse (outcomes) eines Programms oder einer Policy, in Bezug auf ein Set von expliziten oder impliziten Standards, und zwar als ein Mittel, um zur Verbesserung des Programmes oder der Policy beizutragen.“⁶⁰

Carol Weiss konzipiert Evaluation als *practical craft*, als Ressource zur Optimierung von Programmen und Policies und hat ein breites Verständnis von deren Zuständigkeit: Neben die Programmresultate tritt gleichberechtigt der Programmprozess als Evaluationsgegenstand; die Bewertungskriterien sind nicht beschränkt oder von vornherein konzentriert auf die Dimension der Programmwirksamkeit und sollen expliziert sein. Bildungsprogramme innerhalb und außerhalb der Schule sind typische Gegenstände der Evaluation. Die Evaluatorin muss ihre Rolle je nach Auftrag und Machtsituation für jede Evaluation neu bestimmen. Die Präferenz für die neutrale Rolle wird bei Carol Weiss durch eine Anfangsvermutung für stärkeren Kontakt mit den Beteiligten und Betroffenen im Programm abgelöst.

„Programmevaluation ist die systematische Sammlung von Informationen über die Aktivitäten, Charakteristiken und Ergebnisse von Programmen mit dem Zweck, Beurteilungen zum Programm vorzunehmen, die Programmwirksamkeit zu verbessern und/oder Entscheidungen über künftige Programme zu begründen. Nutzenfokussierte Evaluation (in Abgrenzung zur Programmevaluation im Allgemeinen) wird durchgeführt für und mit spezifischen, primär vorgesehenen Nutzern für spezifische, vorgesehene Nutzungen.“⁶¹

Noch stärker will Michael Patton Evaluation so anlegen, dass ihre Befunde und bereits ihre Durchführung⁶² möglichst umfassend genutzt werden. Er sieht Evaluation oft eingebunden in sozialpolitische, milliardenschwere Programme zur Bekämpfung von Armut, Ungleichheit oder anderer „societal ills“, so dass Nicht-

nutzung inakzeptable Verschwendung sozialer und finanzieller Ressourcen sei. Die Gegenstandsfelder von Evaluation werden bei Patton oft aus pädagogischer Sicht erschlossen – er versteht seinen Ansatz als „*learning approach*“ – „Unterstütze die vorrangigen Beteiligten und Betroffenen darin, eine Langfristperspektive des Lernens, der Verbesserung und der Wissensnutzung zu entwickeln.“⁶³ Folglich ist zentral für das Gelingen nicht Kontakt mit Gremien oder Funktionsträgern sondern der mit den Personen, deren Selbstverpflichtung (commitment) zur Evaluation und Nutzung ihrer Befunde gefördert werden soll. Verhandlungsgeschick, Kompetenzen zur Konfliktklärung und Fähigkeit zu wertschätzendem und gleichzeitig kritischem Feedback sind wichtige Bestandteile der damit weit mehr als Beherrschung der Erhebungsmethoden umfassenden Qualifikationen einer Evaluatorin.

Diese drei Beispiele verdeutlichen einerseits die Breite der methodologischen Plattform, auf der ein Großteil der teilweise in Konkurrenz zueinander stehenden Evaluationsmodelle Platz haben.⁶⁴ Einen Kernunterschied sehe ich in der angestrebten Wirkungsweise von Evaluation und im Professionsverständnis der Evaluatorin, wobei Rossi u.a. einerseits und Patton andererseits Pole markieren, Weiss eine Mittelposition einnimmt. Soll Evaluation über (veröffentlichte) Berichte einwirken, die in einer organisationsinternen oder einer politischen Öffentlichkeit diskutiert werden und in Entscheidungen und Programmverbesserungen einmünden – Wirkmodell der experimentierenden Gesellschaft – ? Oder soll die Evaluationswirkung wesentlich in der Interaktion zwischen Evaluatorin und wichtigen Beteiligten und Betroffenen durch Verhandlungen vorbereitet werden, mit hoher Relevanz bereits des „geplanten“ Prozessnutzens – Wirkmodell der lernenden Organisation – ?⁶⁵ Ich denke eine Orientierung an beiden Modellen ist unverzichtbar.

Die nächsten Aufgaben

Soll Evaluation als Profession und Wissenschaft ihre Möglichkeiten zur Entwicklung sozialer wie pädagogischer Einrichtungen im 21. Jahrhundert entfalten, sind erhebliche Anstrengungen erforderlich. Die jungen Gesellschaften für Evaluation in der Schweiz und in Deutschland markieren den Beginn der Professionalisierung. Es gibt von diesen und auch von Regierungsbehörden teils schon beschlossene Regelwerke zur Evaluation.⁶⁶ In den gesetzlichen Grundlagen sollte Evaluation nicht nur als Pflicht- oder Wahlpflichtaufgabe genannt, sondern darüber hinaus sollten auch Anforderungen an Evaluation konkretisiert werden.⁶⁷ Noch mehr aktuelle Standardwerke⁶⁸ sollten übersetzt werden und schließlich originäre deutschsprachige Lehr- und Handbücher erscheinen. Zu den ersten berufsbegleitenden Ausbildungen für Evaluatorinnen und ersten Hochschulprofessuren mit „Evaluation“ – neben anderen – in ihrer Bezeichnung sollte eine (sinnvollerweise) als Aufbau-, Kontakt- oder Weiterbildungsstudium anzuliegende Evaluationsausbildung auf Hochschulniveau durch die Kultus- und Wissenschaftsbehörden anschubfinanziert werden. Die nächste Hauptetappe wäre dann genommen, wenn sich eine anerkannte Fachzeitschrift für Evaluation etabliert hätte. – In der Mitte des nächsten Jahrzehnts dürften diese Ziele erreicht sein, und es wird der Profession schwer fallen, sich an die Zeit zu erinnern, als diese Vorhaben noch große Herausforderungen waren.

Fußnoten

- ¹ Erschienen in: Klaus Künzel, (Hrsg) Internationales Jahrbuch für Erwachsenenbildung 27 (1999) „Evaluation der Weiterbildung“, S. 29-48.
- ² Kant (1901), S. 79; den Hinweis verdanke ich Grohmann (1997), S. 41.
- ³ Patton (1997), S. 26.
- ⁴ Längere Zitate insbesondere mit Fachtermini habe ich übersetzt, um Verständnisschwierigkeiten zu vermeiden; vgl. das von Thomas Widmer und mir übersetzte deutsch-englische Fachglossar in Joint Committee (1999), S. 225 ff.
- ⁵ Wobei Controlling (anglophon gesprochen) eine begriffliche Prägung der deutschen Betriebswirtschaftslehre ist.
- ⁶ Für die Erziehungswissenschaft richtungweisend der Sammelband von Wulf (1972), in dem „klassische“ Aufsätze amerikanischer Evaluationstheoretiker übersetzt sind. Vgl. für die Sozialwissenschaften die Übersetzung des Buches von Weiss (1974).
- ⁷ Dies gilt besonders für die Curriculumevaluation in der ersten Hälfte der 70er Jahre, als Versuch methodologischer Verdichtung. Vgl. Neul (1977).
- ⁸ In den 80er Jahren erschienen etwa zehn Werke als deutschsprachige „Platzhalter“ in der winterschlafenden Evaluationsdiskussion, z.B.: (1) Das „Handbuch zur Evaluierungsforschung“ der Politikwissenschaftler Hellstern, Wollmann (1984), von dem ein 2. Band nicht erschien; (2) die – nicht weitergeführte – Übersetzung der 3. Auflage des Standardwerkes „Evaluation – A Systematic Approach“ von Rossi und Freeman (Rossi, Freeman, Hofmann (1988)), (3) der pädagogisch-pragmatisch orientierte Sammelband von Will, Winteler, Krapp (1987) zur Evaluation in der beruflichen Aus- und Weiterbildung, (4) meine erziehungs- und sozialwissenschaftlich geprägte Darstellung der methodologischen Neuausrichtung amerikanischer Evaluation seit Mitte der 70er Jahre, Beywl (1988, 1999); sowie (5) das methodologisch eigenständige Psychologie-Lehrbuch „Evaluation“ von Wottawa, Thierau (1990), das auch in der aktualisierten Zweitausendachtzig (1998) neuere amerikanische Entwicklungen kaum verarbeitet.
- ⁹ Vgl. als Beleg für die „lateinisierte“ Aussprache: Duden (1996), S. 266. Die sprachkritische Argumentation verdanke ich dem Lateinlehrer Manfred Karbacher aus Lünen, der auf <http://schulen.hagen.de/GSGE/ew/EvalW.html> Evaluation etymologisch herleitet (Stand 5.4.1999).
- ¹⁰ „To determine or to set the value or amount of“; Webster’s (1989), S. 493.
- ¹¹ Collins (1991), S. 219.
- ¹² Joint Committee (1999), S. 25 („The systematic investigation of the worth or merit of an object“; im amerikanischen Original von 1994 auf S. 3).
- ¹³ Um der Gleichstellung der Geschlechter auch sprachlich zu ihrem Recht zu verhelfen, wird die evaluierende Person durchgängig mit der weiblichen Form bezeichnet; die anderen Akteure erscheinen in der Regel in der männlichen Form; vgl. zu dieser in der Evaluationsliteratur verbreiteten Konvention Weiss (1998), S. 7.
- ¹⁴ AEA (1995), zitiert nach der deutschen Übersetzung in: Beywl, Geiter (1997), S. 89, S. 87.
- ¹⁵ Diese unvollständige Liste ist ohne Anspruch auf Trennschärfe den „Guiding Principles for Evaluators“ entnommen; vgl. AEA (1995).
- ¹⁶ Diese Auflistung ist nicht abschließend; illustriert wird sie beispielsweise durch die Liste der knapp 30 „Topical Interest Groups“ der AEA oder durch die 97 im Jahr 1998 angekündigten Titel im „Verzeichnis lieferbarer Bücher“ (VLB).
- ¹⁷ Im Originaltext der Standards des Joint Committee (1994) kommt der Begriff evaluation research bis auf zwei Literaturangaben nicht vor; research wird in keinem Fall synonym mit evaluation genutzt; auf S. 193 wird einmal von research designs gesprochen, um das forschungsartige Vorgehen bei der Erhebung quantitativer Daten zu kennzeichnen (S. 166). Als eine Aussage für viele: „Programmevaluation unterscheidet sich fundamental von Forschung, sowohl was die Datensammlung als auch was die Bewertungsmaßstäbe betrifft“; Patton (1997), S. 24.
- ¹⁸ Bortz, Döring (1995), S. 95.
- ¹⁹ Wottawa, Thierau (1998), S. 23-53.
- ²⁰ Rossi, Freeman, Lipsey (1999), S. 398-400.
- ²¹ Weiss (1998), S. 5.
- ²² Heiner (1998), S. 25; S. 41.
- ²³ Bortz, Döring (1995), S. 29.
- ²⁴ Moser (1995), S. 94.

- ²⁵ vgl. Weiss (1998), S. 15ff.
- ²⁶ Patton (1997), S. 72.
- ²⁷ Vgl. Beywl (1991), S. 266.
- ²⁸ Ich möchte die Differenz Forschung-Evaluation konturieren, die Systemgrenzen sichtbar machen. Ich folge der Empfehlung von Sommersfeld (1998, S. 21): „(Es) müssen die Systemreferenzen durch Austausch- und Kooperationsprozesse höherer Ordnung neu gestaltet werden“, wofür gegenseitige Anerkennung und Wertschätzung unverzichtbare Voraussetzungen ist, um zu einem „heterarchischen und bilateralen oder zirkulären Wissenschaftstransfermodell“ (S. 25) zu gelangen.
- ²⁹ Vgl. auch die ausführliche Argumentation mit Literaturlauswertung bei Nadeau, Baumann, Claude (1996).
- ³⁰ Wottawa, Thierau (1998), S. 14; Kromrey (1999), S. 76 geht weiter: „Grundlagenforschung darf sich irren ... Angewandte Sozialwissenschaft darf sich Irrtümer nicht erlauben. ... Jede falsche Schlussfolgerung im Verwertungskontext, die wegen falscher Daten gezogen wird, hat Konsequenzen für einen nicht absehbaren Kreis von Betroffenen ...“ (Kromrey spricht von „Forschungsdaten“, da hier verwirrend durch „Daten“ ersetzt).
- ³¹ Die Aufbau-Logik der „Standards für die Evaluation von Programmen“ – zunächst Nützlichkeit, dann Durchführbarkeit sowie Korrektheit und schließlich Genauigkeit unterstreicht, dass die Methoden den Untersuchungszwecken folgen sollen (Joint Committee 1994, 1999); hilfreich hierzu Patton (1997), S. 239-338; siehe auch Heiner (1998).
- ³² Vgl. exemplarisch Sachverständigenrat Bildung (1999), Krüger, Olbertz (1997), Landesinstitut für Schule und Weiterbildung (1997), Merchel (1998), Bildungskommission NRW (1995).
- ³³ vgl. Beywl (1999).
- ³⁴ Vgl. das umfassende Handbuch von Friebertshäuser, Prengel (1997), in vielfacher Hinsicht Höhepunkt der Entwicklung qualitativer Forschungsmethoden in der Erziehungswissenschaft.
- ³⁵ Terhart (1997), S. 29.
- ³⁶ Dahinter stehen starke Wertkonflikte: Operationale Zielsetzung und „Messung“ (insbesondere von pädagogischen) Humandienstleistungen, stößt nicht nur auf Meßprobleme, sondern auch auf Widerstand, menschliche Interaktion über Sinnzusammenhänge, das „Lernen des autonomen Subjektes“ zu objektivieren und damit ihrer Würde zu berauben; danebensind öffentliche Finanzknappheit und erhöhtes Interesse der Steuerzahler an der Effektivität und Effizienz der von ihnen aufgebrauchten Mittel auch Ausdruck fortschreitender Individualisierung und Entsolidarisierung (vgl. Bildungskommission NRW 1995, S. 21ff). Für sozial? und bildungspolitisch engagierte Evaluatorinnen markiert dies eine große Herausforderung; vgl. Chelimsky (1997), S. 5.
- ³⁷ Auch auf der Makroebene z.B. der Landespolitik begannen sich industrielle Qualitätssteuerungsverfahren durchzusetzen, z.B. in Form des von der Firma prognos simma entwickelten Förderprogrammcontrolling für alle durch das Ministerium für Arbeit, Soziales, Stadtentwicklung, Kultur und Sport des Landes NRW aufgelegten Förderprogramme.
Vgl. für die betriebliche Weiterbildung die Übersicht in Beywl, Geiter (1997), für den Bereich der erzieherischen Hilfen EREV (1998).
- ³⁸ Vgl. als Beispiel für die wenig geklärte Stellung von Evaluation zu Qualitätsmanagement und -sicherung das Diskussionspapier „Ein neues Leitbild für das Bildungssystem“ des Sachverständigenrat Bildung (1998). Ist in den allgemeinen Empfehlungen von „Qualitätssicherung“ die Rede, die auf Evaluierung auf „Grundlage genereller Standards“ nachfolgen soll (s. 7f), so werden an anderer Stelle unter „Evaluation ... alle Verfahren der Qualitätsprüfung verstanden“ ..., deren Ziel die Akkreditierung von Bildungseinrichtungen“ ist. - Der methodologisch unscharfe Gebrauch des Konzepts „Evaluation“ ist in diesem wie vielen anderen aktuellen Diskussionspapieren, Erlassen und Gesetzeskommentaren augenfällig.
- ³⁹ Diese „einzige akkreditierte Zertifizierungsstelle in Europa, die schwerpunktmäßig im Bildungsbe- reich tätig ist“ bietet als Gründung der Deutschen Wirtschaft „Bildungseinrichtungen die Möglichkeit, ihr Qualitätsmanagementsystem nach der internationalen Normenreihe DIN EN ISO 9000 ff. zu zertifizieren.“ (<http://www.certqua.de/> vom 21.6.1999).
- ⁴⁰ Ähnlichkeiten entdeckte ich in Israel, wo Schulen vielfach ISO-zertifiziert sind und „Qualität“ im ingenieurwissenschaftlichen „Sinne“ auch als Unterrichtsgegenstand festgeschrieben ist.
- ⁴¹ Anders z.B. im Hamburgischen und im Hessischen Schulgesetz, die Evaluation von Schulprogrammen zur Pflichtaufgabe der Schulen bestimmen; ähnlich in NRW, wo dies über Erlass verpflichtend gemacht wurde; vgl. den Überblick über die Bundesländer in Diegelmann, Porzelle (1998).

- ⁴² Vgl. zum Status der Evaluation in Bildungsabteilungen der Privatwirtschaft Becker (1999) und die dort zitierte Literatur.
- ⁴³ Einige Beispiele: merit & worth als „Qualitätsdimensionen des Evaluationsgegenstandes“ (s.o., Anm. 11), „Nützlichkeit“, „Durchführbarkeit“, „Korrektheit“ und „Genauigkeit“ als „Qualitätsdimensionen für die Evaluation“ selbst und Kriterien der Meta-Evaluation (Joint Committee 1994, 1999, Widmer 1996); „konzeptionelle“ und „instrumentelle“ „Nutzenart“ (Patton 1997); „formative“ (gestaltende) und „summative“ (bilanzierende) „Leistungsart“, (Scriven 1991), die zwei Quartette der „Gegenstandsdimensionen“ „Kontext“, „Input“, „Prozeß“ und „Produkt“ (Stufflebeam 1972) bzw. „Konzept“, „Struktur“, „Prozeß“ und „Ergebnis“ (Heiner 1998); die drei „Hauptphasen“ „Gegenstandsbestimmung“, „Informationsgewinnung“, „Ergebnisvermittlung“ (Owen, Rogers 1999).
- ⁴³ Die Unterscheidung ist idealtypisch: soll z.B. ein Lehrbuch oder eine Selbstlern-CD-Rom als Produkt oder als Programm angesehen werden? Ich tendiere dazu, komplexere „Interaktionen“ mit Nutzern, Lernern auslösende Medien und Materialien, die wie ein Selbstlern-Buch oder eine Lernsoftware ein eigenständiges Handeln der Zielgruppenmitglieder erfordern, damit das gewünschte Ergebnis „Lernen“ zustandekommt, als „Programme“ zu bezeichnen. Heiner (1998), argumentiert, dass Evaluationen (für alle Typen von Gegenständen W.B.) besonders geeignet seien, das Selbststeuerungspotential von Organisationen zu erhöhen und zur Erweiterung der organisationalen Wissensbasis beizutragen, was auf die oft wünschenswerte Verknüpfung von Evaluationen mit verschiedenen Gegenstandstypen verweist.
- ⁴⁴ Mit *policies* ist im Unterschied zu *politics* (formaler Aspekt, bezogen auf das „Politiktreiben“ und „Regieren“) der inhaltliche Aspekt von Politik angesprochen, ...bei dem es stets um die Frage der Effizienz und der von ihr zu erbringenden inhaltlichen Leistungen und um die Qualität der von ihr erzeugten Güter und deren Verteilung geht. (Rohe 1986, S. 352).
- ⁴⁵ Vgl. Bussmann, Klöti, Knoepfel (1997).
- ⁴⁶ „Programm“ wird hier als evaluationsfachsprachlicher Begriff eingeführt; Verwechslungsgefahr besteht insbesondere mit dem ausgedruckten, in Heft- oder Broschürenform verteilten „Programm(überblick)“ einer Volkshochschule / einer betrieblichen Weiterbildungsabteilung. Es ist kein alternativer generischer Terminus bekannt, der die aufgeführten Ausformungsmöglichkeiten intentionaler Interventionsbündel“ in so unterschiedlichen Größenordnungen und Gegenstandsfeldern befriedigend bezeichnet; Definitionen bei Smith (1989), Rossi, Freeman, Lipsey (1999, S. 23), Bundesamt für Gesundheit (1997, S. 12).
- ⁴⁷ Hier hat sich – auch mit der Ausdifferenzierung von Evaluation aus der Forschung – als zentraler Bezugspunkt das Konzept „Programmtheorie“ (Chen 1990, Weiss 1998, S. 55ff; Rossi, Freeman, Lipsey 1999, S. 98ff, Patton 1997, S. 215ff) durchgesetzt: im Kern sind dies logisch miteinander verknüpfte, begründete Annahmen darüber, in welcher gedachten „einfachen“ (chronologisch) angeordneten Verkettung bestimmte Programminterventionen dazu beitragen, die Programmziele zu erreichen.
- ⁴⁸ Oft dann, wenn sie zeitlich befristet sind; vgl. Joint Committee (1999), S. 231.
- ⁴⁹ Versteht man unter Curriculum nicht nur Lehr-, Lernplan sondern gleichzeitig auch dessen Umsetzung, ist Curriculumevaluation eine Form der Programmevaluation.
- ⁵⁰ So die drei konstitutiven Elemente im institutionellen Organisationsbegriff von Schreyögg (1996), S. 9ff.
- ⁵¹ Auch unter der Überschrift „Schul-Selbst-Evaluation“, vgl. z.B. MSW NRW (1998); darin auch ein Beispiel der Organisationsevaluation von Volkshochschulen (Jürgen Heinen-Tenrich am Beispiel der Qualitätssicherung und Selbstevaluation im Landesverband der Volkshochschulen Niedersachsens, S. 51-55 und die dort zitierte Literatur).
- ⁵² Eine in der Fachwelt umstrittene Ausnahme davon im deutschsprachigen Raum ist das auf dem Modell des European Foundation Quality Management (EFQM) aufbauende Zertifizierungsmodell in Südtirol, das öffentliche Finanzierung freier Weiterbildungseinrichtungen in der Höhe abhängig macht von den bei Audits erreichten Punktzahlen; vgl. Frey (1997); vgl. auch Anm. 39.
- ⁵³ Beispielhaft ist das Konzept von Peterson (1995), das dazu anleiten soll, die Qualität der Lehrperson möglichst genau einzuschätzen und als Bestandteil einer partizipativen Personal- und Organisationsentwicklung anzulegen.
- ⁵⁴ Um diese Trennung zu unterstreichen, hat das Joint Committee for Educational Evaluation (1998) neben dem Set von Qualitätsstandards für Programmevaluationen ein anderes, ebenfalls als Buch herausgegebenes Set von Standards der Personenevaluation verfaßt.

- ⁵⁵ Bei der Produktevaluation (vgl. Scriven 1991) ist die Perspektive der Kunden oder Konsumenten (*customers*) zentral; bei der Programmevaluation ist dies eine nach- oder untergeordnete Perspektive gegenüber der von „Beteiligten & Betroffenen“ (*stakeholders*).
- ⁵⁶ Eine Ausnahme könnte der Schulbereich bleiben, wo die Organisationsevaluation gegenwärtig dominiert. Ich erkläre mir dies durch die besonderen organisatorischen Merkmale von Schulen: Unkündbare Lebenszeitbeamte als Lehrpersonen, damit geringe Fluktuation und kaum Gelegenheit, Programmverantwortliche – etwa durch Neueinstellung – abzulösen oder neu zu bestimmen; in der Regel nur eine Hierarchieebene mit geringer Anweisungsbefugnis und hohem Wert der Lehrfreiheit der einzelnen Lehrperson; geringe Marktmacht der Kunden (die da sind: weiterführende Schulen und Hochschulen, Arbeitgeber der Region, Eltern, nachrangig (erwachsene) Schüler). So fordert die Bildungskommission NRW (1995, S. 195): „... externe Evaluation ... soll sich auf die Schule als Ganzes, als pädagogische Handlungseinheit beziehen.“ - Es läßt sich allerdings feststellen, dass in wichtigen Randfeldern von Schule Programme bereits verstärkt eingeführt (und z.T. auch evaluiert) werden, z.B. Programme zur schulnahen Referendarausbildung, Programme im Schnittfeld von Jugendhilfe und Schule. Ob die durch Erlasse, Gesetze angeordneten „Schulprogramme“ tatsächlich Programme im hier gemeinten Sinne darstellen, welche das Handeln der Beteiligten maßgeblich konfigurieren, oder ob es eher bei (stärker als bisher operationalisierten) Konzeptionen/Programmwürfen bleibt, ist nicht abzuschätzen.
- ⁵⁷ Vgl. – in den Grundzügen noch immer zutreffend - meine Dissertation aus dem Jahr 1988 von der ein Reprint vorliegt (Beywl 1999). Patton (1997, S. 192ff) ironisiert die hauptsächlich den 80er Jahren zuzuordnende Modelldebatte in der Evaluation, indem er über 50 Ansätze aufführt und kurz definiert (von „accreditation focus“ bis „utilization focused“).
- ⁵⁸ Rossi, Freeman, Lipsey (1999), S. 20.
- ⁵⁹ Pawson, Tilley (1997, S. XIV) sind ebenfalls Anhänger dieser auf „piecemeal social reform“ angelegten Evaluationstradition.
- ⁶⁰ Weiss (1998), S. 4.
- ⁶¹ Patton (1997), S. 23.
- ⁶² Zur Relevanz des Prozeßnutzens vgl. Patton (1998).
- ⁶³ Patton (1997), S. 356.
- ⁶⁴ Kodex dieser Einigung auf einen politisch-fachlichen Minimalkonsens über Aufgabe und Gütekriterien von Evaluation in einer demokratischen Gesellschaft, über die verschiedenen „Schulen“ und disziplinären Hintergründe hinweg sind die Standards for Program Evaluation des Joint Committee (1994/1999).
- ⁶⁵ Vgl. Preskill, Torres (1999).
- ⁶⁶ Bundesamt für Gesundheit (1998); Widmer, Landert, Bachmann (1999), siehe auch <http://www.admin.ch/bj/rspm/seval/bulletin/bull11.pdf>.
- ⁶⁷ Weitgehend im § 13. (2) des Österreichischen Fachhochschulgesetz in der Fassung vom 5.5.1998, der für die Verlängerung der auf maximal fünf Jahre befristeten Anerkennung eines Fachhochschul-Studiengangs festlegt, dass ein „Evaluationsbericht“ vorzulegen ist. Dieser hat insbesondere Aufschluss zu geben über die didaktische Qualität der Lehre und die Adäquanz des Studienplans hinsichtlich der beruflichen Verwertbarkeit der Ausbildung und muß als Voraussetzung für die Verlängerung von der zuständigen Fachhochschul-Kommission abgenommen werden.
- ⁶⁸ Bisläng: Rossi, Freeman, Hofmann (1988); Nadeau, Baumann, Claude (1996); Joint Committee (1999).

Literatur

- AEA - American Evaluation Association (1995): Task Force on Guiding Principles for Evaluators, Guiding principles for evaluators, in: Shadish, W. R., Newman, D. L., Scheirer, M. A., Wye, C. (ed.): Guiding principles for evaluators. San Francisco, S. 16-29. Deutsche Übersetzung in: Beywl, W., Geiter, C. (1997), S. 85-94.
- Becker, F. (1999): Evaluation in der Personalentwicklung, Funktionen und Gestaltung, in: Schwuchow, K., Gutmann, J.: Jahrbuch Personalentwicklung und Weiterbildung 1999/2000, Neuwied, S. 198-203.
- Beywl, W., Geiter, C. (1997): Evaluation - Controlling - Qualitätsmanagement in der betrieblichen Weiterbildung. Kommentierte Auswahlbibliographie. (2. Auflage) Bielefeld.
- Beywl, W. (1991): Entwicklung und Perspektiven praxiszentrierter Evaluation, in: Sozialwissenschaften und Berufspraxis, 14 (1991) Heft 3, S. 265-279.
- Beywl, W. (1999): Nutzenfokussierte Evaluation von Humandienstleistungen - Plädoyer für eine sozialwissenschaftliche Rückbesinnung in der Qualitätsdebatte, in: Sozialwissenschaften und Berufspraxis 22 (1999), Heft 2. S. 143-156.
- Beywl, W. (1988, Reprint, Köln 1999): Zur Weiterentwicklung der Evaluationsmethodologie. Grundlegung, Konzeption und Anwendung eines Modells der responsiven Evaluation. Frankfurt
- Bildungskommission NRW (1995): Zukunft der Bildung – Schule der Zukunft. Denkschrift der Kommission beim Ministerpräsidenten des Landes Nordrhein-Westfalen, Neuwied.
- Bortz, J., Döring, N. (1995): Forschungsmethoden und Evaluation für Sozialwissenschaftler, Berlin.
- Bundesamt für Gesundheit (BAG) (Hg.) (1997): Leitfaden für die Planung von Projekt- und Programmevaluationen, Bern (Bundesamt für Gesundheit, Fachbereich Evaluation).
- Bussmann, W., Klöti, U., Knoepfel, P.: (Hg.) (1997): Einführung in die Politikevaluation, Basel.
- Chelimsky, E. (1997): The coming transformations in evaluation, in: Chelimsky, E. & Shadish, W. R. (eds.): Evaluation for the 21st century. A handbook. Thousand Oaks, S.1-26.
- Chen, H. (1990): Theory-driven evaluation, Newbury Park.
- Collins German-English, English-German Dictionary (1991). Glasgow.
- Diegelmann, E., Porzelle, K. (1998): Schulprogramm und Evaluation. Aktivitäten, Materialien und Programme der Bundesländer; in: Pädagogik Nr. 2, (1998), S. 36-40.
- EREV (Evangelischer Erziehungsverband e.V.) (Hg.) (1998): Qualitätsentwicklung und –bewertung in der Sozialen Arbeit. Systematische Ansätze und Verfahren, Hannover.
- Frey, K. (1997): Qualitätsmanagement-Systeme im Vergleich. Empfehlung für die Jugend- und Erwachsenenbildung, in: BMFSFJ (Hg.): Materialien zur Qualitätssicherung in der Kinder- und Jugendhilfe, Qs Nr. 9, (1997), S. 31-35.
- Friebertshäuser, B., Prengel, A. (Hg.) (1997): Handbuch Qualitative Forschungsmethoden in der Erziehungswissenschaft, Weinheim.
- Grohmann, R. (1997): Das Problem der Evaluation in der Sozialpädagogik. Bezugspunkte zur Weiterentwicklung der evaluationstheoretischen Reflexion, Frankfurt.
- Heiner, M. (1998): Experimentierende Evaluation und lernende Organisation, in: Heiner, M. (Hg.): Experimentierende Evaluation. Ansätze zur Entwicklung lernender Organisationen, Weinheim, S. 11-53.
- Heiner, M. (1996): Qualitätsentwicklung und Evaluation, Freiburg.
- Hellstern, G.M., Wollmann, H. (Hg.) (1984): Handbuch zur Evaluierungsforschung, Bd. 1, Opladen.

- Joint Committee on Standards for Educational Evaluation (Hg) (1999): Handbuch der Evaluationsstandards, Opladen. (Originalausgabe 1994: dass.: The program evaluation standards: How to assess evaluations of educational programs, Thousand Oaks); vgl. auch die Besprechung in diesem Band.
- Joint Committee on Standards for Educational Evaluation (1988): The personnel evaluation standards. Newbury Park.
- Kant, I. (Erstausgabe 1803): Über Pädagogik. In: Vogt, T. (Hg.): Kant, Pädagogik, Langensalza 1901, S. 69-123.
- Kromrey, H. (1999) : Von den Problemen anwendungsorientierter Sozialforschung und den Gefahren methodischer Halbbildung, in: Sozialwissenschaften und Berufspraxis, 22 (1999) Heft 1, S. 58-77.
- Krüger, H.-H., Olbertz, J.-H. (Hg.) (1997): Bildung zwischen Markt und Staat, Opladen.
- Landesinstitut für Schule und Weiterbildung (Hg.) (1997): Evaluation der Weiterbildung – Gutachten, Soest.
- Merchel, J. (Hg.) (1998): Qualität in der Jugendhilfe, Münster.
- Ministerium für Schule und Weiterbildung NRW (MSW) (Hg) (1998): Evaluation in der Schulpraxis. Beiträge zur Qualitätsentwicklung und Qualitätssicherung von Schule. Schriftenreihe des MSW Nr. 9023, Düsseldorf.
- Moser, H. (1995): Grundlagen der Praxisforschung, Freiburg.
- Nadeau, M.-A., Baumann, G., Claude, A. (1996): Evaluation. Sonderpublikation der Zentralstelle für Weiterbildung der Mittelschullehrpersonen WBZ, Zürich.
- Neul, W. (1977): Strategien der Evaluation von Curricula und der Curriculumentwicklung, Stuttgart.
- Owen, J. M., Rogers, P.J. (1999): Program evaluation: Forms and approaches, Thousand Oaks.
- Patton, M. (1998): Die Entdeckung des Prozeßnutzens. Erwünschtes und unerwünschtes Lernen durch Evaluation, in: Heiner (1998), S. 55-66.
- Patton, M. (1997): Utilization-Focused evaluation. The new century text, (3rd edition) Thousand Oaks.
- Pawson, R., Tilley, N. (1997): Realistic evaluation, Thousand Oaks.
- Peterson, K. D. (1995): Teacher evaluation, Thousand Oaks.
- Preskill h., Torres, r.T. (1999): Evaluative inquiry for learning in organizations, Thousand Oaks.
- Rohe, K.: Politikbegriffe, in: Mickel, W. (Hg.) (1998): Handlexikon zur Politikwissenschaft, München, S. 349-354.
- Rossi, P. H., Freeman, H. E., Lipsey, M. W. (1999): Evaluation. A systematic approach. 6th ed. Thousand Oaks.
- Rossi, P. H. , Freeman, H. E. , Hofmann, G. (1988): Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung, Stuttgart.
- Sachverständigenrat Bildung bei der Hans-Böckler-Stiftung (1998): Ein neues Leitbild für das Bildungssystem – Elemente einer künftigen Berufsbildung, Diskussionspapiere Nr. 2, Düsseldorf.
- Sachverständigenrat Bildung bei der Hans-Böckler-Stiftung (1999): Jugend, Bildung und Zivilgesellschaft. Anregungen zur Bildungsdiskussion, Diskussionspapiere Nr. 3, Düsseldorf.
- Schreyögg, G. (1996): Organisation. Grundlagen moderner Organisationsgestaltung, Wiesbaden.
- Scriven, M. (1991): Evaluation thesaurus, Newbury Park.

- Scriven, M. (1981): Product Evaluation; in: Smith, M.F.: New technologies for evaluation, Beverly Hills, S. 121-166.
- Smith, M.F. (1989): Evaluability assessment. A practical approach, Norwell.
- Sommersfeld, P. (1998): Erkenntnistheoretische Grundlagen der Sozialarbeitswissenschaft und Konsequenzen für die Forschung, in: Steinert, E. (Hg.) (1998): Sozialarbeitsforschung: was sie ist und leistet; eine Bestandsaufnahme, Freiburg, S. 13-31.
- Stufflebeam, D. (1997): Evaluation als Entscheidungshilfe, in: Wulf C. (Hg.) (1972): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen, München, S. 113-145.
- Terhart, E. (1997): Entwicklung und Situation des qualitativen Forschungsansatzes in der Erziehungswissenschaft, in: Friebertshäuser, Prengel, S. 27-42.
- Webster's Encyclopedic Unabridged Dictionary of the English Language (1989). New York.
- Weiss, C.H. (1998): Evaluation. Methods for studying programs and policies, 2nd ed., Prentice Hall.
- Widmer, T., Landert, C., Bachmann, N. (1999): Evaluations-Standards empfohlen von der Schweizerischen Evaluationsgesellschaft (SEVAL), Bern, Genève.
- Widmer, T. (1996): Meta-Evaluation, Bern.
- Will H., Winteler A., Krapp, A. (1987): Evaluation in der beruflichen Aus- und Weiterbildung. Konzepte und Strategien, Heidelberg.
- Wottawa, H., Thierau, H. (1990): Lehrbuch Evaluation, Bern, 2. Auflage.
- Wulf, C. (Hg.) (1972): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen, München.

Zusammenfassung

Der Beitrag führt ein in das im angelsächsischen Sprach- und Wissenschaftsraum verwurzelte Konzept der Evaluation, leistet Begriffsklärungen, unterscheidet wichtigste Typen von Evaluationsgegenständen, grenzt Evaluation zu Forschung ab und stellt prominente Modelle der Programmevaluation vor. Evaluation als im deutschen Sprachraum wenig verbreitete reife Methodologie zwischen Wissenschaft, Politik und Fachpraxis wird durch eine am exemplarischen Prinzip orientierte Darstellung erschlossen. Wer vertiefend daran interessiert ist, wie Evaluationen geplant, durchgeführt und selbst evaluiert werden, sei auf die teils kommentierte aktuelle Literatur verwiesen.

Konzepte der Qualität in der Erwachsenenbildung

Bericht

Elisabeth Wappelshammer, *Österreichisches Institut für Erwachsenenbildung*

Die österreichischen Bemühungen um Qualitätssicherung und Qualitätsmanagement von Erwachsenenbildung erfolgen auf mehreren Ebenen – in den einzelnen Bildungseinrichtungen und Verbänden, in einer ARGE von Verbänden und der Abteilung 18 des BMBWK. Das Bundesinstitut für EB in Strobl sammelt, koordiniert und organisiert solche Bemühungen. Das Seminar vom 2.-3.10.2000 mit Dr. Wolfgang Beywl setzte einen Impuls zu Abgrenzung und Klarheit verwendeter Begriffe, gab Raum für persönlichen Austausch sowie für eine Debatte zur Perspektive spezifisch österreichischer Entwicklung von Standards.

Die Gewandtheit in der Verwendung der Begrifflichkeit rund um Qualitätssicherung bewirkt sehr rasch Phänomene von Inklusion und Exklusion: Man wird rasch als Insider oder als Outsider identifiziert und gestempelt. Weiß man schon Bescheid über die „Balanced Score-Card“ oder quält man sich noch mit der Aussprache herum? Hat man schon die höheren Weihen einer ISO-Zertifizierung erlebt oder kommt man sich vor „wie ein Eingeborener, dem zum ersten Mal Missionare begegnen“ (Zitat einer Teilnehmerin) ?

Die Annäherung an Qualitätssicherung soll die Bewusstheit tagtäglicher Routine sicherstellen und deren Handhabung vereinfachen. Indem sie die Klarheit des eigenen Handelns unterstützt, kann sie auch lustvoll sein. In jedem Fall erinnert Qualitätssicherung an die Pflicht und die Lust zur dauernden Bewusstheit von Selbstverständlichem.

Dieser Bericht protokolliert zum einen die Veranstaltung in Strobl. Er bietet zum anderen eine Übersetzung in die Praxis der tagtäglichen Bildungsarbeit an. *Diese Passagen sind im Text kursiv gesetzt.*

1. Zur Klarheit der Begriffe

Es gibt eine Reihe von klassischen Definitionen, denen jeweils eine bestimmte Philosophie zugrunde liegt – sich etwa näher oder ferner von technischer Machbarkeit zu positionieren. Bei der Definition von Qualität sind prinzipiell fünf Metadimensionen im Spiel:

- ◇ Der Preis bzw. das Verhältnis von Preis und Leistung und die Frage, was preiswert ist.
- ◇ Die Mustergültigkeit, d.h. die Mindeststandards – etwa als gesetzlich vorgegebene Norm oder als Norm industrieller Fertigung oder auch als einheitliches Begriffssystem.

- ◇ Die Besttauglichkeit, d.h. Anforderungen, die der Zufriedenheit der Kunden/ Klienten/TeilnehmerInnen dienen, etwa Serviceangebote oder Anschlussfähigkeit an die Alltagssprache.
- ◇ Die Höchstgedeihlichkeit, d.h. die fachlichen Prinzipien und Standards von einem guten Produkt, einer guten Leistung.
- ◇ Das Außerordentliche: darunter fällt alles, was außerhalb des Berechenbaren und Technischen liegt – etwa Kunst oder Charisma.

Qualität ereignet sich in diesem Spannungsfeld, die Sicherung von Qualität bemüht sich um ein ausgewogenes Verhältnis – etwa von Kundenerwartungen und fachlichen Standards.

Was heißt das für die Praxis: Zunächst einmal heißt es, sich dieser Begriffe für die eigene Praxis zu bemächtigen, ihre Relativität zu erkennen und ihnen einen bestimmten und zugleich begrenzten Platz im eigenen Arbeitsfeld zuzuerkennen. Sie dienen dazu, das eigene tagtägliche Handeln übersichtlicher und kontrollierbarer zu gestalten. Die Begriffe sind ein mehr oder weniger brauchbares Handwerkszeug.

2. Zur Philosophie der Begriffe

Die Betonung einer der fünf Dimensionen verweist auf eine spezifische Philosophie. Es gibt eine Reihe von Definitionen und Instrumenten zur Steuerung von Qualität, die jeweils an einer der Dimensionen ansetzen. Das Fünfeck gibt daher wiederum ein Instrument zur Beurteilung von Definitionen und Ansätzen der Qualitätssteuerung in die Hand:

Total Quality Management (= TQM) der European Foundation for Quality Management (= EFQM) etwa setzt bei der Dimension Besttauglichkeit an, die „ISO-Familie“ bei der Mustergültigkeit. Evaluationen setzen bei der Höchstgedeihlichkeit an, Controlling und ökonomische Modelle der Neuen Steuerung bei der Dimension des Preiswerten.

Was heißt das für die Praxis: Es heißt, sich über den eigenen Begriff von Qualität klarer zu werden: An welcher Ecke des begrifflichen Spannungsfeldes setze ich selber üblicherweise an? Woher beziehe ich etwa meine Vorstellungen von Mustergültigkeit und Höchstgedeihlichkeit, was weiß ich von den Erwartungen meiner Zielgruppen? Wann erlebe ich Spannungen zwischen diesen beiden Polen? Mit wem tausche ich mich in solchen Fragen aus?

3. Zur Definition von Qualitätsmanagement

Unter dem Begriff von Qualitätsmanagement werden alle Führungsaufgaben zusammengefasst, die zur Feststellung und Verwirklichung der Qualitätspolitik sowie der Qualitätsziele und Verantwortlichkeiten dienen. Die oberste Leitung fixiert und dokumentiert ihre Zielsetzungen und Verpflichtungen zur Qualität:

Qualität

- ◇ fixiert Kernpunkte ihres Selbstverständnisses
- ◇ bestimmt die organisationsspezifische Definition von Qualität (Stichwort Leitbild)
- ◇ stellt eine Einigung über Kernprozesse der Organisation her
- ◇ beschreibt überprüfbare, in Messgrößen umgesetzte angestrebte zukünftige Zustände: Was soll genau wann erreicht werden?
- ◇ macht ein Konzept, legt eine bestimmte Struktur fest
- ◇ legt typische Prozesse und deren erwartbare Ergebnisse fest
- ◇ eruiert Erwartungen von Kundschaft/Teilnehmerschaft
- ◇ legt Ziele bezüglich der MitarbeiterInnen fest
- ◇ positioniert sich gegenüber der Gesellschaft

Das Modell des Total Quality Managements (= TQM) oder das Instrument der „Balanced Score-Card“ stellen prozessorientierte, mehr oder weniger strikte Führungsmodelle dar. Zentrales Ziel ist es, allgemeine Absichten in konkrete Messgrößen zu verwandeln, z.B. das allgemeine Ziel, neue TeilnehmerInnen zu gewinnen, in die konkrete Angabe zu verwandeln: Man nimmt sich z.B. vor, 20% mehr TeilnehmerInnen zu gewinnen.

Speziell die „Balanced Score-Card“ unterstützt streng hierarchisch und zugleich transparent angelegte Führung. „Balanced“ meint dabei den Ausgleich zwischen angegebenen Zielen und erreichten Ergebnissen. Nach vier groben Bewertungsdimensionen werden Strategien der eigenen Arbeit sowie Messdimensionen und Messkriterien des Erfolgs angegeben. Der Vorteil ist ein permanenter Überblick über Gewichtungen von Anstrengungen – eigener und jener von MitarbeiterInnen, über Ziele, Strategien, Erfolge und Defizite. Der Nachteil kann im ebenso permanenten Druck auf alle Beteiligten liegen.

Was heißt das für die Praxis: Angesichts der Tatsache, dass diese Art der Führung und Kontrolle aus der industriellen Produktion stammt, sollte man aus der Warte der in der Regel ja freiwilligen Arbeit in einem Bildungswerk sich vielleicht zunächst einmal tief durchatmend zurücklehnen, um sich dann dem Thema mit entspannter Neugier zu widmen. Es kann sinnvoll sein, einen Prozess zum Qualitätsmanagement durchzubuchstabieren, um die dabei erworbenen Fertigkeiten dann dort einzubauen, wo es auf Dauer sinnvoll erscheint. Dabei gilt es natürlich auch zu sehen, welchen Unterschied es macht, solch ein Führungsinstrument für qualitative Prozesse in der Bildungsarbeit einzusetzen. In der Praxis kann das bedeuten, in strukturierter Weise über eine Rationalisierung der eigenen Arbeit nachzudenken, etwa eine eigene „Score-Card“ zu entwerfen: Welche unterschiedlichen Ziele habe ich, nach welchen Strategien gehe ich vor, woran messe ich Erfolg oder Misserfolg?

Hier also das Beispiel einer Balanced Score-Card (Wolfgang Beywl, 2000):

Dimension	Strategie	Messdimension	Messkriterium
Finanzorientierung (45%)	Kurzfristige Ergebnissituation verbessern	Kapitalrendite	Erhöhung des Kapitalwerts auf 110%
Markt- und Kundenorientierung (20%)	Steigerung beim Neukundengewinn	Anteil der Neukunden	Quote der Neugeschäfte steigt um 11%
Prozessorientierung (10%)	Abteilungs-spezifische QM-Systeme	Anzahl genutzter Verfahrensanweisungen	mindestens 60% der Anweisungen werden in jeder Gruppe genutzt
Lern- und Entwicklungsorientierung (25%)	Steigerung problemlösender Mitarbeiter-Aktivitäten	Vorschlagswesen	Anerkannte Verbesserungsvorschläge steigen um 6%

Was heißt das für die Praxis: Zunächst ist festzulegen, wie viel Gewicht welcher Dimension verliehen werden soll, wenn man das Schema so übernimmt: den Geldfragen, den Kunden bzw. TeilnehmerInnen, den Prozessen der Angebote oder der Weiterentwicklung des eigenen Aufgabenbereichs. Eine auf die Arbeit eines Bildungswerkes zugeschnittene Balanced Score-Card könnte dann etwa folgendermaßen aussehen:

Dimension	Strategie	Messdimension	Messkriterium
Finanzorientierung (25%)	Gespräche mit einzelnen Gemeinderäten verbessern und dem	Geld- oder Sach-Sachleistungen	Erhöhung des Beitrages der Gemeinde um 20%
Markt- und Kundenorientierung (40%)	dem Bürgermeister zur Bedarfs-erhebung	aktuelle Themen im Ort	bis Ende des Jahres stehen drei neue Themen fest.
Prozessorientierung (20%)	Systematische Auflistung aller notwendigen Details zur Durchführung einer Veranstaltung zu einem der neuen Themen	Klarheit und Effizienz	die Veranstaltung kommt ohne Hektik zustande.
Lern- und Entwicklungsorientierung (15%)	Ansprechen jüngerer Leute in der Gemeinde	Effektivität der generationsübergreifenden Arbeit	Bis Ende des Jahres sind zwei jüngere Leute zur Mitarbeit gewonnen.

Die probeweise Orientierung von Bildungsarbeit an der Konfektion aktueller „QS“ und „QM“- Instrumente erinnert bald daran, wie schwierig der Vergleich

mit der Kultur industrieller Arbeitsprozesse ist. Messkriterien wirken leicht platt, die Dimensionen scheinen zentrale Fragen der Bildungsarbeit auszusparen. Wie auch immer: hier geht es darum, handwerkliche Fertigkeiten zu gewinnen, und die Balanced Score-Card kann darin unterstützen, Übung darin zu bekommen, in reflektierende Distanz zur eigenen Arbeit zu treten: Welche anderen Gewichtungsmöglichkeiten gäbe es noch, welche weiteren Strategien, Messdimensionen und Messkriterien ließen sich den vier Dimensionen noch zuordnen, und was würde das für die gesamte Entwicklung des örtlichen Bildungswerkes bedeuten?

Zugleich vertreten Nonprofit-Organisationen aber – nicht zuletzt mit Hilfe von Subventionen der öffentlichen Hand - auch öffentliche Interessen. Ihre Angebote und deren Qualität gehen daher über eine bloße Orientierung am Markt und dessen Mechanismen hinaus. So sind verschiedene gesellschaftliche Gruppen gar nicht „marktfähig“, d.h. sie sind weder in der Lage, ihre Wünsche zu artikulieren, noch sich einen entsprechenden Überblick über passende Angebote zu verschaffen. Dieser Umstand macht es sinnvoll und notwendig, Instrumente des QM auf ihre diesbezügliche Tauglichkeit zu prüfen bzw. sich entsprechende Ergänzungen einfallen zu lassen. Wie könnte das im Fall der Balanced Score-Card aussehen? Man könnte als neue Dimension das öffentliche Interesse einführen und eine neue Gewichtung vornehmen, etwa 25% seiner Aufmerksamkeit dieser Dimension widmen und bei den anderen Dimensionen entsprechend reduzieren:

Dimension	Strategie	Messdimension	Messkriterium
Öffentliches Interesse (25%)	Kooperation mit einem örtlichen Pflegeheim zur Bildung hochaltriger Pflegebedürftiger	Die Lebensfreude Pflegebedürftiger	Die Einnahme von Schmerz- und Schlafmittel bei den Teilnehmenden des Bildungsangebotes geht im Laufe eines halben Jahres deutlich zurück.

4. Die spezielle Qualität humaner Dienstleistungen

Qualitätssicherung und Qualitätsmanagement haben ihre Wurzeln in der industriellen Produktion, speziell in Auto- und Rüstungsindustrie. Ausgehend vom Bemühen, Fehlerquoten zu minimieren und besser mit anderen Firmen zu kooperieren, entwickelten sie „QS“ und „QM“ - Instrumente.

Humane oder soziale Dienstleistungen gehören in der Sprache industrieller Fertigung zu den nicht lagerfähigen Gütern. Das heißt, sie haben ein rasches Verfallsdatum und folgen dem „uno actu“ Prinzip, d.h. Produktion und Konsumation finden in einem Akt statt. Es braucht die Mitwirkung des Kunden/Klienten für das Zustandekommen der Dienstleistung. Zu beachten sind die Gefahr von schlechtem „Schnittstellen-Management“ und die generelle Störanfälligkeit. Dabei sind drei Dimensionen im Spiel:

- ◇ wie einig sind sich Fachleute über spezielle Wirkungszusammenhänge?
- ◇ wie stabil ist der zu beurteilende Gegenstand über die Zeit?

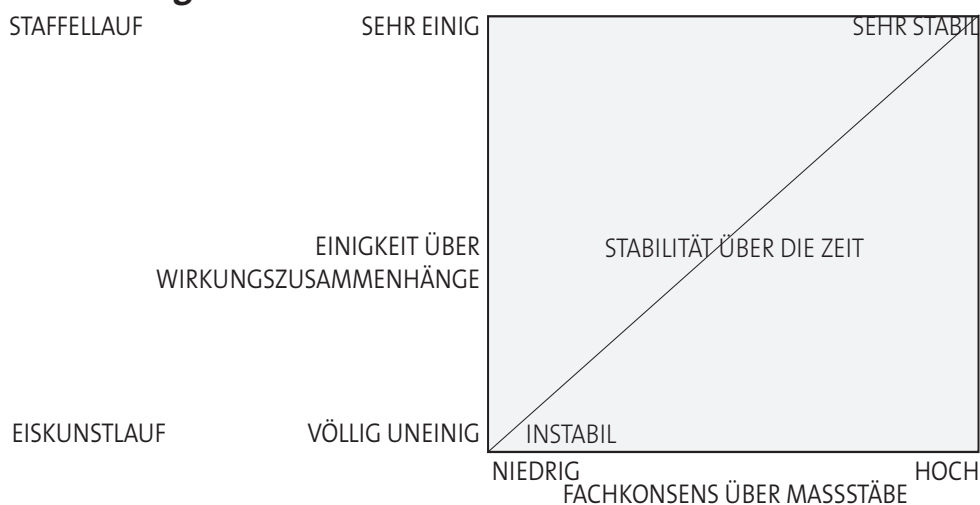
- ◇ wie hoch ist der Fachkonsens über Qualitätsmaßstäbe? Welche Werte und Grundüberzeugungen über das So-sein-sollen gibt es?

Was heißt das für die Praxis – z.B. bezogen auf die Dienstleistung des Essens in einem Bildungshaus: Der Speiseplan eines Bildungshauses mag eine über die Zeit hinweg relativ stabile Dienstleistung sein. Die Einigkeit über den Wirkungszusammenhang zwischen Essen und Bildungsprozessen ist relativ: Welchen Wirkungszusammenhang sieht man zwischen der Art der Kost und den jeweiligen Lernprozessen? Dies hängt sehr stark von den angelegten Maßstäben ab. Hier aber ist der Fachkonsens über eine gute Bildungshaus-Küche höchst instabil: Sollen die TeilnehmerInnen vor allem Wahlfreiheit haben, will man sie eher für ein bestimmtes ernährungsphysiologisches Konzept gewinnen oder legt man vor allem Wert auf eine bestimmte Art der Tischkultur als Teil der Corporate Identity des Hauses? Der Gegensatz zwischen traditioneller Hausmannskost und durchgängig vegetarischer Küche markiert nur zwei der bekanntesten Pole der Maßstäbe, die hier im Spiel sind. Eine Herausforderung für alle Beteiligten – ProduzentInnen und KoproduzentInnen der Dienstleistung Essen im Bildungshaus ist die „Schnittstelle“ des Übergangs von der Bildungsarbeit zur Essenspause. Das Essen sollte zeitgerecht zubereitet sein, die Arbeitseinheit entsprechend zeitgerecht abgeschlossen werden. Diese komplexe Situation beinhaltet verschiedenste Möglichkeiten für Pannen und Lösungen. Je nach den Grundsätzen des „So-sein-Sollens“ werden die Verantwortlichen ihre Entscheidungen treffen.

Dimensionen bei den Werten

- ◇ Grundüberzeugung über das So-sein-sollen, aufgrund derer eine Gruppe von Menschen handelt
- ◇ Wertekonflikte, z.B. zwischen Waschmitteln und Umweltverträglichkeit
- ◇ Standards von Fachleuten

Dienstleistung zwischen Staffellauf und Eiskunstlauf



- ◇ Was ist das Ziel?
- ◇ In welchen Einheiten wird die Erreichung des angegebenen Ziels gemessen?
- ◇ Welche Prüfmittel stehen zur Verfügung?
- ◇ Welche statistischen Methoden sind einsetzbar?
- ◇ Welche KundInnen werden beliefert?

In dem Bild vom Staffellauf steckt die Schnittstelle als Übergabe des „Produkts“. Im Bild des Eiskunstlaufs steckt die Frage nach einer komplexen Messung, bei der die Mitglieder der Jury selbst das Messinstrument darstellen.

Standards für Evaluation

Evaluation ist ein Verfahren, Aussagen über den Wert eines Angebots von Dienstleistungen auf der Basis systematisierter Daten zu treffen. Aspekte des Evaluierens:

- ◇ Ermittlung der Beteiligten und Betroffenen: Wo sind die „Stakeholder“, also jene Gruppen, die in irgendeiner Form Interessen vertreten, die mit dem Gegenstand der Evaluation verbunden sind?
- ◇ Umfang und Auswahl der Informationen
- ◇ Feststellung von Werten
- ◇ Klarheit des Berichts
- ◇ Verbreitung des Berichts
- ◇ Glaubwürdigkeit der Evaluation
- ◇ Wirkung der Evaluation

Was heißt das für die Praxis: Wenn man Interesse an einer Evaluation von außen hat, behält man diese Fragen selbst gut im Auge und überlegt Fragen an das Programm: Was genau soll durch die Evaluation erhoben werden, was genau ist das Ziel? Welcher Art könnten die Empfehlungen sein, die aus den Ergebnissen folgen? Welche Informationen sollten die externen EvaluatorInnen erreichen, welches Material sollten sie zur Verfügung haben?

Vier Gegenstandsarten der Evaluation eines Programms:

Beschreibung

- ◇ Welche Ziele werden verfolgt?
- ◇ Welche Erwartungen/Erfordernisse gibt es?
- ◇ Welche Bedingungen/Voraussetzungen/Instrumente sind gegeben?
- ◇ Wie verläuft die Erstellung der Leistung?
- ◇ Welche Prozesswirkungen gibt es?
- ◇ Was wird in welchem Umfang, in welcher Qualität bereitgestellt?
- ◇ Werden Ziele bei Zielgruppen erreicht?

Bewertung

- ◇ Wie gut ist das Konzept?
- ◇ Wie gut ist die Struktur?
- ◇ Wie gut ist der Prozess?
- ◇ Wie gut ist das Ergebnis?

KONZEPT	STRUKTUR	PROZESS	ERGEBNIS
Ziele und deren a) Passung auf Bedürfnisse/ Anfor- derungen b) Logische Umsetzung in Interventionen	Personelle, finanzielle, sächliche und räumliche Aus- stattung ... als Basis des ge- planten Prozesses	Geplante Ereignisse/ Aktivitäten/ Interventionen auf Ergebnissen	Leistungen/ Produkte (out puts) und Wirk- kungen (out comes)
Beschreibung			
Welche Ziele werden verfolgt? Neugeschäfte	Welche Be- dingungen/ Vor- aussetzungen/ Ins- trumente sind gegeben?	Wie verläuft die Erstellung der Leistungen?	Was wird in welchem Umfang, in welcher Qualität bereitgestellt?
Welche Erwartungen/ Erfordernisse gibt es? Verfahrensanweisungen		Welche Prozesswirkungen gibt es?	Werden Ziele bei Zielgruppen erreicht?
Bewertung			
Wie gut ist das Konzept?	Wie gut ist die Struktur?	Wie gut ist der Prozess?	Wie gut ist das Ergebnis?

Messinstrumente für die Qualität in der Erwachsenenbildung sollen ...

- ◇ immer in der Mehrzahl auftreten
- ◇ sichtbar machen, was gemessen werden soll („face validity“)
- ◇ von allen wichtigen Beteiligten verstanden werden
- ◇ allen wichtigen Beteiligtegruppen eine Chance zur Mitgestaltung des Instruments geben
- ◇ durch eine gesättigte Anzahl von Pre-Tests die Kundenmitbestimmung gewährleisten
- ◇ im Kontakt mit der Zielgruppe produziert werden, sozialen „Sinn“ machen
- ◇ Angaben über die Rahmenbedingungen enthalten:
wer hat z.B. die Instrumente gemacht?
- ◇ gesellschaftlich relevant sein

5. Qualitätsmanagement und Evaluation: Gemeinsamkeiten und Unterschiede

Beide Zugangsweisen sind von den USA nach Europa gekommen. Beide laufen Gefahr, als reine Legitimation für Subventionen missverstanden zu werden, beide werden häufig durchgeführt, aber nicht genutzt. So entstehen mehrbändige QS-Handbücher in einer Art Überentwicklung von Verfahrensanweisungen und verstauben in den Schränken von Einrichtungen. Evaluationsberichte sind oft zu dick, um gelesen zu werden, generell werden Nutzungsgewohnheiten von Adressaten zu wenig bedacht. Die kleine Schwester der Evaluation ist das Monitoring. Es bildet Grund- und Kennzahlen ab, die als Pflichtenkatalog gelten und durch episodische Evaluation ergänzt werden. Instrumente des Controllings beziehen sich speziell auf den Finanzbereich von Organisationen und Programmen. Sie richten sich auf eine sinnvolle Gliederung der Prozesse, in denen Geld ausgegeben wird sowie auf eine plausible Abbildung von Erlösen.

SEKTOREN	QM: industriell/profit	EVALUATION: non profit
Externe	Auftakt mit Externen	dauerhaft auf Externe angewiesen
Perspektive	soll fest eingebaut werden	ist stärker als ein Stück Weges fokal/ auf Ausschnitte orientiert
Periodizität	kontinuierlich	episodisch
Relation Management	Instrument der obersten Leitung	Entscheider: eher Staff oder Endnutzer
„Audit“	Rollenteilung: werden Q-Kriterien getroffen oder nicht?	EvaluatorInnen geben Endurteil ab

Funktionen von Qualitätsmanagement (= QM) für Evaluation:

- ◇ QM klärt Leitbilder, Ziele, Q-Politik etc. Das ist für Evaluation eine günstige Voraussetzung, da dies eine höhere Klarheit über Prioritäten und Ziele verleiht.
- ◇ QM klärt Zuständigkeiten. In Nonprofit-Organisationen ist oft sehr intransparent, wer entscheidet, wie protokolliert wird etc. Klarheit darüber ist eine wichtige Voraussetzung für Evaluationen.
- ◇ QM liefert Basis Zahlen für Ist-Analysen,
- ◇ Durch QM werden Schlüsselprozesse, die Organisation bestimmter Abläufe ein Stück weit geklärt und kommuniziert, man entwickelt eine gemeinsame Sprache, grenzt bestimmte Bereiche voneinander ab (unterscheidet etwa Weiterbildungsberatung von Weiterbildungsverbesserung)
- ◇ QM kann Anstoß für Evaluation geben.

Funktionen von Evaluationen (= E) für Qualitätsmanagement

- ◇ hat die Fähigkeit, Wertedimensionen zu thematisieren und Prozesse der Aushandlung in Gang zu setzen.
- ◇ verbindet die Dimensionen von **F**aktum - **I**nterpretation - **W**erte - **E**mpfehlungen: „FIWE“ (z.B.: 90% der inskribierten Studierenden kommen nicht in die Vorlesung an der Universität. Was bedeutet das aufgrund welcher Werte, was folgt daraus?)
- ◇ ist ein didaktisches Denken zu eigen. Sie denkt immer in Dimensionen von Ziel, Intervention und Ergebnis und kann Fragen und Ergebnisse fachlich vernünftig ordnen. Sie gibt Ziele an, die von didaktischen Modellen abgeleitet sind und schlägt Interventionen vor.
- ◇ Während QM beschreibt, wie Prozesse ablaufen sollen, leuchtet E die „black box“ der Kernprozesse aus: Was läuft da genauer ab zwischen Produzent und Koproduzent einer Dienstleistung?
- ◇ kann Anstoß für QM geben.

Zur „ISO-Familie“

Die Serie ISO 9000 ist ein Set allgemeiner internationaler Standards für Qualität. Mit Hilfe dieses Zugangs streben Betriebe und Organisationen an, sich selbst, ihre GeschäftspartnerInnen und KundInnen zu vergewissern, dass Produktion und/oder Dienstleistungen unter sorgfältiger Routine und permanenter Kontrolle ablaufen. Viele Betriebe und Organisationen unterwerfen sich in diesem Sinne einem Verfahren zur „Zertifizierung“: Nach einer gewissen Vorbereitungszeit, in der man den eigenen Betrieb einer systematischen Kontrolle gemäß den allgemeinen Richtlinien unterwirft und die Arbeit bis ins Detail „durchleuchtet“, wird der gewählte Prozess/das gewählte System durch einen bestellten Prüfer („Auditor“) geprüft. Der Schwerpunkt der Entwicklung von betriebseigenen Standards liegt auf der genauen Dokumentation. Zentrales Augenmerk der Kontrolle richtet sich auf die Standards und Verfahrensanweisungen der Produktion, nicht so sehr auf die Zufriedenheit der Kundschaft. Im Vergleich zum später entwickelten TQM hat dieses Verfahren daher einen eher bürokratischen Charakter. Bei Organisationen, die vor allem Dienstleistungen anbieten, sind ISO-Zertifizierungen in letzter Zeit gegenüber den kundenorientierten QM-Modellen der European Foundation for Quality Management (= EFQM) in den Hintergrund getreten.

Nun liegt als Zusammenfassung und Weiterentwicklung von ISO 9001, 9002 und 9003 ein neues Zertifizierungs-Instrument vor: 9000:2000. Dieses Konzept ist als Antwort auf die Bedürfnisse von Dienstleistungsbetrieben gedacht. Im Stobler Seminar berichtete Frau Gunhild Leye, eine Teilnehmerin vom Institut für Erziehungswissenschaft in Leipzig, von ihren Erfahrungen mit ISO-9000:2000. Im Prinzip finden sich die Inhalte von 9001 ff. wieder, doch wird mehr Gewicht auf Management-Prozesse gelegt, und auch Personalentwicklung wird thematisiert.

Was bedeutet das für die Praxis: Sprache und Vorgangsweisen rund um Zertifizierungen können Neulinge leicht verschrecken und MitarbeiterInnen zu verschiedensten Formen des passiven Widerstands verleiten (Stichwort Zeitaufzeichnungen). Zudem

sind sie nicht ganz billig. Hier gilt es auf besondere Weise, untenstehende Regel zu beherzigen. Die konkreten Unterlagen für eine Zertifizierung sind beim technisch-wissenschaftlich orientierten Beuth-Verlag erhältlich - die zentrale Bezugsquelle für Normen, Literatur aus dem Normungsbereich und technische Regeln in Deutschland. Unter der Internet-Adresse <http://www.beuth.de> kann man sich auch einen Überblick über Angebote und Preise verschaffen.

6. Goldene Regel für Prozesse von QS und QM in der Praxis:

Generell zu beherzigen ist, Instrumente für QM und QS nicht allzu umfangreich und detailliert anzulegen, damit die Ressourcen und Energien nicht schon durch die Einführung aufgezehrt werden. Das ist nämlich die beste Garantie dafür, dass Handbücher nur noch mit einem Seufzer der Überforderung ins Regal gestellt werden – und ansonsten alles beim Alten bleibt. Hier gilt es, die Machbarkeit im Auge zu behalten und sich selbst und die MitarbeiterInnen nicht allzu stark unter Druck zu setzen. Es geht darum, sich lernend auf den Weg zu machen – bei Prozessen „wo's klemmt“ (Zitat Wolfgang Beywl).

Literatur

Arnold, Rolf: Qualitätssicherung in der Erwachsenenbildung.- Verlag Leske und Budrich: Opladen 1997

Badelt, Christoph(Hrsg.): Handbuch der non profit Organisation. Strukturen und Management.- Schäffer-Poeschel Verlag: Stuttgart 1997

Decker, Franz: Das große Handbuch. Management für soziale Institutionen.- mi Verlag (moderne Industrie): Landsberg/Lech 1997

Egger, Rudolf, Wolfgang Grilz (Hrsg.): Bildung an der Grenze. Erwachsenenbildung im Übergang.- Leykam Verlag: Graz 1999

Evers, Adalbert, Riitta Haverinen, Kai Leichsenring, Gerald Wistow (Eds.): Developing Quality in Personal Social Services. Concepts, Cases and Comments (= Public Policy and Social Welfare, Volume 22). – Published by Ashgate: Aldershot 1997

Heiner, Maja: Qualitätsentwicklung durch Evaluation.- Lambertus Verlag: Freiburg i.B. 1996

Heiner, Maja (Hrsg.): Experimentierende Evaluation. Ansätze zur Entwicklung lernender Organisationen. – Juventa Verlag: Weinheim und München 1998

Kalny, Edith, Rudolf Pusterhofer: ISO Management. Chancen und Risiken bei der Zertifizierung, ISO 9001:2000.- Linde Verlag: Wien 1999

Sanders, James R. (Hrsg.): Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“. - 2. Auflage 2000. Aus dem Amerikanischen übersetzt, bearbeitet und ergänzt von Wolfgang Beywl und Thomas Widmer, Verlag Leske und Budrich: Leverkusen, 2. Auflage 2000

(Bei Edgar Forster und Daniela Ingruber möchte ich mich herzlich für Korrekturen und kritische Anmerkungen bedanken.)

Dr. Elisabeth Omran-Wappelshammer ist Geschäftsführerin des Österreichischen Instituts für Erwachsenenbildung, St. Pölten, <http://www.oieb.at>

