

Jan Hense und Heinz Mandl

Selbstevaluation – Ein Ansatz zur Qualitätsverbesserung
pädagogischer Praxis und seine Umsetzung am
Beispiel des Modellversuchprogramms SEMIK

Dezember 2003



Hense, J. & Mandl, H. (2003). Selbstevaluation – ein Ansatz zur Qualitätsverbesserung pädagogischer Praxis und seine Umsetzung am Beispiel des Modellversuchsprogramms SEMIK (Forschungsbericht Nr. 162). München: Ludwig-Maximilians-Universität, Department Psychologie, Institut für Pädagogische Psychologie.

Forschungsbericht Nr. 162, Dezember 2003

Ludwig-Maximilians-Universität München
Department Psychologie
Institut für Pädagogische Psychologie
Lehrstuhl Prof. Dr. Heinz Mandl
Leopoldstraße 13, 80802 München
Telefon: (089) 2180-5146 – Fax: (089) 2180-5002
<http://lsmndl.emp.paed.uni-muenchen.de/>
email: mandl@edupsy.uni-muenchen.de, hense@emp.paed.uni-muenchen.de

Selbstevaluation – Ein Ansatz zur Qualitätsverbesserung
pädagogischer Praxis und seine Umsetzung am
Beispiel des Modellversuchprogramms SEMIK

Jan Hense und Heinz Mandl

Forschungsbericht Nr. 162

Dezember 2003

Ludwig-Maximilians-Universität München
Department Psychologie
Institut für Pädagogische Psychologie
Lehrstuhl Prof. Dr. Heinz Mandl

Zusammenfassung

Als Selbstevaluationen werden Evaluationsverfahren bezeichnet, bei denen die Praxis gestaltenden Fachleute „Eigentümer des Prozesses“ sind. Das bedeutet, dass sie über Durchführung, Ziele und Vorgehen bei der Evaluation sowie über die Verwendung von Ergebnissen selbst entscheiden bzw. einen maßgeblichen Einfluss bei diesen Entscheidungen ausüben. In diesem Beitrag wird die Entwicklung von Selbstevaluation als eigenständigem Ansatz vor dem Hintergrund der jüngeren Evaluationsgeschichte nachgezeichnet. Ein Vergleich möglicher Funktionen von Fremd- und Selbstevaluation sowie vier Fallbeispiele aus Sozialpädagogik, Schule, Hochschule und Aus- und Weiterbildung dienen als Ausgangspunkte einer Diskussion der Leistungsfähigkeit des Ansatzes. Ein konkretes Umsetzungsbeispiel von Selbstevaluation und erste Ergebnisse werden anhand des bundesweiten Modellversuchsprogramms SEMIK, welches die Implementation neuer Medien in Schulen zum Inhalt hatte, ausführlich dargestellt. Abschließend werden Desiderata für die weitere Entwicklung in Theorie und Praxis formuliert.

Schlüsselwörter: Selbstevaluation, Evaluation, Evaluationsforschung, Qualitätssicherung im Bildungswesen, Modellversuch, Modellversuchsprogramm

Abstract

Self-evaluations are evaluations where the practitioners who are responsible for the evaluation object are “owners of the process”. They can decide on their own on initiation, goals, and procedure of an evaluation, and on the use of evaluation results, or at least they have considerable influence on these decisions. In this report we trace back the development of self-evaluation as an independent approach in recent evaluation history. To discuss its potentials and limits, we compare possible functions of conventional evaluation and self-evaluation, and portray four case examples from the fields of social work, school, university, and further education. As a concrete example, concept and realization of the self-evaluation in SEMIK, a national German program to foster implementation of new media in schools, are described together with early results. Concluding, we outline some desirable developments for future theory and practice.

Keywords: self-evaluation, evaluation, evaluation research, quality assurance in education, project, program

Inhalt

1. Evaluation und der Qualitätsdiskurs im Bildungswesen	4
2. Entwicklungslinien der Evaluation	5
2.1 Definition und die jüngere Entwicklung	5
2.2 Funktionen von Evaluation	7
2.3 Probleme und Grenzen traditioneller Evaluationsansätze	9
3. Selbstevaluation	11
3.1 Definition und Bestimmungsmerkmale	11
3.2 Funktionen von Selbstevaluation	12
3.3 Beispiele aus verschiedenen pädagogischen Praxisfeldern	14
4. Selbstevaluation im BLK-Programm SEMIK.....	22
4.1 Konzept der Selbstevaluation	22
4.2 Umsetzung	25
4.3 Ergebnisse der SEMIK-Selbstevaluation	31
5. Ausblick.....	32

SELBSTEVALUATION – EIN ANSATZ ZUR QUALITÄTSVERBESSERUNG PÄDAGOGISCHER PRAXIS UND SEINE UMSETZUNG AM BEISPIEL DES MODELLVERSUCHPROGRAMMS SEMIK

1. Evaluation und der Qualitätsdiskurs im Bildungswesen

Im Bildungswesen spielt die Qualitätsdebatte vor dem Hintergrund knapper werdender Ressourcen der öffentlichen und privatwirtschaftlichen Hand eine wichtige Rolle. In der bildungspolitischen Diskussion der letzten Jahre ist der Legitimationsdruck, der insbesondere auf öffentlich finanzierten Bildungsangeboten lastet, erheblich gestiegen. Gefragt wird dabei in erster Linie, ob und in welchem Maße unser Bildungssystem die Leistungen hervorbringt, die von ihm erwartet werden.

In diesem Zusammenhang hat die Evaluation von Bildungsmaßnahmen zunehmend an Bedeutung gewonnen. Historisch gesehen (vgl. Madaus & Stufflebeam, 2000) handelte es sich dabei in der Regel um externe bzw. Fremdevaluationen, bei denen der Evaluationsgegenstand von einem eigens dafür beauftragten Evaluator oder Evaluationsteam begutachtet wird und die Personen, deren Handeln evaluiert wird, geringen Einfluss auf den Evaluationsprozess ausüben. Daneben gewinnen aber in den letzten Jahren Evaluationsansätze an Gewicht, die den Praktikern einen aktiveren Part bei der Evaluation einräumen wollen und die versuchen, die Sichtweisen aller Beteiligten stärker zu berücksichtigen (vgl. Madaus & Kellaghan, 2000).

Im deutschsprachigen Raum hat sich die Bezeichnung *Selbstevaluation* für Evaluationsansätze etabliert, die sich von herkömmlichen Modellen der Fremdevaluation durch externe Personen oder Institutionen abgrenzen wollen: Bestimmendes Merkmal ist, dass die Personen, deren Tätigkeit evaluiert wird, selbst den Evaluationsprozess durchführen oder zumindest wesentliche Teile des Prozesses in der eigenen Hand haben. Wie weit diese Rollenidentität zwischen Evaluierenden und Evaluierten gehen muss, damit von einer Selbstevaluation gesprochen werden kann, wird später noch zu klären sein.

Dieser Beitrag geht aus von einer Darstellung allgemeiner Ansätze und Entwicklungslinien im Feld Evaluation (Abschnitt 2). Vor diesem Hintergrund soll auf die Entwicklung von Selbstevaluation als eigenständigem Ansatz zur Evaluation in pädagogischen Praxisfeldern eingegangen und wesentliche Merkmale des Ansatzes anhand verschiedener Praxisbeispiele verdeutlicht werden (Abschnitt 3). In Abschnitt 4 wird dann die Umsetzung der Selbstevaluation im

Modellversuchsprogramm SEMIK (Systematische Einbeziehung von Medien, Informations- und Kommunikationstechnologien in Lehr- und Lernprozesse) dargestellt und über deren Ergebnisse berichtet. Abschließend werden Konsequenzen für die weitere Forschung und Praxis diskutiert.

2. Entwicklungslinien der Evaluation

Die Entwicklung von Selbstevaluation als eigenständiger Ansatz innerhalb des Praxisfelds Evaluation ist nur vor dem Hintergrund allgemeiner Trends im Evaluationsfeld der vergangenen Jahre zu verstehen. Daher ist es notwendig, zunächst auf wesentliche Bestimmungsmerkmale von Evaluation einzugehen. Dazu gehören neben einer definitorischen Abgrenzung die Funktionen, die Evaluation erfüllen kann, einige wichtige theoretische Entwicklungslinien und Probleme und Grenzen, welche sich in der Praxis gezeigt haben und zur Entwicklung ergänzender und alternativer Herangehensweisen geführt haben.

2.1 Definition und die jüngere Entwicklung

Evaluation kann definiert werden als die systematische Untersuchung von Nutzen oder Wert einer Sache („the systematic investigation of the worth or merit of an object“, Joint Committee on Standards for Education, 1994, S. 3). Drei wesentliche Elemente machen diese Definition aus: Systematische Untersuchung, Nutzen bzw. Wert, sowie der Evaluationsgegenstand („Sache“ bzw. „object“).

Obwohl das Element *systematische Untersuchung* in der Definition des Joint Committee nicht näher bestimmt wird, besteht doch weitgehend Konsens darüber, dass es sich dabei um ein wissenschaftlich begründbares Vorgehen handeln muss und wissenschaftliche Methoden zum Einsatz kommen müssen, um zu einem Urteil über Wert bzw. Nutzen einer Sache zu kommen. In Anlehnung an den von Suchman (1967) geprägten Begriff „evaluative research“ wurde vorgeschlagen, explizit „Evaluationsforschung“ zu verwenden, um wissenschaftsgestützte Evaluation vom alltäglichen Begriffsverständnis abzugrenzen. Diese Trennung konnte sich jedoch nicht nachhaltig durchsetzen, vermutlich schon deswegen, weil sie dem intuitiven Verständnis von Evaluationsforschung als Forschung über Evaluation (analog zu „Sozialforschung“, „Genforschung“ und ähnlichen Komposita) zuwiderläuft.

Die Differenzierung von *Wert und Nutzen* in der Definition des Joint Committee ist folgendermaßen zu verstehen: Wert meint den „inneren“ Wert des untersuchten Gegenstands unabhängig von einem bestimmten Verwendungskontext. Nutzen oder Nützlichkeit einer Sache ergeben sich dagegen erst

angesichts konkreter Anforderungen und sind damit immer vom Verwendungskontext abhängig (vgl. Wittmann, 1985). So kann beispielsweise der Wert eines bestimmten Web Based Trainings (WBT) aufgrund einer Qualitätsanalyse allgemein als hoch beurteilt werden; seine Nützlichkeit kann sich allerdings in einem bestimmten Anwendungskontext als niedrig erweisen, wenn etwa die technischen Anforderungen für das intendierte Einsatzfeld zu hoch sind oder die intendierte Zielgruppe nicht über das notwendige Vorwissen verfügt.

Gegenstand einer Evaluation („object“) können je nach Kontext und Praxisfeld ganz unterschiedliche Sachverhalte sein. Scriven (1999) nennt Programme, Personal, Leistungen, politische Strategien, Vorschläge und Produkte als „Big Six“ möglicher Evaluationsfelder. Für den Bildungsbereich ist vor allem der Bereich Programmevaluation relevant, da es hier meist um die Beurteilung von innovativen oder etablierten Bildungsmaßnahmen geht. Jedoch spielen auch die Evaluation von Personal (etwa bei der Lehrevaluation), von Leistungen (etwa bei der Evaluation pädagogischer Institutionen) und von Produkten (etwa bei Lehrmitteln oder Lernumgebungen) hier eine Rolle. Die Darstellungen im Rahmen dieses Beitrags konzentrieren sich auf den Bereich der Programmevaluation, wobei der Begriff „Programm“ im Deutschen synonym mit „Maßnahme“ und „Intervention“ gesetzt werden kann (vgl. Beywl & Taut, 2000).

Eine Praxis der Evaluation lässt sich bis in die Renaissance und die ihr zugrunde liegende Denktradition sogar bis in die Antike zurückverfolgen (vgl. Cronbach et al., 1980; Henninger, 2000). Dennoch werden üblicherweise als Ausgangspunkt der gegenwärtigen Entwicklung die großen Innovations- und Reformvorhaben der 1960er Jahre angesehen (vgl. Stufflebeam, 2001). Sie schufen einen allgemein erhöhten Bedarf nach Überprüfung und objektiverer Steuerung gesellschaftlicher Entscheidungsprozesse, der durch Evaluationen gedeckt werden sollte.

Zu Beginn dieser Entwicklung wurde Evaluation im Wesentlichen als bloße angewandte Sozialwissenschaft betrachtet und betrieben. Aus deren Forschungslogik übernimmt man die obligatorische Trennung zwischen Evaluierenden und dem Evaluationsgegenstand bzw. den Evaluierten, so dass bis zum Aufkommen der Selbstevaluation jede Evaluation immer automatisch eine Fremdevaluation ist. In (quasi-)experimentellen Designs sollte die Wirkung von Programmen und Maßnahmen überprüft werden, um politisch-administrative Entscheidungsprozesse zu objektivieren. In Campbells (1969) Paradigma von „reforms as experiments“ hat sich dieser Gedanke am deutlichsten manifestiert.

Erst allmählich kam es zur Ausdifferenzierung eines eigenständigen Profils als Disziplin und Profession (vgl. Altrichter, 1999a; Shadish, Cook & Leviton, 1991). Sie schlug sich etwa in der Gründung von Berufsverbänden, der Etablierung

eigenständiger Ausbildungsprogramme, der Entwicklung evaluationsspezifischer Forschungsmethoden und der Ausarbeitung von normativen Standards für die Evaluationspraxis nieder. Begleitend stellte sich die Erkenntnis ein, dass dem Praxisfeld Evaluation spezifische Anforderungen, Methoden und auch Probleme zu Eigen sind, die sich aus ihrem Verwendungskontext in der Praxis ergeben (vgl. Wittmann, 1985) und die weit über das Feld der angewandten Sozialforschung hinausgehen.

Die Entwicklungen im angloamerikanischen Raum sind dabei mit Verzögerung ähnlich im deutschen Sprachraum zu beobachten (vgl. Stangel-Meseke & Wottawa, 1993; Will & Krapp, 1983), auch wenn ihnen hier jeweils spezifische Begründungszusammenhänge und Bedürfnislagen zugrunde lagen (vgl. etwa Frey, 1975; Wulf, 1975).

2.2 Funktionen von Evaluation

Was kann Evaluation leisten? Die Definition des Joint Committee bleibt gegenüber möglichen Funktionen, die Evaluation erfüllen kann, neutral. Eine intensive Auseinandersetzung mit der Frage nach Funktionen von Evaluation findet im Rahmen der Diskussion des Evaluationsnutzens (utilization) statt (vgl. Alkin, 1985; Alkin, Daillak & White, 1979; Legge, 1984; Patton, 1997). Zwei wichtige Dimensionen werden dort in Bezug auf den Nutzen von Evaluationen unterschieden:

1. Evaluation bezeichnet sowohl den Prozess der systematischen Untersuchung einer Sache, als auch das Produkt dieses Prozesses, nämlich konkrete Evaluationsergebnisse, die meist in Form von Berichten niedergelegt werden. Evaluationsnutzen kann nicht nur aus den Ergebnissen, sondern gleichermaßen als Folge des Evaluationsprozesses entstehen (vgl. Beywl, 1991; Patton, 1998; Will & Blickhan, 1987). Dabei ist festzuhalten, dass als Produkte auch Zwischen- und nicht nur Endergebnisse einer Evaluation anzusehen sind.
2. Der Nutzen von Evaluation kann sich auf zwei Ebenen niederschlagen (vgl. Alkin, 1985). Auf der Handlungsebene kann sie konkrete Veränderungen und Konsequenzen auslösen (action use). Auf der Wissens-ebene kann sie vertiefte Informationen über einen Sachverhalt oder eine Maßnahme bereitstellen und damit Überzeugungen beeinflussen (conceptual use).

Verbindet man die zwei Dimensionen Prozess/Produkt und Handlungsebene/Wissensebene schematisch (s. Abbildung 1), so lassen sich die vier wichtigsten Funktionen von Evaluation (vgl. etwa Alkin, 1985; Chelimsky, 1997; Legge, 1984; Will, Winteler & Krapp, 1987; Wulf, 1975) konzeptionell verorten:

	Prozessnutzen	Produktnutzen
Handlungsebene	1. Entwicklungsfunktion	2. Entscheidungsfunktion
Wissensebene	3. Kontrollfunktion	4. Lernfunktion

Abbildung 1: Vier Hauptfunktionen von Evaluation

1. **Entwicklungsfunktion:** Eine häufige Erwartung an Evaluationen ist, Erkenntnisse über mögliche Verbesserungen und Optimierungen des Evaluationsgegenstands zur Verfügung zu stellen. Leisten kann Evaluation dies, indem sie im Prozess Schwachstellen einer Maßnahme oder Ist-Soll-Diskrepanzen feststellt. Veränderungen können bereits zu Beginn des Evaluationsprozesses ausgelöst werden (vgl. Alkin, 1985), etwa wenn der Erstkontakt mit dem Evaluationsteam dazu führt, dass die Ziele einer Maßnahme deutlicher formuliert werden, um ihre Operationalisierung zu ermöglichen.
2. **Entscheidungsfunktion:** Evaluationsergebnisse sollen hier das notwendige Wissen zur Verfügung stellen, um Entscheidungen zu treffen, die sich auf den evaluierten Gegenstand beziehen. Gewöhnlich handelt es sich dabei um Entscheidungen über die Einstellung, Fortführung oder Implementation von evaluierten Maßnahmen wie beispielsweise bei Modellversuchen. Evaluationstheoretisch ist die Entscheidungsfunktion die ursprünglichste Funktion von Evaluation (vgl. etwa Campbell, 1969; Weiss, 1972).
3. **Kontrollfunktion:** Die Einrichtung einer Evaluation kann dazu dienen, im Prozess Wissen darüber bereitzustellen, ob sich die für eine bestimmte Maßnahme investierten Mittel bezahlt machen und im Sinne des Auftraggebers der Maßnahme verwendet werden. Aus Sicht der Durchführenden der Maßnahme handelt es sich also um die Kontrolle ihrer Arbeit.
4. **Lernfunktion:** Hier konzentriert sich der Anspruch an die Evaluation darauf, mehr über Wirksamkeit und Funktionsweise einer untersuchten Maßnahme zu erfahren. Dieses Wissen kann dann unterschiedlich genutzt werden, etwa um für die Maßnahme zu werben (PR-Funktion), oder im wissenschaftlichen Bereich, um es in die weitere Forschungsarbeit einfließen zu lassen (Wissenschaftsfunktion).

Grundsätzlich schließen sich die genannten Funktionen nicht gegenseitig aus, das heißt, in einem Evaluationsprojekt können durchaus mehrere oder gar alle dieser Funktionen verfolgt werden. Zwei Aspekte spielen dabei aber eine wichtige Rolle:

Erstens haben die verschiedenen Funktionen eine mehr oder weniger enge Affinität zu unterschiedlichen methodischen Herangehensweisen. Chelimsky (1997) spricht aufgrund dieser Verschränkung von Funktion und Methode von verschiedenen Paradigmen bzw. conceptual frameworks der Evaluation. So legt etwa das Entwicklungsparadigma ein eher qualitativ orientiertes Vorgehen nahe, bei dem das Evaluationsteam mehr beratend als kontrollierend auftritt; das Kontrollparadigma dagegen legt ein stärker quantitativ und output-orientiertes Vorgehen nahe, bei dem auf eine strenge Trennung zwischen Evaluierten und Evaluierenden geachtet wird. Sollen mehrere Evaluationsfunktionen in einem Evaluationsprojekt kombiniert werden, müssen diese methodischen Affinitäten entsprechend berücksichtigt werden.

Der zweite Aspekt, der bei der Kombination mehrerer Funktionen beachtet werden muss, ist, dass häufig die unterschiedlichen Beteiligengruppen auch unterschiedliche Interessensperspektiven in einem Evaluationsprojekt verfolgen (vgl. Legge, 1984; Wittmann, 1985; Wottawa, 2001). So wird der Auftraggeber einer Maßnahme meist vorrangig an der Kontrolle von Ergebnissen interessiert sein, während für Durchführende (und Zielgruppe) der Maßnahme eher die Entwicklungsfunktion von Evaluation wichtig ist. Evaluationsteams ist oft stark an der Lernfunktion gelegen und für politische oder sonstige Entscheidungsträger dürfte in der Regel die Entscheidungsfunktion die wichtigste sein. Auch hier gilt: Die verschiedenen Funktionen schließen sich nicht gegenseitig aus und unterschiedliche Interessen können durchaus gleichzeitig berücksichtigt werden. Nur müssen diese bei Planung und Durchführung eines Evaluationsprojekts angemessen berücksichtigt werden.

Neben den vier genannten „regulären“ Funktionen lassen sich in der Praxis noch verschiedene sogenannte „Pseudo-Funktionen“ von Evaluation beobachten. Diese liegen etwa vor, wenn Evaluationen beauftragt werden, um eine Entscheidung hinauszuzögern (vgl. Suchman, 1970), um eine bereits gefallene Entscheidung nachträglich zu legitimieren (vgl. Legge, 1984) oder wenn sie nur pro forma und dem Augenschein nach durchgeführt werden („symbolic use“, vgl. Alkin, 1985).

2.3 Probleme und Grenzen traditioneller Evaluationsansätze

Die optimistische Sichtweise, welche die frühen Jahre der Evaluationspraxis dominierte und die in Anlehnung an Kraus (1991) die „positivistische“ genannt werden kann, geriet im Laufe der 70er Jahre in die Kritik. Es hatte sich gezeigt, dass die in Evaluationen gesetzten Erwartungen häufig nicht erfüllt werden konnten: „Everybody seems to think that evaluation is not rendering the service it should“ merken Cronbach et al. (1980, S. 44), ursprünglich selbst Vertreter

einer objektivistischen Herangehensweise, resigniert an. Legge (1984) diagnostizierte zu Beginn der 80er Jahre eine dreifache Krise der Evaluation:

1. Evaluationsergebnisse würden zu selten in der Praxis rezipiert und genutzt (Nutzungskrise).
2. Das traditionell positivistische Forschungsparadigma und (quasi-)experimentelle Designs seien in der Praxis nicht wirklich umsetzbar oder erbrächten nur wenig praxisrelevante Informationen (Methodenkrise).
3. Bei vielen Evaluationsprojekten bewegten sich die Evaluierenden in einem Spannungsfeld von divergierenden Interessen und sähen sich schnell dem Vorwurf ausgesetzt, einseitig die Interessen der Auftraggeber (der Maßnahme und/oder der Evaluation) zu transportieren (Wertekrise).

Neben eher abwehrenden Reaktionen, welche im Wesentlichen das positivistische Paradigma aufrechterhalten wollten (z. B. Sechrest & Figueredo, 1993), lassen sich zwei Grundtendenzen infolge dieser Kritik erkennen. Einerseits wurde versucht, die existierende Praxis zu reformieren und so die genannten Krisen zu lösen. Als Resultate dieser Bemühungen können die Formulierung von normativen Standards für die Evaluationspraxis (Deutsche Gesellschaft für Evaluation, 2002; ERS Standard Committee, 1982; Joint committee on standards for educational evaluation, 1988, 1994, 2003) und die verstärkte Forderung nach Meta-Evaluationen zur Qualitätskontrolle von Evaluation (z. B. Cooky, 1999) gesehen werden.

Die zweite Grundtendenz zielte auf die konzeptionelle und methodische Neuorientierung der existierenden Praxis (vgl. Beywl, 1991). Sie äußerte sich in Forderungen nach partizipativer Einbeziehung aller Beteiligengruppen in den Evaluationsprozess, nach multiperspektivischer Sichtweise, nach einem Primat des Evaluationsnutzens vor methodischen Gütekriterien und nach einer Aufwertung qualitativer und alternativer Forschungsmethoden. In der Folge wurde eine Vielzahl unterschiedlicher Evaluationsmodelle entwickelt, welche diese Forderungen oder zumindest Teile davon umsetzen wollten. Dazu gehören etwa Participatory Evaluation (Cousins & Whitmore, 1998; Whitmore, 1998), Responsive Evaluation (Beywl, 1991; Stake, 1975), 4th Generation Evaluation (Guba & Lincoln, 1989), Empowerment Evaluation (Fetterman, 1996, 2000) oder Utilization-focused Evaluation (Patton, 1997). Gleichzeitig sind auch Integrationsversuche zu erkennen, welche sich bemühen, beide Perspektiven zusammenzuführen (vgl. Shadish, Cook & Leviton, 1991).

3. Selbstevaluation

Die im vorigen Abschnitt dargestellten Entwicklungen im Praxisfeld Evaluation liefern den Hintergrund, vor dem das Aufkommen der Selbstevaluation zu sehen ist. Im deutschsprachigen Raum ist die Debatte um den richtigen Weg Evaluation zu betreiben weniger terminologisch vielfältig geführt worden, auch wenn häufig ein Rückbezug auf die genannten angloamerikanischen Entwicklungen stattfindet. Selbstevaluation hat sich hier als wichtigstes Konzept herauskristallisiert, das (je nach Sichtweise) eine Ergänzung oder Alternative zu herkömmlichen Evaluationsmodellen sein soll.

3.1 Definition und Bestimmungsmerkmale

Als Terminus wurde Selbstevaluation zuerst vom Verwaltungswissenschaftler Wildavsky (1972) geprägt. Sein Konzept der „self-evaluating organization“ beinhaltet die Forderung nach einem kontinuierlichem Wandel von Institutionen, der durch Evaluation permanent mit wissenschaftlichen Methoden auf Erfolg hin überprüft werden soll. Damit kann es als Versuch gesehen werden, Campbells (1969) Idee der „experimenting society“ auf den Bereich der Organisationsentwicklung zu übertragen. Mit dem Element der kontinuierlichen Selbstvergewisserung hat Wildavsky ein wichtiges Element heutiger Selbstevaluationskonzepte vorweggenommen.

Im deutschen Sprachraum ist der Begriff Selbstevaluation von Heiner (1988, 1994, 1996) erneut aufgegriffen und für die Bereiche Sozialpädagogik und soziale Arbeit zu einem elaborierten Konzept ausgebaut worden. Teils unter Rezeption dieser Entwicklungen, teils aber auch unabhängig davon hat das Konzept auch in den Praxisfeldern Schule, Universität und Weiterbildung jeweils auf Basis der spezifischen Traditionen und Ausgangslagen Eingang gefunden. Auf diese fachspezifischen Entwicklungen werden wir weiter unten im Rahmen der Darstellung von Fallbeispielen genauer eingehen.

In einer Vorversion für die Standards der Selbstevaluation, welche derzeit von der Deutschen Gesellschaft für Evaluation (DeGEval) diskutiert werden, wird folgende Definition von Selbstevaluation gegeben: „Als Selbstevaluationen werden Verfahren bezeichnet, bei denen die Praxis gestaltenden Fachleute identisch sind mit den Evaluatorinnen bzw. Evaluatoren. D.h. die Akteure überprüfen ihre eigene Tätigkeit und deren Konsequenzen“ (Müller-Kohlenberg & Beywl, 2003, S. 1).

Die Rollenidentität, die hier *expressis verbis* zwischen Evaluatoren und Evaluierten verlangt wird, ist insofern zu relativieren als durchaus nicht verlangt wird, dass sämtliche Tätigkeiten, die zum Evaluationsprozess gehören, durch die Praktiker vollzogen werden. Vielmehr machen die sogenannten „Muss-

Standards“ der Selbstevaluation (vgl. Müller-Kohlenberg & Beywl, 2003) deutlich, dass bestimmendes Merkmal für eine Selbstevaluation ist, wer Eigentümer des Evaluationsprozesses ist (engl. *owner of the process*). Selbstevaluation liegt demnach dann vor, wenn die Praktiker diejenigen sind, welche die maßgebliche Verantwortung für den Prozess tragen, welche Konsequenzen beschließen und welche über die Verbreitung von Ergebnissen entscheiden; Fremdevaluation liegt vor, wenn diese Kompetenzen bei anderen Beteiligten-Gruppen wie Auftraggeber oder Evaluationsteam liegen.

Häufig werden die Begriffe *Selbstevaluation* und *interne Evaluation* synonym verwendet. Eine konzeptionelle Trennung nimmt König (1998) vor. Demnach ist eine Evaluation intern, wenn die bewertenden Akteure aus derselben Organisation stammen, die auch den Evaluationsgegenstand zu verantworten hat. Selbstevaluation liegt vor, wenn die für den Evaluationsgegenstand Verantwortlichen selbst die Evaluation durchführen. Bei einer Fremdevaluation dagegen sind Evaluierete und Evaluierende nicht identisch. Nach dieser Logik ist eine externe Evaluation immer eine Fremdevaluation, während eine interne Evaluation als Fremd- oder als Selbstevaluation durchgeführt werden kann. Eine interne Fremdevaluation liegt also vor, wenn innerhalb einer Organisation eine Abteilung eine andere evaluiert. Da dies in der pädagogischen Praxis jedoch eher selten vorkommt, ist es verständlich, wenn interne Evaluation und Selbstevaluation bzw. externe Evaluation und Fremdevaluation hier meist gleichgesetzt werden, auch wenn sie sich konzeptionell trennen lassen.

3.2 Funktionen von Selbstevaluation

Selbstevaluation ist eine spezifische Form von Evaluation. Daher sollte man erwarten können, dass sie grundsätzlich in der Lage ist, die gleichen Funktionen wie sonstige Evaluationsansätze zu erfüllen (vgl. Abschnitt 2.2). Aufgrund ihrer spezifischen Rollenkonstellation, die sie von anderen Modellen unterscheidet, ist allerdings damit zu rechnen, dass sich bei der Selbstevaluation Schwerpunkte bzw. funktionale Verschiebungen bei den möglichen vier Hauptfunktionen ergeben.

Bei der *Entwicklungsfunktion* sind deutliche Vorteile zu erwarten, da die Generierung von Erkenntnissen, das Ziehen von Konsequenzen und deren Umsetzung idealerweise in einer Hand liegen. Da die Praktiker selbst die evaluationsrelevanten Daten erheben bzw. eng daran beteiligt sind, kann es bereits im Evaluationsprozess immer zu einer sofortigen Rückspeisung von Ergebnissen in die Praxis kommen. Die Fragestellungen der Selbstevaluation wurden zuvor selbst ausgewählt, so dass gewährleistet sein sollte, dass die Ergebnisse als relevant aufgegriffen werden und auch zu Konsequenzen führen.

Inwieweit Selbstevaluation im Sinne der *Entscheidungsfunktion* den Beschluss über Fortsetzung, Implementierung oder Beendigung von Maßnahmen unterstützen kann, dürfte von der spezifischen Umstellung abhängen. Falls die Praktiker die fragliche Entscheidung eigenständig treffen können, sind hier keine Probleme zu erwarten, vorausgesetzt die Selbstevaluation hat tatsächlich die entscheidungsrelevanten Informationen geliefert. Liegt die Entscheidungshoheit dagegen bei anderen Akteuren, so wird es darauf ankommen, dass die Selbstevaluationsergebnisse von den Entscheidern als glaubwürdig und relevant eingeschätzt werden. Hier gewinnt also die Frage von Gütekriterien der Selbstevaluation an Gewicht.

Die *Kontrollfunktion* von Evaluation widerspricht bis zu einem gewissen Grad der Definition von Selbstevaluation. Wenn die Praktiker selbst *owner of the process* sind impliziert das, dass es keinen externen Auftraggeber gibt, der eine Kontrolle verlangen kann. Die Kontrollfunktion verschiebt sich hier also zur *Selbstkontrollfunktion*, indem Zielsetzungen expliziert und deren Verfolgung im Sinne einer Selbstvergewisserung systematisch überwacht werden. Allerdings kann auch die externe Kontrollfunktion eine Rolle bei Selbstevaluationen spielen, wenn Mischformen vorliegen, bei denen die Praktiker nur teilautonom bei Initiierung und Realisierung der Selbstevaluation sind, wie es etwa bei einer „delegierten Selbstevaluation“ (Müller-Kohlenberg & Beywl, 2003) der Fall ist.

Die *Lernfunktion* von Evaluation schließlich kann bei einer Selbstevaluation insbesondere da zum Tragen kommen, wo vorher Wissen und Klarheit über Bedingungen und Folgen des eigenen professionellen Handelns wenig bewusst oder explizit waren.

Meist werden vor allem Entwicklungs- und Lernfunktion als wichtigste Ziele von Selbstevaluation betont (vgl. etwa Müller-Kohlenberg & Beywl, 2003). Die obige Diskussion hat zeigt, dass auch die Kontroll- und Entscheidungsfunktion unter bestimmten Bedingungen eine Rolle spielen kann. Ausführlichere Auseinandersetzungen mit Funktionen der Selbstevaluation finden sich bei von Bühren, Killus & Müller (1998), Burkard (1995), König (1998), Liebold (1998), MacBeath, Schratz, Meuret & Jakobsen (2000), Nevo (2001) oder Spiegel (1994).

Über die Erfüllung von allgemeinen Evaluationsfunktionen hinaus werden mit dem Ansatz der Selbstevaluation häufig auch weitergehende Erwartungen an ihren Nutzen verknüpft, die sich nicht im engeren Sinne als Evaluationsfunktionen bezeichnen lassen. Sie ergeben sich aus der spezifischen Rollenkonstellation, da ja wesentliche Funktionen der Evaluierung durch Praktiker vorgenommen werden. Im Einzelnen sind dabei vor allem die Aspekte Professionalisierung (vgl. Darling-Hammond & Wise, 1992; Lüders, 1998) und Organisationsentwicklung (vgl. Bühren, 1995) zu nennen.

Mit einer *Professionalisierung* der selbstevaluierenden Praktiker kann in zweifacher Hinsicht gerechnet werden. Erstens erweitern sie ihr Kompetenzspektrum um spezifische Evaluationskompetenzen, unabhängig davon ob sie die Selbstevaluation völlig eigenständig oder mit Hilfe einer externen Begleitung realisieren. Zweitens können Entwicklungs-, Lern- und Selbstkontrollfunktion der Selbstevaluation zur Professionalisierung ihres allgemeinen beruflichen Tuns beitragen.

Ein Nutzen von Selbstevaluation in Form von *Organisationsentwicklung* ist infolge des partizipativen Vorgehens zu erwarten. Denn durch die Einbeziehung der Praktiker in den Evaluationsprozess wird eine zentrale Forderung der Organisationsentwicklung, nämlich Betroffene zu Beteiligten zu machen, verwirklicht. In der Praxis lassen sich oft enge Verbindungen zwischen Selbstevaluation und Organisationsentwicklungsprozessen feststellen.

3.3 Beispiele aus verschiedenen pädagogischen Praxisfeldern

Ansätze zur Selbstevaluation werden heute in allen wichtigen pädagogischen Praxisfeldern angewandt. Eine Vorreiterrolle übernahm dabei in den 80er Jahren der Bereich Sozialpädagogik und soziale Arbeit (s. 3.3.1). Schule (s. 3.3.2), Hochschule (s. 3.3.3) und Weiterbildung (s. 3.3.4) griffen in der Folge die Entwicklung auf bzw. entwickelten eigenständige Formen. Die konkrete Ausprägung von Selbstevaluation in den verschiedenen Feldern erfolgte jeweils stark unter dem Einfluss von fachspezifischen Traditionen und Entwicklungen. Während etwa im schulischen Bereich eine starke Verbindung zum Themenkomplex Schulentwicklung, Schulprogramme und Schulautonomie festzustellen ist, besteht in der Weiterbildung ein enger Zusammenhang mit der Diskussion von übergeordneten Qualitätsmanagementsystemen. Im Folgenden werden die wichtigsten Strömungen der Selbstevaluation in den vier genannten Bereichen dargestellt und jeweils anhand eines Praxisbeispiels illustriert.

3.3.1 Sozialpädagogik und soziale Arbeit

Auf breiterer Basis und als eigenständiges Konzept wurde Selbstevaluation zuerst im Bereich der sozialen Arbeit etabliert. Dies verwundert nicht, wenn man bedenkt, dass der soziale Bereich bereits sehr früh in das Spannungsverhältnis zwischen stagnierenden Mittelzuweisungen und schwer nachweisbaren Leistungen geraten ist und sich damit sehr früh mit der Qualitätsfrage konfrontiert sah. Selbstevaluation ist hier als Versuch zu sehen, einem gestiegenen Qualitätsbewusstsein gerecht zu werden und gleichzeitig Eigenständigkeit und Autonomie gegenüber externen Kontrollansätzen zu bewahren.

Heiner (1988), die als Wegbereiterin der Selbstevaluation in der sozialen Arbeit bezeichnet werden kann, bezieht sich zunächst auf angelsächsische Autoren

wie Tripodi (1983), welche die Tradition einer systematischen Evaluation von Praxis durch Praktiker begründet haben. Heiner kritisiert diesen Ansatz jedoch als zu forschungs- und methodenfixiert und betont die Notwendigkeit, dass Selbstevaluation in erster Linie der Praxis zu dienen habe. Im Hinblick auf forschungsmethodische Standards sei der Anspruch relativ niedrig anzusetzen, da es nicht um verallgemeinerbare Aussagen gehe, sondern um konkrete Veränderungen auf der Mikroebene des eigenen Handelns.

Die meisten Ansätze zur Selbstevaluation in Sozialpädagogik und sozialer Arbeit, die in der Folge veröffentlicht wurden (z. B. König, 1998, 2000; Liebald, 1998; von Spiegel, 1993), betonen in besonderem Maße den Bezug auf das eigene professionelle Handeln der Praktiker, also die individuelle Ebene. Erklärtes Ziel ist es den beruflichen Alltag besser bewältigen zu können, indem Fragen wie „Leiste ich gute Arbeit? Könnte ich anders mehr erreichen? Habe ich mich richtig verhalten?“ (Heiner, 1988, S. 7) auf Grundlage systematisch erhobener Informationen und deren Bewertung beantwortet werden. Daneben wird aber auch die Qualitätsentwicklung auf überindividueller Ebene angestrebt. Liebald (1998) etwa unterscheidet die personenbezogene Selbstevaluation von jener, die sich auf Teams, Organisationseinheiten oder kleinere Organisationen bezieht.

Praxisbeispiel 1: Angeleitete Selbstevaluation im Team (Eisenlohr, 1998)

Ausgangslage: Im Rahmen eines Modellprojekts mit dem Titel „Qualitätsentwicklung durch Selbstevaluation“ hat sich ein Sachgebiet eines Allgemeinen Sozialen Dienstes (ASD), das aus acht Mitarbeitern besteht, das Ziel gesetzt, gemeinsam Qualitätsentwicklung zu betreiben und Standards zu entwickeln.

Ziele: Auf der Basis einer Ist-Stand-Analyse werden allgemeine Praxisziele (z. B. „Wir wollen die Zufriedenheit unserer KundInnen erhöhen“) und die Untersuchungsziele (z. B. „Überprüfung von Standards: Welche gibt es bereits und welche sind methodisch sinnvoll und ‚kundenorientiert‘?“) der Selbstevaluation festgelegt. Konkret soll es um die Untersuchung und Verbesserung des individuellen Beratungsverhaltens und den dabei gefällten Entscheidungen gehen.

Vorgehen: Nach der Phase der gemeinsamen Zielfindung werden verschiedene Instrumente zur begleitenden Dokumentation von Beratungsprozessen erstellt und mehrfach weiterentwickelt. So dient etwa ein „Kooperationsbogen“ zur Erfassung der Zusammenarbeit mit externen Partnern wie z. B. einer Tagesgruppe für Problemkinder. Er wird nicht nur zur Dokumentation eingesetzt, sondern auch zur Vorbereitung der Kooperationskontakte und zur Überprüfung der dabei getroffenen Vereinbarungen. Zur Auswertung werden eigene Bögen entwickelt, die quantitativ z. B. den Erfolg der Kooperationskontakte erfassen sollen.

Ergebnisse: Die Auswertung erbrachte eine Vielzahl von Erkenntnissen über die ursprünglich festgelegten Untersuchungsziele. Beispielsweise wurde festgestellt, dass die kollegiale Beratung einen wichtigen Beitrag bei der Wahrnehmung von „blinden Flecken“ im Beratungsalltag leisten kann. Insgesamt nennt die Autorin als Ergebnis der Auseinandersetzung mit diesen Ergebnissen: (1) Eine Verbesserung individueller Arbeitsweisen; (2) Das Sichtbarmachen und Sichern von Qualität; (3) Die Integration von Evaluation und Alltagshandeln; (4) Verbesserte Klientenbeziehungen; (5) Mehr Teamarbeit und Kooperation; (6) Stärkung des Transfers von Erkenntnissen und Methoden auf andere Kollegen.

Wie im Praxisbeispiel 1 deutlich wird, sind die Prozessschritte der Selbstevaluation hier von vornherein auf eine Verbesserung der Praxis angelegt. Durch die starke Einbeziehung der Mitarbeiter dient bereits die Entwicklung und Diskussion der Instrumente zur Selbstevaluation der individuellen Professionalisierung, der Teamentwicklung und der Übernahme von Verantwortung. Schon die Anwendung der Instrumente soll eine Qualitätsverbesserung bewirken, da sie Abläufe durch deren permanente Reflexion systematisieren und strukturieren. Obwohl ein ganzes Sachgebiet aus mehreren Mitarbeiterinnen und Mitarbeitern die Selbstevaluation gemeinsam durchführt, stehen die einzelnen Mitarbeiter und ihre professionelle Weiterentwicklung im Zentrum.

3.3.2 Schule

Im schulischen Bereich wird Selbstevaluation meist im Kontext von Schulentwicklung und Schulprogrammarbeit betrieben (vgl. Burkard & Eikenbusch, 1998; Klafki, 1998; Posch & Altrichter, 1998; Radnitzky & Schratz, 1999). So kann sie etwa der Bedarfserhebung und Ist-Stand-Analyse zu Beginn von Schulentwicklungsprozessen oder der kontinuierlichen Überprüfung von im Schulprogramm formulierten Zielen dienen. Teilweise sind Selbstevaluationen im schulischen Bereich mit Fremdevaluation durch die Schulaufsicht komplementär verknüpft, was aber kontrovers diskutiert wird (vgl. Almer, Schratz, Pendl & Rieger, 1998; Specht, 1998; Strittmatter, 1997).

Als Einflüsse fungierten hierzulande oft Vorbilder aus dem europäischen Ausland, z. B. Schottland, Skandinavien oder die Niederlande, wo Selbstevaluation bereits mehr oder weniger institutionalisierter Bestandteil der jeweiligen Schulsysteme ist (vgl. Burkard, 1995). In Deutschland wurde Selbstevaluation von Schule dann vor allem im Rahmen verschiedener Schulversuche seit Beginn der 90er Jahre erprobt und eingeführt (z. B. Buhren, Killus & Müller, 1998, 1999; Burkard, 1995).

Wichtigste Ziele der schulischen Selbstevaluation sind die Weiterentwicklung und Qualitätsverbesserung der schulischen Arbeit auf allen Ebenen des

Schullebens. Durch die Verknüpfung mit Schulentwicklung und Schulprogrammarbeit liegt der Schwerpunkt dabei zunächst auf der Schule als Organisation, aber auch das konkrete unterrichtliche Handeln der einzelnen Lehrkräfte kann unter die Lupe der Selbstevaluation genommen werden. Im letzteren Fall ist der Begründungszusammenhang dann weniger die Schulentwicklung, sondern eher die Forderung, Evaluation zum Bestandteil der individuellen Arbeitskultur von Pädagoginnen und Pädagogen zu machen (vgl. Burkard, 1995).

Methodisch kommen bei der schulischen Selbstevaluation tendenziell die herkömmlichen Methoden sozialwissenschaftlicher Forschung zum Einsatz, beispielsweise Fragebögen, Interviews und Beobachtungen (z. B. Buhren, Killus & Müller, 1999). Oft wird aber auch auf „weichere“ Methoden zurückgegriffen, die wesentlich offener und assoziativer angelegt sind, wie etwa Rollenspiele oder Gedankenlandkarten (z. B. Eikenbusch, 1997; MacBeath, Schratz, Meuret & Jakobsen, 2000).

Praxisbeispiel 2: Selbstevaluation von offenen Lernformen mit Hilfe von Qualitätsindikatoren (Buhren, 1997)

Ausgangsbedingungen: Ein österreichisches Gymnasium hat vor zwei Jahren offene Lernformen in einigen Fächern der Jahrgänge 7 und 8 eingeführt. Eine Lehrergruppe entscheidet sich, diese Innovationsmaßnahme selbst zu evaluieren.

Ziele: Erfolge und Misserfolge sollen analysiert werden, um eine Basis für eventuell notwendige Veränderungen zu schaffen. Außerdem will man den Kritikern der Innovation im Kollegium etwas Konkretes entgegenhalten können.

Vorgehen: Die ursprünglichen Ziele der Maßnahme werden rekapituliert und in Form von mehreren Qualitätsindikatoren für jedes Ziel operationalisiert. Für den Zielbereich „Selbstständigkeit und Eigenverantwortlichkeit der Schüler steigern“ lauten etwa zwei der Indikatoren „Die Schüler versuchen Lösungen zunächst ohne die Hilfe des Lehrers zu finden“ und „Die Schüler wenden sich bei Problemen an ihre Mitschüler“. Diese Indikatoren fließen dann in die Entwicklung eines Lehrerfragebogens und eines Beobachtungsbogens für Unterrichtshospitationen ein. Ergänzt werden diese Instrumente um einen kurzen offenen Schülerfragebogen. Parallel zum Testlauf der Instrumente werden Maßstäbe dafür festgelegt, welche Werte als erfolgreiches Ergebnis gelten sollen. Nach der Datensammlung können die Daten von 14 Lehrerfragebögen, 6 Unterrichtshospitationen und 220 Schülerfragebögen in die abschließende Auswertung einbezogen werden.

Ergebnisse: Insgesamt und auch auf Ebene der klassenspezifischen Auswertung ist man mit den erreichten Werten zufrieden. Die Profilauswertung zeigt

aber, dass im Zielbereich Gruppenarbeit noch Optimierungen notwendig sind. Hierzu werden konkrete Maßnahmen geplant, wie etwa die Erstellung von Materialien für Teamarbeit und gezielte Anleitung. Eine weitere Erhebung wird für das kommende Halbjahr geplant, um Veränderungen sichtbar zu machen.

Die Selbstevaluation dient im Praxisbeispiel 2 der Verbesserung einer schulischen Innovation, ist darüber hinaus aber mit dem Ziel der Organisations- bzw. Schulentwicklung verknüpft. Daneben wird hier aber auch explizit das Ziel genannt, die Ergebnisse der Evaluation als Argumentationshilfe gegenüber kritischen Kollegen zu nutzen (Lern- bzw. PR-Funktion). Da von den getroffenen Konsequenzen auch unterrichtliche Maßnahmen einzelner Lehrkräfte betroffen sind, hat die Selbstevaluation neben der schulischen Ebene auch die personale Ebene im Fokus. Methodisch orientiert sich dieses Beispiel an „klassischen“ Methoden der Evaluation (Fragebögen und Beobachtungen).

3.3.3 Hochschule

Auch der Hochschulbereich ist in den letzten Jahren verstärkt in den Fokus der Qualitätsdebatte gerückt. Erklärt wird das mit einem Mehr an Entscheidungsgewalt, die in Anwendung des Subsidiaritätsprinzips zunehmend vom Staat auf die Hochschulen übertragen wird (Fischer-Bluhm, 2000). Außerdem steige angesichts sich rasant verändernder gesellschaftlicher Bedingungen das Bewusstsein, dass die bisher vorherrschende Input-Steuerung von Qualität, die im Wesentlichen durch Haushaltsvorgaben, Vorschriften und Berufungen realisiert wurde, ergänzt werden muss um Methoden der Output-Steuerung, die nach den konkreten Leistungen der Hochschulen fragt (Müller-Böling, 1997). In der öffentlichen Wahrnehmung äußert sich das in den umstrittenen Hochschulrankings, wie sie inzwischen regelmäßig zu Semesterbeginn in den größeren Wochenzeitschriften erscheinen (vgl. Pechar, 1997).

Welche Rolle spielen dabei nun Ansätze zur Selbstevaluation? Zwar gibt es auch hier Vorschläge zur individuellen Professionalisierung von Dozenten (Abs, Raether, Tippelt, & Vögele, 2000; Schratz, 1997). In der Regel wird jedoch im universitären Qualitätsdiskurs unter Selbstevaluation etwas anderes verstanden, nämlich die Selbstbeurteilung ganzer Fachbereiche, welche in den Kontext eines übergreifenden Modells zur Qualitätsentwicklung eingebettet ist (Reisert & Carstensen, 1998). Dieses Modell, welches Selbstevaluationen mit externen *peer-reviews* verknüpft, wurde in den 90er Jahren durch das Vorbild der niederländischen Hochschulevaluation angeregt (vgl. Richter, 1997). Basis ist dabei eine selbstevaluierende Selbstbeschreibung, die von den verschiedenen Beteiligten eines Fachbereichs erstellt wird. Mit Hilfe eines Fragenkatalogs, der unter anderem Angaben über verschiedene Leistungsindikatoren (vgl. Müller-Böling, 1997) wie Studierendenzahlen oder angeworbene Drittmittel beinhaltet,

soll ein spezifisches Profil des Fachbereichs herausgearbeitet werden. Dieser wird dann von einer externen Kommission begutachtet, die zusätzlich eigene Datenerhebungen vor Ort durchführt, um das Bild zu vervollständigen. Es handelt sich hier also um ein Mischmodell, das Fremd- und Selbstbeurteilung verbindet.

Praxisbeispiel 3: Selbst- und Fremdevaluation von Studium und Lehre (Verbund Norddeutscher Universitäten, 2001)

Ausgangslage: Im Verbund Norddeutscher Universitäten spielt Selbstevaluation seit 1994 eine wichtige Rolle. Im Gesamtkonzept zur Evaluation von Studium und Lehre ergänzt sie die externe Evaluation durch eine Expertenkommission. Meist wird an den beteiligten Hochschulen jedes Jahr ein bestimmter Fachbereich evaluiert, so z. B. das Fach Psychologie 1999/2000.

Ziele: Zentrales Ziel der Evaluation ist die Qualitätsentwicklung in den Bereichen Lehre, Studien- und Prüfungspraxis und Kommunikation zwischen Lehrenden und Studierenden. Ausgangspunkt sind dabei immer die Meinungen der Beteiligten und die jeweils vor Ort gegebenen Bedingungen.

Vorgehen und Ergebnisse: Im jeweiligen Sommersemester erfolgt per Selbstevaluation eine Stärken-Schwächen-Beschreibung, die sich an einer detaillierten Checkliste orientiert. Dazu wird eine Arbeitsgruppe aus Lehrenden und Studierenden gebildet, die über Schwerpunktsetzungen innerhalb der Checkliste entscheidet und anschließend die dafür notwendigen Erhebungen durchführt. Am Ende des Prozesses steht eine offiziell verabschiedete Selbstbeschreibung. Sie dient einer externen Expertenkommission zur Vorbereitung ihrer Begehung des Fachbereichs im darauffolgenden Wintersemester. Vor der Veröffentlichung diskutiert die Kommission ihren Bericht in einer auswertenden Konferenz gemeinsam mit den Betroffenen. Dabei stehen die Konsequenzen der Ergebnisse im Mittelpunkt. Als letzten Schritt trifft der evaluierte Fachbereich konkrete Zielvereinbarungen mit der Universitätsleitung über Maßnahmen, die als Folge der Evaluation eingeleitet werden sollen.

Wie am Praxisbeispiel 3 zu sehen ist, unterscheiden sich hier Gegenstand und Vorgehen von den ersten beiden Praxisbeispielen aus den Bereichen soziale Arbeit und Schule. Hauptunterschied ist die Einbeziehung einer externen Expertenkommission (*peer-review*). Deutlich wird aber auch, dass zwei zu den obigen Beispielen ähnliche Aspekte betont werden: Einerseits die Verbesserung der Praxis, die sich in den konkreten Zielvereinbarungen mit der Hochschulleitung niederschlägt, andererseits die Partizipation aller Betroffenen, die sowohl beim Anfertigen der Selbstbeschreibung als auch beim Diskutieren von Ergebnissen und Konsequenzen unverzichtbar ist und damit einen Beitrag zur Organisationsentwicklung leistet.

3.3.4 Erwachsenen- und Weiterbildung

Der Bereich der Erwachsenen- und Weiterbildung ist deutlich weniger homogen strukturiert als die Bereiche Sozialpädagogik, Schule und Hochschule (vgl. Prenzel, Mandl & Reinmann-Rothmeier, 1997). Daher liegt hier auch eine größere Vielfalt von unterschiedlichen Ansätzen vor, bei denen Selbstevaluation eine Rolle spielt. Die drei wichtigsten sind übergeordnete Qualitätsmanagementansätze wie das Total Quality Management, Benchmarking und die persönliche Selbstevaluation.

Die Auseinandersetzung mit dem *Total Quality Management* (TQM) spielt vor allem im Bereich der beruflichen Weiterbildung eine wichtige Rolle. Der Fokus der Qualitätsentwicklung liegt hier in erster Linie auf den Weiterbildungsträgern und -anbietern, also auf der Ebene der Organisation (Franz, 1999; Stark, 2000). Der wichtigste TQM-Ansatz in diesem Bereich stammt von der European Foundation for Quality Management (EFQM). Er beschreibt einen kontinuierlichen und selbstevaluierenden Verbesserungsprozess, der die folgenden Schritte umfasst (Franz, 1999): (1) Selbstbewertung durch eine Stärken- und Schwächenanalyse anhand eines Fragenkatalogs; (2) Finden von Verbesserungsmöglichkeiten; (3) Umsetzung je nach Prioritäten der Organisation.

TQM unterscheidet sich damit deutlich von einem anderen, in der Weiterbildung häufig diskutierten Ansatz des Qualitätsmanagements, nämlich der Normenreihe DIN EN ISO 9000-9004 (vgl. Wunder, 1995; Wuppertaler Kreis, 1996). Während die DIN-Normen eine einmalige und standardisierte Zertifizierung von Weiterbildungsanbietern durch externe Agenturen ermöglichen, sieht TQM interne Selbstbewertungen vor und betont die fortlaufende Organisationsentwicklung. Ohnehin ist im Weiterbildungsmarkt insgesamt eine Verschiebung der Perspektive von der Qualitätssicherung zur Qualitätsverbesserung zu beobachten (Küchler, 1999). Damit geht u. a. die Erkenntnis einher, dass alle Beteiligten am Qualitätsdialog zu beteiligen sind und dass die systematische Selbstvergewisserung im Vordergrund zu stehen habe (Meisel, 1999). Im Bereich der allgemeinen Erwachsenenbildung sind Verfahren zu finden, die dem TQM ähneln, etwa bei der Selbstevaluation mittels eines Fragenkatalogs, wie sie in niedersächsischen Volkshochschulen praktiziert wird (Heinen-Tenrich, 1999; vgl. Praxisbeispiel 4).

Benchmarking ist ein weiterer Ansatz der Selbstevaluation auf Ebene der Organisation, der in erster Linie den Vergleich mit anderen Bildungsträgern erlauben soll (Halfar & Lehnerer, 1997). Dabei bewertet sich eine Organisation anhand von Checklisten und Instrumenten selber. Durch den Vergleich mit den Kennzahlen anderer Organisationen, die etwa in Form einer Datenbank vorliegen oder auf bilateraler Basis ausgetauscht werden, steht dann ein Maßstab für die eigene Qualitätsentwicklung zur Verfügung.

Neben diesen Ansätzen, die auf Ebene von Organisationen ansetzen, wird stellenweise auch in der Erwachsenenbildung der Ansatz *persönlicher Selbstevaluation* (vgl. Abschnitt 3.3.1) praktiziert. So propagiert Brödel (1999) etwa den Einsatz pädagogischer Tagebücher in Kombination mit standardisierten Teilnehmerbefragungen zur datengestützten Reflexion des eigenen Handelns als Dozent.

Praxisbeispiel 4: Qualitätsentwicklung durch Selbstevaluation in der Volkshochschule (Heinen-Tenrich, 1999)

Ausgangsbedingungen: Im Land Niedersachsen haben sich 34 Volkshochschulen (VHS) in einem „Qualitätsring“ zusammengeschlossen, um einen gemeinsamen Ansatz zur Qualitätssicherung mittels Selbstevaluation zu verfolgen.

Ziele: Anstelle der externen Kontrolle anhand vorgegebener Normen und Standards soll eine kontinuierliche selbstevaluative Überprüfung und Anpassung der eigenen Qualität erfolgen. Als Basis will man dabei immer die individuell gegebenen Rahmenbedingungen berücksichtigen.

Vorgehen: Ausgangspunkt ist ein Selbstverpflichtungskatalog, der auf den Qualitätsstandards des Vereins „Weiterbildung Hamburg“, einer extern zertifizierenden Organisation (vgl. Krüger, 1999), basiert. Diese Standards werden um konkrete Indikatoren und Erschließungs- und Beschreibungskriterien ergänzt, so dass ein Merkmalsraster entsteht. Dieses erfasst die Einrichtungen, Anforderungen, Leistungen und Aufgaben der jeweiligen VHS und deckt dabei pädagogische, institutionelle, organisatorische und finanzielle Fragen umfassend ab. So lautet etwa eine Frage aus dem Qualitätsbereich „Präsenz und Service“: „Wie wird die zeitliche, räumliche und verkehrstechnische Erreichbarkeit der VHS sichergestellt?“ Das gesamte Instrumentarium wird in Form einer Arbeitshilfe veröffentlicht, die freiwillig bezogen werden kann. Im Rahmen von extern moderierten Workshops besteht darüber hinaus für eine einzelne VHS die Möglichkeit, im Mitarbeiterteam den Fragenkatalog gemeinsam zu bearbeiten.

Ergebnisse: Infolge der Auseinandersetzung mit dem Fragenkatalog werden Veränderungen in den identifizierten Problembereichen beschlossen und umgesetzt, etwa die Neuregelung von Öffnungszeiten oder systematische Mitarbeiterfortbildungen. Als Nebeneffekt der Qualitätsarbeit wird eine allgemeine Verbesserung der „Arbeitskultur“ berichtet, die sich etwa in der intensivierten innerbetrieblichen Kommunikation und Kooperation zeigt.

Charakteristisch für den Bereich Erwachsenen- und Weiterbildungsbereich ist beim Praxisbeispiel 4, dass die Selbstevaluation sich stark auf die Organisationsebene bezieht. Die Weiterbildungsinstitution soll als Ganzes auf

den Prüfstand und gemeinsam sollen Konsequenzen gezogen werden. Das Handeln der einzelnen Mitarbeiter, etwa der Dozenten, kann dabei zwar auch in den Fokus der Aufmerksamkeit geraten, wird jedoch immer unter der Perspektive organisationaler Qualitätsmerkmale thematisiert. Deutlich ist also auch hier die Betonung der kontinuierlichen Entwicklungsorientierung. Die strukturierte Auseinandersetzung und gemeinsame Reflexion soll nicht dem einmaligen Feststellen des Vorhandenseins oder Fehlens von Qualitätsnormen dienen, sondern zu Veränderung und Verbesserung auffordern.

4. Selbstevaluation im BLK-Programm SEMIK

In diesem Abschnitt wird die Umsetzung von Selbstevaluation anhand eines Beispiels aus dem Modellversuchsprogramm SEMIK („Systematische Einbeziehung von Informations- und Kommunikationstechnologien in Lehr- und Lernprozesse“) genauer dargestellt. SEMIK war ein von der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK) im Zeitraum 1998 bis 2003 finanzierter Zusammenschluss von 25 Einzelprojekten, deren Zielsetzung es war, Medien, insbesondere die neuen Medien in allen Schularten und -stufen auf breiter Basis zu implementieren und dies mit der Umsetzung innovativer Lehr- und Lernformen zu verbinden (Mandl, Hense & Kruppa, 2002, 2003; Mandl, Reinmann-Rothmeier & Gräsel, 1998). Didaktisches Leitkonzept von SEMIK war das problemorientierte Lernen (Reinmann-Rothmeier & Mandl, 2001).

Vier der 25 SEMIK-Projekte schlossen sich Anfang 2000 im sogenannten SEMIK-Selbstevaluationsverbund zusammen. Die folgenden Darstellungen beziehen sich auf die Arbeit dieses Verbunds, der von der LMU München moderiert, koordiniert und wissenschaftlich betreut wurde.

4.1 Konzept der Selbstevaluation

Mit der Arbeit im SEMIK-Selbstevaluationsverbund wurden mehrere Ziele auf verschiedenen Ebenen verfolgt:

- Wichtigste evaluationsbezogene Funktion war die Entwicklung bzw. Verbesserung von Arbeit und Ergebnissen der vier beteiligten Einzelprojekte. Primär sollte die Selbstevaluation der kontinuierlichen Qualitätssicherung und Qualitätsverbesserung während der Laufzeit dienen. Da es in SEMIK zusätzlich eine zentrale externe Evaluation gab, sollte die Selbstevaluation sich ganz auf die Entwicklungsfunktion konzentrieren.
- Auf Ebene der beteiligten Lehrkräfte sollte es sich daneben um einen Beitrag zu deren Professionalisierung handeln.

- Auf administrativer Ebene realisierten die vier Projekte ihre projekt-spezifische wissenschaftliche Begleitung und Evaluation, wie sie im SEMIK-Rahmenkonzept vorgesehen war.

Als Selbstevaluation wurden im SEMIK-Selbstevaluationsverbund Evaluationsmaßnahmen verstanden, welche maßgeblich durch Projektbeteiligte konzipiert und verantwortet, nicht aber notwendigerweise auch selbst durchgeführt wurden. Entscheidend war also nicht, dass die beteiligten Lehrkräfte tatsächlich alle Schritte des Evaluationsprozesses selbst durchführten, sondern dass sie „owner of the process“ waren. Das Ausgehen von den Bedürfnissen der selbstevaluierenden Praktiker und die Hoheit über die Daten werden einhellig als wichtige Faktoren einer gelingenden Selbstevaluation angesehen (z. B. Rolff, 1998; Strittmatter, 2000). Um das zur Umsetzung notwendige Know-how zu ergänzen, war die wichtigste Aufgabe der wissenschaftlichen Begleitung, die im Evaluationsprozess notwendige Expertise zur Verfügung zu stellen.

Das im Selbstevaluationsverbund praktizierte Vorgehen zur Selbstevaluation orientierte sich an den gängigen Modellen, die für den pädagogischen bzw. schulischen Bereich vorgeschlagen werden (z. B. Beywl & Henze, 1999; Buhren, Killus & Müller, 1999; Burkard, 1995; Dubs, 1998). Da sich die in der Literatur üblicherweise dargestellten Schritte sehr stark ähneln, kann der typische Ablauf in Form eines „Evaluationszyklus“ (Abbildung 2) zusammengefasst werden.

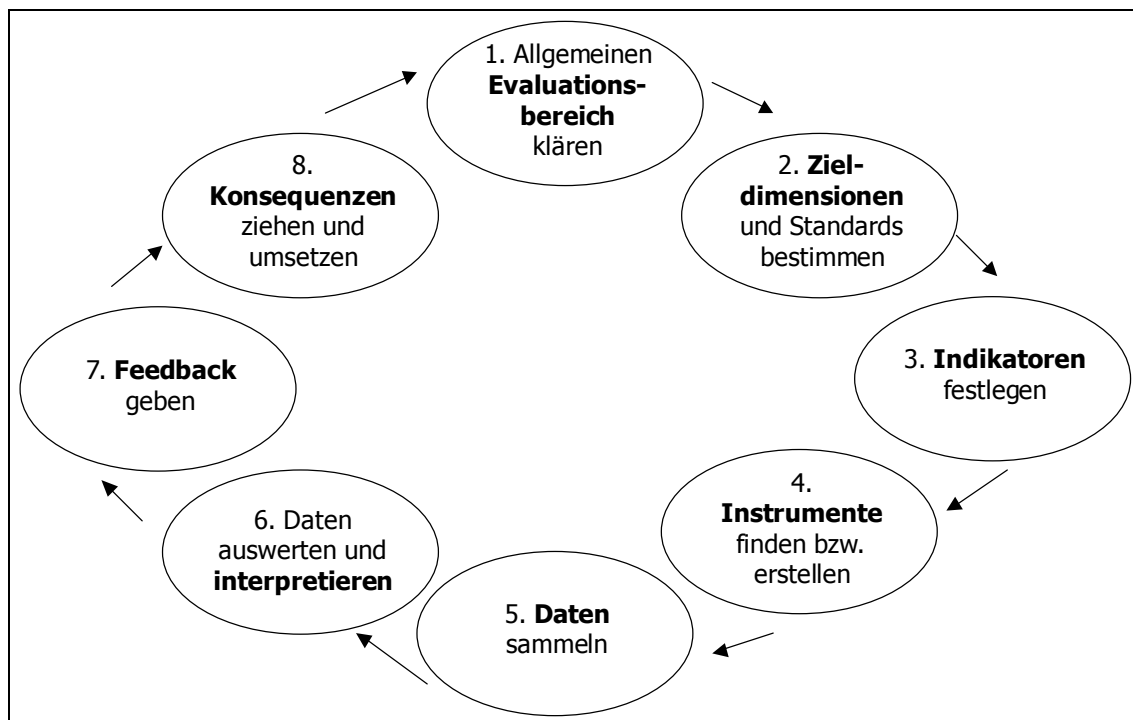


Abbildung 2: (Selbst-)Evaluationszyklus

Dieser Zyklus gleicht weitgehend dem Vorgehen bei einer externen Evaluation, die prozessbegleitend angelegt ist. Demgegenüber ist Selbstevaluation aber dadurch gekennzeichnet, „dass sie von Personen aus dem Kreis der für den Evaluationsgegenstand Verantwortlichen selbst durchgeführt und oft auch selbst beauftragt werden“ (Deutsche Gesellschaft für Evaluation [DeGEVal], 2002, S. 7). Im Idealfall liegen bei einer Selbstevaluation also die in Abbildung 2 skizzierten Schritte vollständig in der Hand jener Personen bzw. Institutionen, deren professionelles Handeln evaluiert wird. Dieser Idealfall wird allerdings selten in Reinform praktiziert, da meist die Hilfe von externen Experten in Anspruch genommen wird (vgl. DeGEVal, 2002). Eine beratende Begleitung von Selbstevaluation durch externe Experten erscheint durchaus angezeigt, wenn man bedenkt, dass für die Durchführung der einzelnen Schritte sozialwissenschaftliches Methodenwissen notwendig ist (vgl. König, 2000).

Der zyklische Charakter der Selbstevaluation und insbesondere der Schritt „Konsequenzen ziehen und umsetzen“ machen deutlich, dass die Selbstevaluation primär auf den Entwicklungsaspekt ausgerichtet war. Ergebnisse sollten bereits zur Laufzeit oder Entwicklungsphase in den Prozess zurückgespeist werden, um so frühzeitig Verbesserungen zu initiieren und Fehlentwicklungen abzufangen.

Eine Besonderheit in SEMIK war, dass Selbstevaluation hier nicht auf Ebene einer Einzelschule, sondern in vier Modellprojekten durchgeführt wurde. In der Literatur wird Selbstevaluation im schulischen Bereich bisher meist auf die Schulentwicklung von Einzelschulen bezogen. Dokumentierte Umsetzungsbeispiele betreffen daher immer die Kollegien von Einzelschulen. In SEMIK dagegen wurde die Projektarbeit jeweils von einer Gruppe von Lehrkräften verschiedener Einzelschulen getragen, welche nur für die Dauer der Projektlaufzeit bestand. Gegenstand der Selbstevaluation waren damit auch nicht, wie sonst im schulischen Bereich üblich, Aspekte der Einzelschule wie etwa Schulklima oder Teile des Schulprogramms, sondern Aspekte der Projektarbeit, die sich aus den Zielen der einzelnen Vorhaben ableiteten.

Die Gegenstände, welche im Fokus der Selbstevaluation standen, lassen sich in zwei Ebenen unterscheiden: Die Ebene der Arbeit an dem jeweiligen Projekt als Ganzes (Makroebene) und die Ebene der alltäglichen Arbeit der einzelnen beteiligten Lehrkräfte (Mikroebene). Auf der *Makroebene* stand die Realisierung allgemeiner Projektziele im Vordergrund des Interesses. Daher spielte hier die enge Kooperation und Abstimmung mit den Projektleitungen der vier Projekte eine wichtige Rolle. Auf der *Mikroebene* sollten die Instrumente der Selbstevaluation der kontinuierlichen Qualifizierung des individuellen professionellen Handelns der einzelnen Lehrkräfte dienen, indem sie eine systematisierte und inhaltlich vertiefte Reflexion des unterrichtlichen Handelns ermöglichten. Dieses

Verständnis von Selbstevaluation auf individueller Ebene war ebenfalls ein wichtiger Aspekt der gemeinsamen Arbeit.

Im SEMIK-Selbstevaluationsverbund wurde ein begleitetes Modell der Selbstevaluation praktiziert, bei dem die jeweils projektbeteiligten Lehrkräfte und die Projektleitungen zwar „owner of the process“ waren, die notwendige Expertise und organisatorische Arbeit aber von externer Seite durch die wissenschaftliche Begleitung eingebracht wurde. Ihre Hauptaufgabe war die Beratung und Unterstützung bei der Planung, Durchführung und Auswertung von Selbstevaluationsmaßnahmen. Im Einzelnen waren die wichtigsten Aufgaben der wissenschaftlichen Begleitung für die Einzelprojekte:

- Beratung der Projektleitung bei der Konzeption der Selbstevaluation
- Durchführung von eintägigen Workshops pro Halbjahr mit den Projektbeteiligten
- Entwicklung von Prototypen der Selbstevaluationsinstrumente
- Rückmeldung von Selbstevaluationsergebnissen, soweit sie zentral ausgewertet wurden
- Unterstützung bei der Interpretation der Daten

4.2 Umsetzung

Die Arbeit in den Projekten konzentrierte sich auf halbjährliche Workshops, welche unter Anleitung und Moderation der LMU in den vier Projekten durchgeführt wurden. Dazwischen erfolgte die Umsetzung durch die Beratung der Projektleitungen, die Koordination des Verbunds und die jährliche Dokumentation von Aktivitäten und Ergebnissen. Im Folgenden wird dargestellt, wie im Rahmen der Workshops die Schritte des Selbstevaluationszyklus umgesetzt worden sind und welche Instrumente und Methoden verwendet wurden.

4.2.1 Selbstevaluationsworkshops mit projektbeteiligten Lehrkräften

In jedem der vier Projekte fand jeweils im Frühjahr und im Herbst ein eintägiger Workshop statt. Teilnehmer waren die Projektleitung und die projekttragenden Lehrkräfte (ca. 8-10 Personen). Die Workshops wurden durch die wissenschaftliche Begleitung in enger Absprache mit den Projektleitungen geplant und durchgeführt.

In den *Auftaktworkshops* wurde den teilnehmenden Lehrkräften jeweils eine Einführung in Ziele, Vorgehen und Methoden der Selbstevaluation gegeben. Zweiter zentraler Bestandteil war die gemeinsame Arbeit an den Zieldimensionen für die weitere Arbeit. Da es sich um die Selbstevaluation eines Modellversuchs handelte, standen dabei die innovativen Ziele des Modellversuchs im Vordergrund. Dazu gehörten etwa die Umsetzung problem-

orientierten Lernens, Schulentwicklung oder der Aufbau von Medienkompetenz.

In den *Folgeworkshops* wurde kontinuierlich an den anfangs gesetzten Zieldimensionen gearbeitet. Typischerweise erfolgte auf Basis von Inputmaterialien und -präsentationen der Begleitung zunächst eine inhaltliche Auseinandersetzung mit den Zieldimensionen, da es sich dabei wie erwähnt um innovative, also den Lehrkräften bisher wenig vertraute Konzepte handelte. Darauf aufbauend wurden Instrumente entwickelt, deren Einsatz besprochen und die damit gewonnenen Daten im folgenden Workshop ausgewertet, um anschließend Konsequenzen für die weitere Arbeit und die folgenden Evaluationsschritte zu ziehen. Die Instrumente sind jeweils als Prototypen von der Begleitung vorbereitet und im Workshop gemeinsam den Bedürfnissen angepasst worden. Oft ergab sich auch infolge des Ersteinsatzes der Instrumente der Bedarf zur Überarbeitung oder Verfeinerung.

In den *abschließenden Workshops* ging es darum, die Arbeit der vorangegangenen Jahre zusammenfassend auszuwerten und im Hinblick auf die Abschlussdokumentation aufzubereiten. Da im Sinne der „ownership of the process“ die Projekte über die Verwendung von Erhebungsergebnissen eigenständig entscheiden konnten, werden diese hier nicht berichtet. Sie haben aber Eingang in die Abschlussberichterstattungen der Einzelprojekte genommen (vgl. Mandl, Hense & Kruppa, 2003).

4.2.2 Umsetzungsschritte

Obwohl die Arbeit in den vier Projekten bzw. fünf Schulsets individuell auf die Ziele der Projekte und Bedürfnisse der Beteiligten zugeschnitten wurde, wird anhand der in Abbildung 2 wiedergegebenen Schrittfolge im Selbstevaluationszyklus das Vorgehen an dieser Stelle zusammenfassend dargestellt.

4.2.2.1 Allgemeinen Evaluationsbereich klären

Der erste Schritt, Klärung des allgemeinen Evaluationsbereichs, war im Rahmen der SEMIK-Selbstevaluation durch die Ziele von SEMIK und durch die Ziele der Einzelprojekte bereits vorgegeben.

4.2.2.2 Zieldimensionen bestimmen

Obwohl der allgemeine Evaluationsbereich damit nicht frei in den Projektgruppen gewählt werden konnte, bestand die Möglichkeit, innerhalb des allgemeinen Evaluationsbereichs Schwerpunkte zu setzen, bei denen die Beteiligten besonderen Handlungsbedarf sahen. Die Zieldimensionen der Selbstevaluation in einem der vier Projekte lauteten beispielsweise:

1. Problemorientierung
2. Selbstlernen verstärken
3. Medienkompetenz
4. Lernerfolgskontrolle
5. Unterrichtsorganisation

Die Intention war, diese in absteigender Priorität geordneten Zieldimensionen während der Projektlaufzeit zu bearbeiten. Im Laufe der Arbeit ergaben sich teilweise Bedürfnislagen, welche zu diesem frühen Zeitpunkt noch nicht absehbar waren. Im Sinne des Konzepts zur Selbstevaluation war es immer Ziel, diese aktuellen Bedürfnislagen soweit wie möglich zu berücksichtigen. Außerdem zeigte sich, dass man sich in der „Euphorie des Anfangs“ teilweise zu viele unterschiedliche Ziele gesetzt hatte und sich manches Ziel im Verlauf der Projektarbeit als sekundär erwies. So erklärt es sich, dass nicht alle der genannten Zieldimensionen in gleich starkem Maße bearbeitet worden sind. Die Vorteile dieses Vorgehens überwiegen aber, da so das partizipative Prinzip der Selbstevaluation adäquat umgesetzt wurde und die Forderung, immer das in der Selbstevaluation zu behandeln, was „unter den Nägeln“ brennt (vgl. Burkard, 1999), berücksichtigt wurde.

4.2.2.3 Indikatoren festlegen

Die Messung von Zieldimensionen kann nur über ihre Operationalisierung erfolgen, was wiederum empirisch beobachtbare Indikatoren erfordert. Im Rahmen der Selbstevaluationsworkshops erfolgte diese Operationalisierung auf unterschiedliche Weise.

Bei Zieldimensionen, welche eine tiefergehende inhaltliche Auseinandersetzung erforderten, weil sie begrifflich noch nicht zum Alltagsrüstzeug von Lehrkräften gehörten, wurden Indikatoren zunächst aus der theoretischen Auseinandersetzung mit den entsprechenden Konzepten gewonnen. So fußten etwa die Indikatoren für die Zieldimension Medienkompetenz auf dem entsprechenden Modell von Tulodziecki (1998), für die Dimension selbstgesteuertes Lernen auf einem Konzept von Simons (1992) und für die Dimension problemorientiertes Lernen auf dem SEMIK-Gutachten (Mandl, Reinmann-Rothmeier & Gräsel, 1998).

Bei anderen Zieldimensionen konnte in stärkerem Maße auf die eigenen Vorerfahrungen der Lehrkräfte zurückgegriffen werden. So basierten etwa die Indikatoren eines Beobachtungsbogen zum Thema Rollenverhalten von Lehrern und Schülern, welches für die gegenseitigen Hospitationen in einem der Projekte entwickelt werden sollte, auf einer gemeinsamen Metaplanarbeit im Workshop.

4.2.2.4 Instrumente finden bzw. erstellen

Unter den Instrumenten, welche im Laufe der Selbstevaluation in den vier Projekten neu entwickelt bzw. adaptiert wurden, finden sich u. a. Checklisten, Fragebögen, pädagogische Tagebücher, Beobachtungsbögen, Dokumentationsschemata, Gesprächsleitfäden und andere Instrumente. Diese Instrumente hatten unterschiedliche Quellen.

Bei einigen handelte es sich um völlige Neuentwicklungen, da die Zieldimension sehr speziell gewählt war oder kein adaptierbares Instrument zur Verfügung stand. Beispiele dafür sind etwa der bereits erwähnte Beobachtungsbogen für Rollenverhalten im Unterricht oder Checklisten, welche aus der inhaltlichen Auseinandersetzung mit den Zieldimensionen Problemorientierung, Medienkompetenz oder selbstgesteuertes Lernen hervorgegangen sind.

Teils konnten vorhandene Instrumente adaptiert werden. So etwa bei pädagogischen Tagebüchern, welche in ihrer ursprünglichen Fassung aus Forschungsprojekten der LMU stammten, oder bei einem Dokumentationsschema für Fortbildungssituationen, welches auf in der Literatur dokumentierten Beispielen aus dem Bereich Sozialpädagogik (vgl. Höning, 1988; Meinhold & Radatz, 1988) basierte.

4.2.2.5 Daten sammeln

Die Datensammlung mit neu entwickelten oder überarbeiteten Instrumenten erfolgte in der Regel zwischen den Workshops. Sie gestaltete sich naturgemäß je nach verwendetem Instrument unterschiedlich. So wurden Fragebogenerhebungen in der Regel von Seiten der wissenschaftlichen Begleitung aus in Papier- oder Online-Form realisiert. Checklisten, Tagebücher und Dokumentationsschemata wurden von den beteiligten Lehrkräften selbst bearbeitet. Die strukturierte Beobachtung von Unterrichtssituationen geschah im Rahmen gegenseitiger Unterrichtsbesuche in Lehrertandems. Selbstlernphasen wurden mittels eines Protokollrasters von Schüler/innen angefertigt.

Welche Erhebungen in welchem Umfang zwischen den Workshops realisiert werden sollten, wurde jeweils am Ende der Workshops vereinbart. Ein wichtiger Lerneffekt war dabei, dass die teils durchaus enthusiastische und immer engagierte gemeinsame Arbeit an den Zieldimensionen und den daraus entwickelten Instrumenten nicht immer garantierte, dass die Datensammlung in gleichem Maße engagiert erfolgte. Hier haben sich konkrete Vereinbarungen, welche im Konsens getroffen und im Workshopprotokoll festgehalten wurden, stark bewährt.

4.2.2.6 Daten auswerten und interpretieren

Die Auswertung der Daten erfolgte je nach Erhebung unterschiedlich. Jene, welche auf Grund ihres administrativen Aufwands durch die wissenschaftliche Begleitung durchgeführt wurden, wie etwa Fragebogenerhebungen, oder bei denen eine größere Datenmenge anfiel, wurden durch die Begleitung im Rahmen der Vorbereitung der Workshops ausgewertet.

Erhebungen, welche sehr „unterrichtsnah“ waren, wie Unterrichtsdokumentationen, Checklisten oder pädagogische Tagebücher sind dagegen als Teil der Workshops gemeinsam ausgewertet worden. Hier war auch jeweils das Ziel, als Beitrag zur Lehrerprofessionalisierung Wissen über einfache Verfahren der Datenauswertung an die Beteiligten weiterzugeben und diese einzuüben.

Gleiches gilt für die Interpretation der Daten, welche immer gemeinsam im Rahmen der Workshops realisiert wurde. Prämisse dabei war, dass ein Befund, welcher für die Beteiligten nicht sinnvoll interpretierbar ist, auch nicht zum Ziehen von Konsequenzen führen wird. Wo nötig wurde auch hier fehlendes Methodenwissen, etwa zur Interpretation einfacher statistischer Kennwerte, durch die LMU bereitgestellt.

4.2.2.7 Feedback geben

Das Feedback der wissenschaftlichen Begleitung ergänzte die Auseinandersetzung mit Ergebnissen in der Workshopgruppe soweit wie möglich um eine externe Perspektive. Dem im Konzept formulierten Selbstverständnis der Selbstevaluation zu Folge hatte die Begleitung dabei allerdings nur den Anspruch, eine weitere Sichtweise einzubringen, die ebenso große Berechtigung haben sollte wie jene der Beteiligten.

4.2.2.8 Konsequenzen ziehen und umsetzen

Soweit im Rahmen der Interpretation Handlungsbedarf gesehen wurde, war es auch Teil der Workshops, konkrete Konsequenzen zu diskutieren und zu beschließen. Diese konnten sich sowohl auf die Ebene des gesamten Projekts beziehen als auch auf die individuelle Ebene.

So führte beispielsweise die Rückmeldung von Kollegen, welche erarbeitete Unterrichtsmaterialien erprobt hatten und die dabei gemachten Erfahrungen in einem vorher entwickelten Rückmeldebogen festhielten, zur Überarbeitung kritischer Punkte durch die jeweils betroffenen Lehrkräfte. Auf diese Weise konnten Projektergebnisse auf individueller Ebene optimiert werden.

Aus anderen Ergebnissen dagegen wurden Konsequenzen gezogen, welche die gesamte Projektgruppe betrafen. So etwa, wenn infolge der erweiterten SEMIK-Fragebogenerhebung festgestellt wurde, dass der Aspekt der aktiven

Mediennutzung und -präsentation durch Schüler/innen bisher zu kurz gekommen war.

4.2.3 Instrumente der Selbstevaluation

Im Laufe der Arbeit in den fünf beteiligten Schulsets sind eine Vielzahl von Instrumenten für die Selbstevaluation der Projektarbeit entwickelt, erprobt, überarbeitet und eingesetzt worden. Im Einzelnen sind die folgenden Instrumente erstellt und eingesetzt worden:

Fragebogenverfahren:

- Medienkompetenzfragebogen (für Lehrer)
- Medienkompetenzfragebogen (für Schüler)
- Fragebogen zur Fortbildungsevaluation (für Fortbildungsteilnehmer)
- Fragebogen zum Fortbildungsbedarf (für Fortbildungsteilnehmer)
- Rückmeldebogen für die Erprobungsphase von Unterrichtsmaterial (für Erprober)
- Fragebogen zur Lernumgebung bei selbstgesteuerten Arbeitsphasen (für Schüler)

Gesprächsleitfäden:

- Leitfaden für ein Gruppengespräch mit Lehrern (für Projektleitungen)
- Leitfaden für ein Gruppengespräch mit Schülern (für Lehrer)

Pädagogische Tagebücher:

- Pädagogisches Tagebuch mit Leitfragen – allgemeine Version (für Lehrer)
- Halbstandardisiertes pädagogisches Tagebuch – täglich (für Lehrer)
- Halbstandardisiertes pädagogisches Tagebuch – wöchentlich (für Lehrer)

Beobachtungsbögen:

- Schema zur Unterrichtsbeobachtung – allgemein (für Hospitationen)
- Schema zur Unterrichtsbeobachtung – Lehrer- und Schülerrolle (für Hospitationen)
- Schema zur Fortbildungsbeobachtung (für Hospitationen)

Checklisten zur Unterrichtsplanung und -auswertung:

- Problemorientierter Unterricht (für Lehrer und Fortbildner)
- Selbstgesteuertes Lernen (für Lehrer und Fortbildner)
- Medienkompetenz (für Lehrer und Fortbildner)

Dokumentationsschemata:

- Dokumentation von Unterrichtseinheiten (für Lehrer)

- Dokumentation von Fortbildungseinheiten (für Fortbildner)
- Dokumentation von informellen Fortbildungskontakten (für Fortbildner)
- Dokumentation von Selbstlernphasen (für Schüler)
- Dokumentation von Selbstlernphasen (für Lehrer)

Planungs- und sonstige Dokumentationsinstrumente:

- W-Fragen-Raster (für Lehrer und Fortbildner)
- Verkürztes Sachberichtsschema (für Lehrer und Fortbildner)
- Begleitinformationen zu Lehr- und Lernmaterialien (für Lehrer)

4.3 Ergebnisse der SEMIK-Selbstevaluation

Welchen Ertrag hat die Selbstevaluation im SEMIK-Selbstevaluationsverbund erbracht? Um diese Frage zu beantworten, wurde eine Nachbefragung der beteiligten Projektleitungen per Interview durchgeführt. Diese hatte zum Ziel zu überprüfen, inwiefern das Ziel der allgemeinen Qualitätssicherung auf Ebene der Projektarbeit erreicht worden ist. Zentrale Ergebnisse dieser Überprüfung waren (vgl. Mandl, Hense & Kruppa, 2003):

Die Selbstevaluation wurde insgesamt als erfolgreich im Sinne der Qualitätssicherung für die Projekte gewertet. Besonders betont wurde dabei das Element der systematischen Selbstreflexion. Der Nutzen der Selbstevaluation bestand nach Wahrnehmung der Befragten vor allem darin, dass Prozesse und Ergebnisse der Datengewinnung zu einer vertieften Auseinandersetzung mit dem eigenen professionellen Handeln im Projekt führten. Die Funktionen Entwicklung und Professionalisierung spielten hier also die wichtigste Rolle.

Probleme ergaben sich durch die spezifische Konstellation der Selbstevaluation von Projektgruppen, deren Mitglieder aus unterschiedlichen Schulen stammten. Sie führte zu einer als zu langfristig empfundenen Rhythmisierung der Arbeit in Form der halbjährlichen Workshops.

Als wichtige Bedingung erfolgreicher Selbstevaluation wurde von den Befragten die Zusammenstellung der Selbstevaluationsgruppe genannt. Hierbei ist demnach auf die Prädisposition der Beteiligten zu achten und darauf, dass die Gruppe die Chance erhält, im Sinne einer *learning community* gemeinsame Werte und Normen zu etablieren. Auf Ebene der Organisation und der Rahmenbedingungen wurde als besonders wichtig betont, die notwendigen Ressourcen zu berücksichtigen.

5. Ausblick

Altrichter übertitelte noch 1999 einen Artikel mit der Frage „Selbstevaluation – Alle reden davon, wer macht sie?“. Angesichts der inzwischen vorliegenden Fülle an Praxisbeispielen und -ansätzen aus verschiedenen pädagogischen Feldern scheint sich diese Frage in dieser extremen Form heute nicht mehr zu stellen. Selbstevaluation hat sich als ein Ansatz der Evaluation etabliert, der nicht nur gefordert, sondern auch umgesetzt wird.

Eher problematisch schlägt sich dieser Trend derzeit in einer großen terminologischen und v.a. konzeptionellen Vielstimmigkeit nieder. Die scheinbare Einfachheit des Begriffs Selbstevaluation, die aber doch erhebliche Spielräume für konkrete Umsetzungen lässt, mag dafür ein Grund sein. Ein weiterer ist in der starken Fachgebundenheit vieler Ansätze zu sehen: Selbstevaluation wird meist im Kontext eines einzelnen pädagogischen Handlungsfelds (Schule, Hochschule etc.) behandelt. Obwohl diese Konzentration auf ein einzelnes Praxisfeld durchaus sinnvoll ist, besteht dennoch die Gefahr, dass zu oft „das Rad neu erfunden“ wird und dass Potentiale für Synergieeffekte ungenutzt bleiben. Hier ist daher zu wünschen, dass Selbstevaluation zunehmend stärker aus transdisziplinärer Perspektive diskutiert wird.

Bezüglich des Forschungsstands drängt sich die Frage nach der Effektivität von Selbstevaluation und nach ihrer Leistungsfähigkeit im Vergleich zur „klassischen“ Fremdevaluation auf. Selbstevaluation fordert einen hohen Aufwand von allen Beteiligten, der zusätzlich zur praktischen Arbeit geleistet werden muss. Angesichts dieser Tatsache bleibt zu zeigen, dass tatsächlich die mit einer Selbstevaluation verbundenen Erwartungen erfüllt werden können und unter welchen Bedingungen das der Fall ist. Zur Frage erfolgreicher Selbstevaluation liegen derzeit nur wenige systematische Untersuchungen vor, die über Falldarstellungen hinausgehen (Ausnahmen, z. B. bei Buhren, Killus & Müller, 2000). Hierin ist momentan die wichtigste Aufgabe für zukünftige Forschungsbemühungen zu sehen.

Literatur

- Abs, H. J., Raether, W., Tippelt, R. & Vögele, E. (2000). *Evaluation der Lehre – ein Beitrag zur Qualitätssicherung* (Besser Lehren. Praxisorientierte Anregungen und Hilfen für Lehrende in Hochschule und Weiterbildung, Heft 8). Weinheim: Beltz.
- Alkin, M. C. (1985). *A guide for evaluation decision makers*. London: Sage.
- Alkin, M. C., Daillak R. & White P. (1979). *Using evaluations. Does evaluation make a difference?*. London: Sage.
- Almer, A., Schratz, M., Pendl, R. & Rieger, M. (1998). Qualität durch Kooperation. *Erziehung und Unterricht*, 7-8, 621-627.
- Altrichter, H. (1999a). Evaluation als Alltäglichkeit, als Profession und als Interaktion. In J. Thonhauser & J.-L. Patry (Hrsg.), *Evaluation im Bildungsbereich* (S. 103-120). Innsbruck: StudienVerlag.
- Altrichter, H. (1999b). Selbstevaluation. Alle reden davon, wer macht sie? In E. Rösner (Hrsg.), *Schulentwicklung und Schulqualität. Beiträge zur Bildungsforschung und Schulentwicklung* (Band 8, S. 259-281). Dortmund: IFS.
- Beywl, W. (1991). Entwicklung und Perspektiven praxiszentrierter Evaluation. *Sozialwissenschaften und Berufspraxis*, 14, 265-279.
- Beywl, W. & Henze, B. (1999). Praxisbegleitende Trainings in Selbstevaluation. Ein Einstieg in die Qualitätsentwicklung von unten. *PÄD Forum*, 3, 211-218.
- Beywl, W. & Taut, S. (2000). Standards: Aktuelle Strategie zur Qualitätsentwicklung in der Evaluation. *Vierteljahrshefte zur Wirtschaftsforschung*, 3, 358-370.
- Brödel, R. (1999). Pädagogische Qualitätssicherung als reflexive Kompetenzentwicklung. In F. von Küchler & K. Meisel (Hrsg.), *Qualitätssicherung in der Weiterbildung. Band 1: Auf dem Weg zu Qualitätsmaßstäben* (S. 193-205). Frankfurt/Main: Deutsches Institut für Erwachsenenbildung.
- Buhren, C. G. (1995). Entfaltung der Lernkultur durch Organisationsentwicklung. In H. G. Holtappels (Hrsg.), *Entwicklung der Schulkultur – Ansätze und Wege schulischer Erneuerung* (S. 200-221). Neuwied: Luchterhand.
- Buhren, C. G. (1997). Selbstevaluation mit Qualitätsindikatoren – ein Praxismodell. *Journal für Schulentwicklung*, 3, 70-74.
- Buhren, C. G., Killus, D. & Müller, S. (1998). Selbstevaluation von Schule – und wie Lehrerinnen und Lehrer sie sehen. *Jahrbuch der Schulentwicklung*, 10, 235-269.
- Buhren, C. G., Killus, D. & Müller, S. (1999). *Wege und Methoden der Selbstevaluation. Ein praktischer Leitfaden für Schulen*. Dortmund: IFS.

- Buhren, C.-G., Killus, D. & Müller, S. (2000). Implementation und Wirkung von Selbstevaluation in Schulen. In H.-G. Rolff, W. Bos, K. Klemm, H. Pfeiffer & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven* (Bd. 11, S. 327-364). Weinheim: Juventa.
- Burkard, C. (1995). *Selbstevaluation – ein Beitrag zur Qualitätsentwicklung von Einzelschulen?* (S. 1-42). Bönen: Verlag für Schule und Weiterbildung.
- Burkard, C. (1999). Dos and Don'ts der Selbstevaluation. In E. Radnitzky & M. Schratz (Hrsg.), *Der Blick in den Spiegel: Texte zur Praxis von Selbstevaluation und Schulentwicklung* (S. 203-225). München: Studienverlag.
- Burkard, C. & Eikenbusch, G. (1998). Das Schulprogramm intern evaluieren. In E. Risse (Hrsg.), *Schulprogramm. Entwicklung und Evaluation* (S. 267-283). Neuwied: Luchterhand.
- Campbell, D. T. (1969). Reforms as experiments. *American Psychologist*, 24, 409-429.
- Chelimsky, E. (1997). The coming transformations in evaluation. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation for the 21st century. A handbook* (pp. 1-26). Thousand Oaks: Sage.
- Cooksy, L. J. (1999). The Meta-Evaluand: The evaluation of project TEAMS. *American Journal of Evaluation*, 20, 123-136.
- Cousins, J. B. & Whitmore, E. (1998). Framing participatory evaluation. In E. Whitmore (Eds.), *Understanding and practicing participatory approaches to evaluation* (New Directions in Evaluation, Vol. 80, pp. 5-22). San Francisco: Jossey Bass.
- Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C., Phillips, D. C., Walker, D. F. & Weiner, S. S. (1980). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.
- Darling-Hammond, L. & Wise, A. E. (1992). Teacher professionalism. In M. C. Alkin (Eds.), *Encyclopedia of educational research* (6th ed., pp. 1359-1366). New York: Macmillan.
- Deutsche Gesellschaft für Evaluation [DeGEval] (2002). *Standards für Evaluation*. Köln: Geschäftsstelle DeGEval. Verfügbar unter <http://www.degeval.de/standards/standards.htm> [1.12.2003].
- Dubs, R. (1998). *Qualitätsmanagement für Schulen*. St. Gallen: Institut für Wirtschaftspädagogik der Universität St. Gallen.
- Eikenbusch, G. (1997). Schulinterne Evaluation. Ein Weg zur gemeinsamen Schulentwicklung. *Pädagogik*, 49/5, 6-9.
- Eisenlohr, K. S. (1998). Angeleitete Selbstevaluation im Team am Beispiel eines allgemeinen sozialen Dienstes. In Heiner, M. (Hrsg.), *Experimen-*

- tierende Evaluation. Ansätze zur Entwicklung lernender Organisationen* (S. 229-264). Weinheim: Juventa.
- ERS Standard Committee. (1982). Evaluation research society standards for program evaluation. *New directions for program evaluation: Standards for evaluation practice*, 15, 7-19.
- Fetterman, D. M. (1996). Empowerment evaluation: An introduction to theory and practice. In D. M. Fetterman, S. Kaftarian & A. Wandersman (Eds.), *Empowerment evaluation* (pp. 3-46). Thousand Oaks: Sage.
- Fetterman, D. M. (2000). Steps of Empowerment Evaluation: From California to Cape Town. In D. L. Stufflebeam, G. F. Madaus & T. Kellaghan (Eds.), *Evaluation models – viewpoints on educational and human services evaluation* (pp. 395-408). Boston: Kluwer Academic Publishers.
- Fischer-Bluhm, K. (2000). Qualitätsentwicklung als Antwort auf die Individualisierung im Bildungsbereich – am Beispiel der Hochschulen. In Forum Bildung (Hrsg.), *Erster Kongress des Forum Bildung am 14. und 15. Juli 2000 in Berlin* (S. 680-693). Bonn: Arbeitsstab Forum Bildung.
- Franz, H.-W. (1999). *Integriertes Qualitätsmanagement (IQM) in der Weiterbildung. EFQM und DIN ISO 9001 – Modell, Instrumente, Fallstudie*. Bielefeld: Bertelsmann.
- Frey, K. (Hrsg.) (1975). *Curriculum-Handbuch. Band 2*. München: Piper.
- Guba, E. G. & Lincoln, Y. S. (1989). *Fourth generation evaluation*. London: Sage.
- Halfar, B. & Lehnerer C. (1997). Benchmarking statt Zertifikate. *Grundlagen der Weiterbildung*, 8(3), 134-136.
- Heinen-Tenrich, J. (1999). Niedersächsische Volkshochschulen erproben Qualitätsentwicklung durch Selbstevaluation. In F. von Küchler & K. Meisel (Hrsg.), *Qualitätssicherung in der Weiterbildung. Band 1: Auf dem Weg zu Qualitätsmaßstäben* (S. 114-126). Frankfurt/Main: Deutsches Institut für Erwachsenenbildung.
- Heiner, M. (1988). Von der forschungsorientierten zur praxisorientierten Selbstevaluation. Entwurf eines Konzeptes. In M. Heiner (Hrsg.), *Selbstevaluation in der sozialen Arbeit. Fallbeispiele zur Dokumentation und Reflexion beruflichen Handelns* (S. 7-40). Freiburg: Lambertus.
- Heiner, M. (Hrsg.) (1994). *Selbstevaluation als Qualifizierung in der Sozialen Arbeit*. Freiburg: Lambertus.
- Heiner, M. (Hrsg.) (1996). *Selbstevaluation als Qualifizierung in der Sozialen Arbeit. Fallstudien aus der Praxis*. Freiburg: Lambertus.
- Heiner, M. (1998). *Experimentierende Evaluation. Ansätze zur Entwicklung lernender Organisationen*. Weinheim: Juventa.

- Henninger, M. (2000). Evaluation – Diagnose oder Therapie? In C. Harteis, H. Heid & S. Kraft (Hrsg.), *Kompendium Weiterbildung – Aspekte und Perspektiven betrieblicher Personal- und Organisationsentwicklung* (S. 249-260). Opladen: Leske + Budrich.
- Höning, J. (1988). Strukturierte Verlaufsnotizen als Mittel der Evaluation von Kurzzeit Gruppentherapien. In Heiner, M. (Hrsg.), *Selbstevaluation in der sozialen Arbeit* (S. 140-151). Freiburg: Lambertus.
- Joint committee on standards for educational evaluation. (1988). *The personnel evaluation standards. How to assess systems for evaluating educators*. Newbury Park: Sage.
- Joint committee on standards for educational evaluation. (1994). *The program evaluation standards. How to assess evaluations of educational programs*. Thousand Oaks: Sage.
- Joint committee on standards for educational evaluation. (2003). *The student evaluation standards. How to improve evaluations for students*. Thousand Oaks: Corwin Press.
- Klafki, W. (1998). Schulqualität – Schulprogramm – Selbstevaluation der Kollegien. *Erziehung und Unterricht*, 7/8, 568-582.
- König, J. (1998). „Wie gut sind wir eigentlich?“ Kleiner Praxisleitfaden zur Selbstevaluation in der Sozialen Arbeit. *Verhaltenstherapie und psychosoziale Praxis*, 30, 181-200.
- König, J. (2000). *Einführung in die Selbstevaluation*. Freiburg: Lambertus.
- Kraus, W. (1991). Qualitative Evaluationsforschung. In U. Flick, E. von Kardorff, H. Keupp et al. (Hrsg.), *Handbuch qualitativer Sozialforschung*. München: PVU.
- Krüger, T. (1999). „Freiwillige Selbstkontrolle von Weiterbildungsqualität.“ Das Prüfsiegel des Vereins „Weiterbildung Hamburg e.V.“. In F. von Küchler & K. Meisel (Hrsg.), *Qualitätssicherung in der Weiterbildung. Band 1: Auf dem Weg zu Qualitätsmaßstäben* (S. 102-113). Frankfurt/Main: Deutsches Institut für Erwachsenenbildung.
- Küchler, von F. (1999). Einleitung. In F. von Küchler & K. Meisel (Hrsg.), *Qualitätssicherung in der Weiterbildung. Band 1: Auf dem Weg zu Qualitätsmaßstäben* (S. 7-17). Frankfurt/Main: Deutsches Institut für Erwachsenenbildung.
- Legge, K. (1984). *Evaluating planned organizational change*. London: Academic Press.
- Liebald, C. (1998). Leitfaden für Selbstevaluation und Qualitätssicherung. In *Materialien zur Qualitätssicherung in der Kinder- und Jugendhilfe* (Bd. 19). Bonn: Bundesministerium für Familie, Senioren, Familie, Frauen und Jugend.

- Lüders, M. (1998). Die Bedeutung der Professionalisierung des Lehrers für die Qualitätssicherung von Schule. In H. Ackermann & J. Wissinger (Hrsg.), *Schulqualität managen. Von der Verwaltung der Schule zur Entwicklung von Schulqualität* (S. 77-87). Neuwied: Luchterhand.
- MacBeath, J., Schratz, M., Meuret, D. & Jakobsen, L. (2000). *Self-evaluation in European schools: A story of change*. London: Routledge Palmer.
- Madaus G. F. & Stufflebeam D. L. (2000). Program evaluation: A historical overview. In D. L. Stufflebeam, G. F. Madaus & T. Kellaghan (Eds.), *Evaluation models – viewpoints on educational and human services evaluation* (pp. 3-18). Boston: Kluwer.
- Madaus, G. F. & Kellaghan, T. (2000). Models, metaphors and definitions in evaluation. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models – viewpoints on educational and human services evaluation* (pp. 19-31). Boston: Kluwer.
- Mandl, H., Hense, J. & Kruppa, K. (2002). Die Neuordnung der BLK-Modellversuchsförderung am Beispiel des Programms "Systematische Einbeziehung von Medien, Informations- und Kommunikationstechnologie in Lehr- und Lernprozesse" (SEMIK). *Recht der Jugend und des Bildungswesens, Zeitschrift für Schule, Berufsbildung und Jugenderziehung*, 3/2002, 327-338.
- Mandl, H., Hense, J. & Kruppa, K. (2003). Der Beitrag der neuen Medien zur Schaffung einer neuen Lernkultur: Beispiele aus dem BLK-Programm SEMIK. In W. Vollstädt (Hrsg.), *Zur Zukunft der Lehr- und Lernmedien in der Schule. Eine Delphi-Studie in der Diskussion* (S. 85-102). Opladen: Leske + Budrich.
- Mandl, H., Reinmann-Rothmeier, G. & Gräsel, C. (1998). *Gutachten zur Vorbereitung des Programms "Systematische Einbeziehung von Medien, Informations- und Kommunikationstechnologien in Lehr- und Lernprozesse" (Materialien zur Bildungsplanung und Forschungsförderung, Heft 66)* Bonn: Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung.
- Meinhold, M. & Radatz, I. (1988). Selbstbefragungen, Belastungsanalysen und Verlaufsübersichten als Mittel der Dokumentation und (Selbst-)Evaluation in einem Familiendienst-Projekt. In Heiner, M. (Hrsg.), *Selbstevaluation in der sozialen Arbeit* (S. 194-206). Freiburg: Lambertus.
- Meisel, K. (1999). Dialogische Qualitätsentwicklung im Feld – Erfahrungen und Auswertungen eines bundesweiten Projekts in der öffentlichen Erwachsenenbildung. In F. von Küchler & K. Meisel (Hrsg.), *Qualitätssicherung in der Weiterbildung. Band 2: Auf dem Weg zu besserer Praxis* (S. 234-254). Frankfurt/Main: Deutsches Institut für Erwachsenenbildung.

- Müller-Böling, D. (1997). Evaluationen zur Rechenschaftslegung oder Qualitätsverbesserung? Eine Bestandsaufnahme der Evaluation an deutschen Hochschulen. In Altrichter, H., Schratz, M. & Pechar, H. (Hrsg.), *Hochschulen auf dem Prüfstand. Was bringt Evaluation für die Entwicklung von Universitäten und Fachhochschulen* (S. 88-107). Innsbruck: StudienVerlag.
- Müller-Kohlenberg, H. & Beywl, W. (2003). Standards der Selbstevaluation. *Zeitschrift für Evaluation*, 2, 79-93
- Nevo, D. (2001). School Evaluation: Internal or External? *Studies in Educational Evaluation*, 27, 95-106.
- Patton, M. Q. (1997). *Utilization-focused evaluation. The new century text*. Thousand Oaks: Sage.
- Patton, M. Q. (1998). Die Entdeckung des Prozeßnutzens – Erwünschtes und unerwünschtes Lernen durch Evaluation. In Heiner, M. (Hrsg.), *Experimentierende Evaluation: Ansätze zur Entwicklung lernender Organisationen* (S. 55-66). Weinheim, München: Juventa.
- Pechar, H. (1997). Leistungstransparenz oder Wünschelrute? Über das Ranking von Hochschulen in den USA und im deutschsprachigen Raum. In H. Altrichter et al. (Hrsg.), *Hochschulen auf dem Prüfstand. Was bringt Evaluation für die Entwicklung von Universitäten und Fachhochschulen?* (S. 157-178). Innsbruck: StudienVerlag.
- Posch, P. & Altrichter, H. (1998). Schulen am Weg zu Schulprogramm und Qualitätsevaluation. *Erziehung und Unterricht*, 7-8, 544-557.
- Prenzel, M., Mandl, H. & Reinmann-Rothmeier, G. (1997). Ziele und Aufgaben der Erwachsenenbildung. In F. E. Weinert & H. Mandl (Hrsg.), *Psychologie der Erwachsenenbildung, Enzyklopädie der Psychologie, D/1/4* (S. 1-44). Göttingen: Hogrefe.
- Radnitzky, E. & Schratz, M. (1999). *Der Blick in den Spiegel. Texte zur Praxis von Selbstevaluation und Schulentwicklung*. Innsbruck: Studien-Verlag.
- Reinmann-Rothmeier, G. & Mandl, H. (2001). Unterrichten und Lernumgebungen gestalten. In B. Weidenmann & A. Krapp (Hrsg.), *Pädagogische Psychologie. Ein Lehrbuch* (S. 603-648). Weinheim: Beltz.
- Reisert, R. & Carstensen, D. (1998). *Praxis der internen und externen Evaluation. Handbuch zum Verfahren*. Hannover: Hochschul-Informationssystem.
- Richter, R. (1997). Qualitätsevaluation an niederländischen Universitäten und Fachhochschulen. In H. Altrichter, M. Schratz & H. Pechar (Hrsg.), *Hochschulen auf dem Prüfstand. Was bringt Evaluation für die Entwicklung von Universitäten und Fachhochschulen?* (S. 108-122). Innsbruck: StudienVerlag.

- Rolff, H.-G. (1998). Evaluation von Schulentwicklung als Bestandteil eines neuen Leitungs-Verständnisses. In H. Ackermann & J. Wissinger (Hrsg.), *Schulqualität managen. Von der Verwaltung der Schule zur Entwicklung von Schulqualität* (S. 222-238). Neuwied: Luchterhand.
- Schratz, M. (1997). Hochschullehre unter der Lupe. Auf dem Weg zur professionellen Selbstevaluation. In H. Altrichter, M. Schratz & H. Pechar (Hrsg.), *Hochschulen auf dem Prüfstand. Was bringt Evaluation für die Entwicklung von Universitäten und Fachhochschulen?* (S. 261-288). Innsbruck: Studienverlag.
- Scriven, M. (1999). The nature of evaluation part i: relation to psychology. *Practical Assessment, Research & Evaluation*, 6(11). Verfügbar unter <http://PAREonline.net/getvn.asp?v=6&n=11> [1.12.2003].
- Sechrest, L. & Figueredo, A. J. (1993). Program Evaluation. *Annual Review of Psychology*, 44, 645-674.
- Shadish, W. R., Cook, T. D. & Leviton, L. C. (1991). *Foundations of program evaluation. Theories of practice*. Newbury Park: SAGE.
- Simons, R. J. (1992). Lernen, selbstständig zu lernen – ein Rahmenmodell. In H. Mandl & H. F. Friedrich (Hrsg.), *Lern- und Denkstrategien. Analyse und Intervention* (S. 251-254). Göttingen: Hogrefe.
- Specht, W. (1998). Selbstevaluation der Schule und Qualitätssicherung des Bildungswesens im europäischen Kontext; Bericht über einen internationalen Workshop. *Erziehung und Unterricht*, 7-8, 558-567.
- Spiegel, H. (1993). *Aus Erfahrung lernen. Qualifizierung durch Selbstevaluation*. Münster: Votum.
- Spiegel, H. von (1994). Selbstevaluation als Mittel beruflicher Qualifizierung. In Heiner, M. (Hrsg.), *Selbstevaluation als Qualifizierung in der sozialen Arbeit. Fallstudien aus der Praxis* (S. 11-55). Freiburg im Breisgau: Lambertus.
- Stake, R. E. (1975). *Evaluating the arts in education: A responsive approach*. Columbus, Ohio: Merrill.
- Stangel-Meseke, M. & Wottawa, H. (1993). Evaluation. In A. Schorr (Hrsg.), *Handwörterbuch der angewandten Psychologie* (S. 212-217). Bonn: Dt.-Psychologen-Verlag.
- Stark, G. (2000). *Qualitätssicherung in der beruflichen Weiterbildung durch Anwendungsorientierung und Partizipation; Ergebnisse aus einem Modellversuch*. Bielefeld: Bertelsmann.
- Strittmatter, A. (1997). Mythen und Machbares in der Qualitätsevaluation. Erkenntnisse aus Pilotprojekten und Ländern mit längerer Erfahrung. *Journal für Schulentwicklung*, 3, 22-29.

- Strittmatter, A. (2000). *Worauf bei der Selbstevaluation zu achten ist. 10 Lehren aus 5 Jahren FQS – ein Forschungsbericht*. Sursee: o. V.
- Stufflebeam, D. L. (2001). *Evaluation models* (New directions for evaluation, No. 89). San Francisco: Jossey-Bass.
- Suchman, E. A. (1967). *Evaluative research. Principles and practices in public service and social action programs*. New York: Russel Sage Foundation.
- Suchman, E. A. (1970). Action for what? A critique of evaluative research. In R. O'Toole (Hrsg.), *The organization, management, and tactics of social research*. Cambridge, Mass.: Schenkman.
- Tripodi, T. (1983). *Evaluative research for social workers*. Englewood Cliffs, NJ: Prentice Hall.
- Tulodziecki, G. (1998). Entwicklung von Medienkompetenz als Erziehungs- und Bildungsaufgabe. *Pädagogische Rundschau*, 52, 693-709.
- Verbund norddeutscher Universitäten (Hrsg.) (2001). *Evaluation von Studium und Lehre im Fach Psychologie 1999/2000* (Verbund-Materialien, Band 9). Geschäftsstelle Verbund norddeutscher Universitäten: Hamburg. Verfügbar unter <http://www.uni-nordverbund.de/html/3pub/publi2.html> [01.12.2003].
- Weiss, C. H. (1972). Utilization of evaluation: Toward comparative study. In C. H. Weiss (Ed.), *Evaluating action programs: Readings in social action and education* (pp. 318-326). Boston: Allyn and Bacon.
- Whitmore, E. (Hrsg.) (1998). *Understanding and practicing participatory approaches to evaluation* (New Directions in Evaluation, Vol. 80). San Francisco: Jossey Bass.
- Wildavsky, A. (1972). The self-evaluating organization. *Public Administration Review*, 32, 509-520.
- Will, H. & Blickhan, C. (1987). Evaluation als Intervention. In H. W., A. Winteler & A. Krapp (Hrsg.), *Evaluation in der beruflichen Aus- und Weiterbildung. Konzepte und Strategien* (S. 43-59). Heidelberg: Sauer.
- Will, H. & Krapp A. (1983). Bericht über das Symposium "Methoden und Ergebnisse praxisnaher Evaluation". In R. Olechowski (Hrsg.), *Der Beitrag der empirischen Erziehungswissenschaften zur Praxisverbesserung von Schule, Unterricht und Erziehung. Bericht über die 32. Tagung der Arbeitsgruppe für Empirisch-Pädagogische Forschung (AEPF) vom 20.-21. Sept. 1982 in Wien* (S. 115-120). Braunschweig: Deutsche Gesellschaft für Erziehungswissenschaft / Arbeitsgruppe Empirische Pädagogische Forschung.
- Will, H., Winteler, A. & Krapp, A. (1987). Von der Erfolgskontrolle zur Evaluation. In H. Will, A. Winteler & A. Krapp (Hrsg.), *Evaluation in der beruflichen Aus- und Weiterbildung* (S. 11-42). Heidelberg: Sauer.

- Wittmann, W. W. (1985). Evaluationsforschung: Aufgaben, Probleme und Anwendungen. In D. Albert, K. Pawlik, K.-H. Stapf & W. Stroebe (Hrsg.), *Lehr- und Forschungstexte Psychologie*. Berlin: Springer.
- Wottawa, H. (2001). Evaluation. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (S. 649-674). Weinheim: Beltz.
- Wulf, C. (1975). Funktionen und Paradigmen der Evaluation. In K. Frey (Hrsg.), *Curriculum-Handbuch* (S. 580-600). München: Piper.
- Wunder, H. (1995). ISO 9000 – Entwicklung des Qualitätsmanagements und Vorteile ganzheitlichen Qualitätsmanagements. In J. E. Feuchthofen & E. Severing (Hrsg.), *Qualitätsmanagement und Qualitätssicherung in der Weiterbildung* (S. 127-137). Neuwied: Luchterhand.
- Wuppertaler Kreis. (1996). *Qualitätsmanagement in der Weiterbildung nach DIN EN ISO 9000 ff. Ein Leitfaden*. Köln: Deutscher Wirtschaftsdienst.