
Originalarbeiten

Lars Balzer, Andreas Frey & Peter Nenniger

Was ist und wie funktioniert Evaluation?

Die Beurteilung und Beschreibung von Bildungsmaßnahmen ist besonders im Rahmen der Qualitätssicherungs- und Qualitätssteigerungsdebatte seit Anfang der 80er Jahre ein Feld von großem Interesse. Seit dieser Zeit wurde eine Vielzahl von Modellen und Instrumenten für den Markt entwickelt, erprobt und validiert, um über abgesicherte Vorgehensweisen zu guten Evaluationsergebnissen zu gelangen.

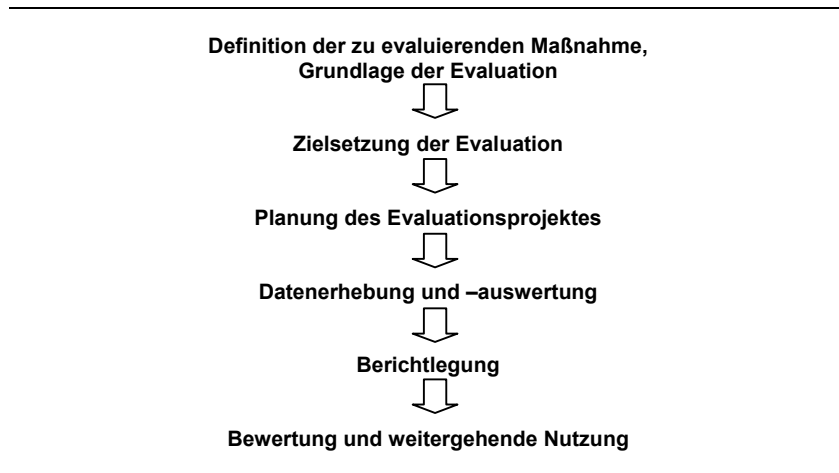
Die Frage, welche grundlegenden Merkmale ein Evaluationsprojekt aufzeigen sollte, erhält im Rahmen der heutigen Forschung eine noch größere Bedeutung, da die Evaluationsvorgehensweise selbst zum Gegenstand der Kritik geworden ist (siehe für den Bereich der Kindergartenforschung: TIETZE, 1999).

Aktuelle Evaluationsbegriffe sind vielschichtig. Beispielsweise definieren BORTZ und DÖRING (1995) Evaluation als "Überprüfung der Wirksamkeit einer sozialen Intervention (...) mit den Mitteln der empirischen Forschung" (S.614), während WOTTAWA und THIERAU (1998) eher auf allgemeine Kennzeichen wissenschaftlicher Evaluation eingehen. "Evaluation dient als Planungs- und Entscheidungshilfe und hat somit etwas mit der Bewertung von Handlungsalternativen zu tun (...). Evaluation ist ziel- und zweckorientiert" und sollte "dem aktuellen Stand wissenschaftlicher Techniken und Forschungsmethoden angepasst sein" (S.14). Deshalb erscheint es in der Folge notwendig, den Evaluationsprozess näher zu untersuchen und die wichtigsten Evaluationsbegriffe zu erläutern, sie zu strukturieren und mit Beispielen aus der gegenwärtigen Praxis zu versehen.

EVALUATIONSPROZESS

In einem ersten Schritt soll der Evaluationsprozess einer Maßnahme näher untersucht werden. Es gibt zwar viele verschiedene Formen, Methoden und Modelle von Evaluationen für viele verschiedene Anwendungsbereiche, aber einige strukturelle Gemeinsamkeiten von empirischen Evaluationsstudien lassen sich dennoch aufzeigen. Abbildung 1 sowie die nächsten Unterkapitel liefern in Anlehnung an WOTTAWA und THIERAU (1998) und DLUGOSCH und WOTTAWA (1994) einen Überblick über den Evaluationsprozess.

Abbildung 1:

Ablauf eines Evaluationsprojektes

(in Anlehnung an WOTTAWA und THIERAU, 1998)

Definition der zu evaluierenden Maßnahme, Grundlage der Evaluation

Noch bevor man sich erste Gedanken über eine Evaluation machen kann, muss das zu evaluierende Evaluationsobjekt genau ausgewählt und definiert worden sein. Ist das Objekt klar beschrieben, sollte sich die Notwendigkeit einer Evaluation von selbst ergeben. Folgende Punkte mögen als Beispiele für eine Argumentationsgrundlage für die Durchführung einer methodisch fundierten Evaluation dienen:

Eine Evaluation gibt bzw. ermöglicht u.a.:

- Entscheidungshilfen über die Einführung, Beibehaltung, Auswertung, Abschaffung oder Kürzung einer Intervention,
- Planungshilfen für eine zukünftige Maßnahme,
- Durchsetzungshilfen für eine geplante Maßnahme,
- eine Entwicklung und Optimierung eines Programms,
- Hinweise auf Mängel und Bedürfnisse,
- einen Vergleich zwischen alternativen Interventionen bzw. Teilmaßnahmen,
- eine Verantwortungsdelegation bei vorhandener großer Unsicherheit,
- eine Erfolgskontrolle,
- die Analyse der Akzeptanz einer Maßnahme,
- einen zusätzlichen Anreiz zur Reflexion über Bestandteile der Intervention.

Schwierigkeiten, die immer wieder auftreten, sind z.B.:

- bestimmte Erwartungen an das Ergebnis,
- ein fehlendes Problembewusstsein bei den Entscheidungsträgern,
- ein Mangel an methodischen Kenntnissen des Evaluierenden, was die Durchführung von Evaluationsmaßnahmen erschwert,
- die Scheu vor den Kosten.

Ist die Notwendigkeit einer Evaluation erkannt, muss man bei der Beschreibung der Intervention darauf achten, eine klare Definition der Ziele, die durch die Intervention erreicht werden sollen, herbeizuführen. Dazu zählen auch die zu Grunde liegenden Theorien und Konstrukte sowie eine Vorstellung über die Wir-

kungsweise der Maßnahme. Es sollte ferner klar sein, welche Kriterien für die Zielerreichung relevant und wie diese zu messen sind. Auch wenn in der Praxis nicht jedem Problem adäquat begegnet werden kann, muss man sich der ‚Fallstricke‘ bewusst sein, um Ergebnisse angemessen interpretieren zu können.

Zielsetzung der Evaluation

Ist eine Intervention klar umschrieben, muss das Evaluationsziel, d.h. die eigentliche Fragestellung, klar abgesteckt werden. Denn "leider sind die Fragen der Zielsetzung in vielen Evaluationsprojekten die größte Schwachstelle" (WOTTAWA & THIERAU, 1998, S.83).

Zumeist wird es sich bei Evaluationen um Auftragsarbeiten handeln, was bedeutet, dass der Wissenschaftler das Ziel seiner Arbeit nicht vollständig selbst bestimmen kann. Bei der Zielsetzung der Evaluation werden Vorstellungen, Wünsche und Interessen des Auftraggebers eine zentrale Rolle spielen.

Es gilt in der Folge, das *Evaluationsobjekt* zu konkretisieren. Verschiedenste Gruppen von Evaluationsobjekten sind nach WOTTAWA und THIERAU (1998) dabei denkbar, z.B.:

- „Personen
- Umwelt-/Umgebungsfaktoren
- Produkte
- Techniken/Methoden
- Zielvorgaben
- Programme
- Projekte
- Systeme/Strukturen
- Forschungsergebnisse/Evaluationsstudien“ (S.59)

Hierbei kommt dem Evaluator eine besonders wichtige Rolle zu: Am Anfang häufig sehr einfach gestellte Fragen (z.B.: Ist die Weiterbildungsmaßnahme erfolgreich?) müssen erst weiter präzisiert werden, damit sie im weiteren Verlauf der Evaluation überprüfbar sind (in Bezug auf obiges Beispiel ist das eine Klärung der Frage, welche Komponenten eine gute Weiterbildungsmaßnahme ausmachen und wie diese Merkmale operationalisiert und messbar gemacht werden können). Diese Präzisierung soll sich sowohl auf theoretische als auch auf praxis-

nahe Konstrukte beziehen. Allerdings ist eine genaue theoretische Fundierung nicht vom Auftraggeber zu verlangen und fällt in den Aufgabenbereich des Evaluators. Meistens muss dem Auftraggeber aber erst der Blick geschärft werden, welche Komponenten der Evaluationsobjekte für die Evaluation relevant sein könnten. Besonders im Bereich des Objektes 'Personen' werden im Sinne der *Zielgruppenfindung* oft nicht alle wirklich Betroffenen vom Auftraggeber erkannt.

Parallel sollte man schon zu diesem Zeitpunkt des Projektes *potenzielle Nebenwirkungen* der Maßnahme und der Evaluation mit ins Kalkül ziehen und mögliche Zielkonflikte inhaltlicher Art wie auch zwischen beteiligten Institutionen und Personen klären.

Auch kann es passieren, dass dem Evaluator mancher Gedankengang und manches versteckte Ziel des Auftraggebers verborgen bleibt.

Planung des Evaluationsprojektes

Stehen die Evaluationsziele zumindest im groben Rahmen fest, kann in die konkrete Projektplanung eingestiegen werden. Wenn möglich sollten in der Planungsphase alle am Evaluationsprozess Beteiligten einbezogen werden. Dabei sind verschiedene Punkte zu beachten:

Zunächst gilt es, Informationen zu Projektdetails und gegebenen *Rahmenbedingungen* vom Auftraggeber zu erfragen oder aus sonstigen Quellen zu beziehen. Dazu gehören Anforderungen, Auflagen, Besonderheiten, Merkmale und Restriktionen von Seiten des Auftraggebers und den von der Evaluation Betroffenen ebenso wie eine endgültige Einigung über die Zielsetzungen. Dies ist nicht immer einfach, besonders wenn man bedenkt, dass der Auftraggeber für viele dieser Dinge erst sensibel gemacht werden muss bzw. nicht bereit oder in der Lage ist, alle notwendigen Informationen zur Verfügung zu stellen. Als allgemeine Informationsquellen können zunächst schriftliche Unterlagen dienen, doch darf auf das persönliche Gespräch mit dem Auftraggeber nicht verzichtet werden.

Es muss ferner eine *Zeit- und Kostenabschätzung* erfolgen, auf die eine solide *Personal- und Materialplanung* aufbauen kann.

Daran anschließend sind *Evaluationsstrategien* zu entwickeln, ein adäquates *Evaluationsmodell* (Abbildung 2, S. 400, liefert einen Überblick) und *Design* auszuwählen und mögliche Fehlerquellen schon im Vorfeld zu analysieren und zu berücksichtigen.

Sind die Bewertungskriterien sorgfältig ausgewählt, ist es ebenfalls Aufgabe während der Planungsphase, sich mit dem Auftraggeber, sofern er bis ins Detail an der Planung beteiligt sein soll, Gedanken über bestehende oder zu entwickelnde *Datenerhebungsverfahren* sowie die *Auswertungsstrategien* zu machen. Das fängt bei einfachen Entscheidungen an, wie der Frage, ob qualitativ und/oder quantitativ vorgegangen werden soll und geht gegebenenfalls bis zur Detailab-sprache über die Art und Weise eines durchzuführenden Interviews. Generell muss aber die Frage geklärt werden, ob für ein bestimmtes Evaluationsvorhaben schon auf bewährte Instrumentarien zurückgegriffen werden kann oder ob neue Instrumente entwickelt werden müssen. Ersteres hat den Vorteil, dass bei der späteren Interpretation der Ergebnisse auf Vergleichsdaten zurückgegriffen werden kann, zweiteres erlaubt es hingegen, auf spezifische Gegebenheiten des Evaluationsprojektes einzugehen. In diesem Fall ist es aber notwendig, sowohl für die Entwicklung als auch für die Vortestung der neuen Instrumentarien genügend Zeit einzuräumen.

Neben dem Zeitbudget für die Datenerhebung und Datenauswertung ist ebenso genügend Raum für die Berichtlegung und eine mögliche weitergehende Publikation der Ergebnisse einzuplanen, denn speziell Letzteres wird allzu oft unterschätzt und deshalb in der Folge vernachlässigt.

Auf der Grundlage dieser Informationen kann ein *Arbeitsplan* erstellt werden, der auch Vorstellungen über mögliche *Kontroll- und Feedbackmöglichkeiten* berücksichtigen sollte.

Datenerhebung und -auswertung

Daraufhin kann in die aktive Phase der Datenerhebung eingestiegen werden. Dabei ist ein guter Arbeitsplan natürlich eine große Hilfe. Dennoch kann man insbesondere bei großen Evaluationsvorhaben nicht davon ausgehen, dass immer alles wie geplant verläuft. Einige unerwartete Störfälle werden immer auftreten, auf die es richtig zu reagieren gilt. Auch ist bei der Datenerhebung besonders dann auf die Qualität zu achten, wenn man als Evaluator nicht selbst für die Erhebung zuständig ist. Gegebenenfalls sind bei der Auswertung besondere Kontrollen einzubauen.

Sind Daten mit Hilfe eines aus dem unüberschaubaren Kanon der Datenerhebungsmethoden ausgewählten Verfahrens gewonnen, steht die Auswertung der Daten an. Hierbei muss generell beachtet werden, dass einerseits je nach Fragestellung und Hypothese, andererseits je nach Datenqualität die passenden statistischen Methoden zur Auswertung des Datensatzes herangezogen werden. Auf eine

ausgewogene Bilanz zwischen Wissenschaftlichkeit und Nutzenbezogenheit der Ergebnisse sowie auf die Einhaltung des Datenschutzgesetzes ist zu achten.

Berichtlegung

Ein besonderes Kapitel im Verlaufe eines Evaluationsprojektes ist die Berichtlegung, da hier all zu oft die Diskrepanz zwischen Wissenschaft und Wirtschaft offenkundig wird. Den Auftraggeber interessieren oft ausschließlich die konkret nutzenbezogenen Ergebnisse, die ihm Entscheidungshilfen liefern können. Deswegen ist auf eine zielbezogene Informationsverdichtung zu achten. Oft ist für den Auftraggeber nur eine einseitige Kurzdarstellung entscheidungsrelevant, obwohl sich für die Forschung ein großer Umfang weiterer Aspekte aufdrängen würde. Auch ist zu bedenken, dass man es zumeist nicht mit in den Evaluationsterminologien ausgebildeten Personen zu tun hat, so dass auf eine verständliche Sprache zu achten ist. Gängig ist auch eine mündliche Abschlusspräsentation der Ergebnisse, für die aber die gleichen Einschränkungen gelten wie für ihre schriftlich formulierten Pendanten. All zu oft endet hier die Evaluation.

Bewertung und weitergehende Nutzung

Doch Evaluation sollte mehr sein als nur Datenerhebung und Auswertung. Auch für die Übertragung der Ergebnisse in die Praxis sollte sich der Evaluator mitverantwortlich zeigen. Inhaltliche Schlussfolgerungen, mögliche Übertragungen auf ähnliche Maßnahmen und Evaluationen und deren Optimierung, eine Endbewertung und ein Vergleich mit anderen Evaluationen in verwandten Gebieten sollten sich anschließen.

Nach welchen Kriterien lässt sich aber eine Evaluation bewerten? "Der Nutzen durch die Ergebnisse ist im Wesentlichen eine Verbesserung der Entscheidungen, um deretwillen das Projekt durchgeführt wurde" (WOTTAWA & THIERAU, 1998, S.165). Dabei werden die Ergebnisse mit den zu Projektbeginn definierten Zielen in Beziehung gesetzt, woran - wie schon bei der Planung - am besten alle Betroffenen beteiligt werden sollten. Als problematisch kann sich erweisen, dass Nutzen stets eine subjektive Größe ist und für alle Beteiligten jeweils etwas anderes bedeuten kann. Unabhängig davon handelt es sich bei der Nutzenbestimmung insbesondere um die Effektivität der Maßnahme sowie um deren Effizienz, d.h. dem Kosten-Nutzen-Verhältnis der Intervention.

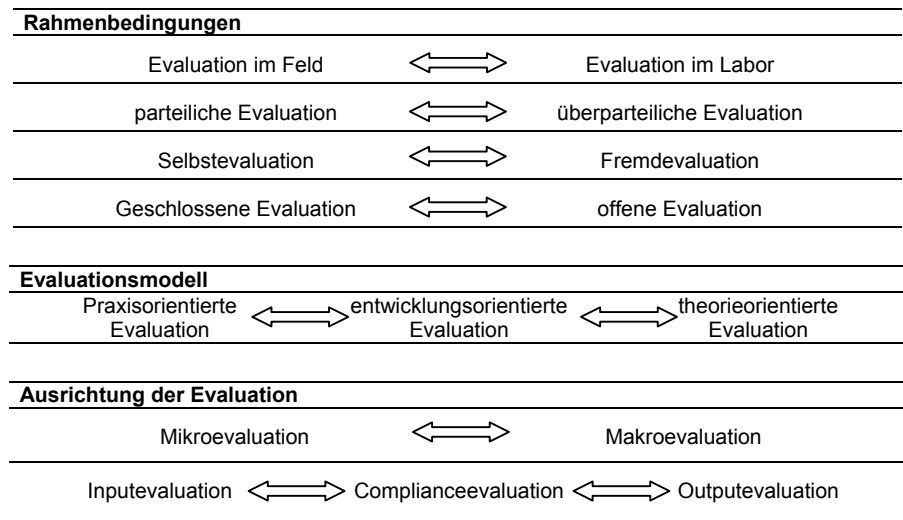
Nur in zweiter Linie von den Ergebnissen abhängig sind die durch die Evaluation entstandenen Aufwendungen finanzieller, materieller und zeitlicher Art. In

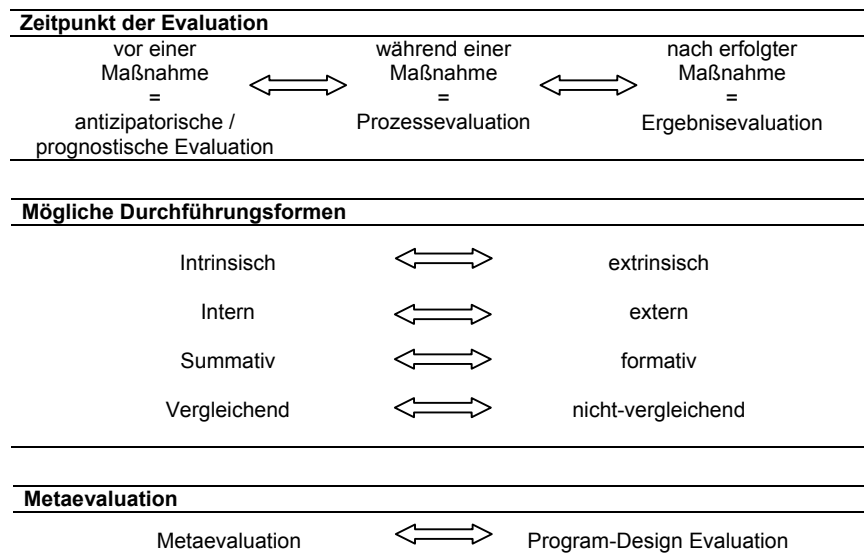
welchem Verhältnis diese zu den erzielten Ergebnissen stehen, ist aber auch ein Bewertungskriterium für die Evaluation selbst.

EVALUATIONSBEGRIFFE

Nach der Beschreibung des Evaluationsprozesses wird im Folgenden eine Systematisierung und Erläuterung der verschiedenen Begrifflichkeiten in Anlehnung an BALZER (1998), WOTTAWA und THIERAU (1998) sowie DLUGOSCH und WOTTAWA (1994) vorgenommen. Abbildung 2 enthält einen Überblick der wichtigsten Evaluationsbegriffe, die allerdings nicht immer eindeutig voneinander abzugrenzen sind.

Abbildung 2:
Übersicht über die wichtigsten Evaluationsbegriffe





(in Anlehnung an BALZER, 1998; DLUGOSCH & WOTTAWA, 1994)

Rahmenbedingungen

Jede Evaluation findet zunächst unter bestimmten Rahmenbedingungen finanzieller wie organisatorischer Art statt. Zumeist wird die Evaluation in der realen Umgebung des Geschehens als *Feldevaluation* ablaufen.

FREY (1999) untersuchte die Effektivität von Ausbildungsmaßnahmen am Beispiel der Erzieherinnenausbildung an Fachschulen für Sozialwesen in Rheinland-Pfalz. Hauptgegenstand der Untersuchung war, mit Hilfe von Einschätzungen von Schülerinnen und Schülern sowie Lehrkräften vor Ort mehrperspektivische Informationen über die Qualität der Ausbildung zu erhalten.

Manche Fragestellungen (Vergleiche bestimmter Bedingungen) eignen sich dabei besser für die *Laborevaluation*, da hier die Rahmenbedingungen sowie die relevanten Variablen strenger kontrolliert werden können. Dabei ist der konkrete Ort der Untersuchung unerheblich. Vielmehr kommt es auf das Ausmaß der Kontrolle an, die der Forscher über die Untersuchungsbedingungen und das Setting ausübt. Die Übergänge zwischen Feld- und Laborevaluation sind aber oft fließend.

BANNERT (1996) variierte in ihrer Untersuchung zum Lernerfolg in EDV-Schulungen quasiexperimentell die Art der Unterrichtsmethode (traditionelles Vorgehen versus Selbstlernmaterial) und die Art der Benutzerschnittstelle (Standardschnittstelle versus einer speziellen Training-Wheels-Schnittstelle, durch die momentan nicht benötigte Funktionalitäten des zu erlernenden Programmes für den Lerner gesperrt wurden).

Um eine *parteiliche Evaluation* handelt es sich, wenn der Verantwortliche eines Projektes selbst in eine Evaluation involviert ist und ihm ein Interesse an einem bestimmten Verlauf oder an bestimmten Ergebnissen der Evaluation unterstellt werden kann.

DLUGOSCH und KRIEGER (1999) berichten von einer Entwicklung, Durchführung und Evaluation eines ganzheitlichen und interdisziplinären ambulanten Gesundheitsförderungsprogramms, welches durch die Techniker Krankenkasse finanziert worden ist und der ein Eigeninteresse an positiven Ergebnissen zugeschrieben werden konnte.

Sind die Verantwortlichen der Evaluation neutrale Instanzen und sind sie vom Ergebnis der Evaluation nicht betroffen, so spricht man von *überparteilicher Evaluation*.

JÄGER, ARBINGER, BANNERT und LISSMANN (1993) berichten von dem Projekt „Computerunterstütztes Lernen an allgemeinbildenden Schulen“, welches im Auftrag des Ministeriums für Bildung und Kultur des Landes Rheinland-Pfalz vom Zentrum für empirische pädagogische Forschung der Universität Koblenz-Landau durchgeführt worden ist.

Wird eine Evaluation von einer oder mehreren Personen durchgeführt, die im zu evaluierenden Projekt selbst als Betroffene eingebunden und somit Bestandteil der Evaluation sind, so spricht man von *Selbstevaluation*. Insiderkenntnisse der Organisation und Wissen über betroffene Personen wie gegebene Strukturen können dabei bei der Planung und Durchführung durchaus von Vorteil sein.

BALZER (1998) berichtet von der Evaluation des Weiterbildungsangebotes eines öffentlichen Trägers im EDV-Bereich, bei dem das Evaluationskonzept von einem selbst an diesem Institut lehrenden Dozenten entwickelt, erprobt und durchgeführt worden ist.

Wird eine Evaluation von einer externen Person oder Organisation durchgeführt, die in keinem Zusammenhang mit der Planung und Organisation der zu evaluierenden Maßnahme steht, so spricht man von *Fremdevaluation*. Dieser Art der Evaluation wird oft mehr Neutralität und eine höhere Fachkompetenz unterstellt, bringt aber auch höhere Kosten mit sich.

WOLF, BECKER und CONRAD (1999) führten im Auftrag des Bundesfamilienministeriums eine externe empirische Evaluation durch. Gegenstand dieser Evaluation war das Projekt

„Kindersituationen“, welches durch eine andere Gruppe von Wissenschaftlern durchgeführt wurde. Ziel dieser Fremdevaluation war es, die Effektivität des pädagogischen Situationsansatzes zu überprüfen.

Die Verwendung der Ergebnisse ist ebenfalls zu klären. Sollen die Ergebnisse einer Evaluation der Öffentlichkeit zugänglich gemacht werden, spricht man von einer *offenen Evaluation*.

Die Deutsche Forschungsgemeinschaft (DFG) hat ein großes Interesse daran, dass Ergebnisse einer Evaluation einer breiten Öffentlichkeit zugänglich gemacht werden. Aus diesem Grund können ohne vorherige Absprachen mit der DFG Ergebnisse oder Teilaspekte der Evaluation publiziert werden. (Im DFG-Projekt „Kindergartenalltag im deutsch-deutschen Vergleich“ wurden unter dem Teilaspekt „Erzieherinnenaus- und -fortbildung“ verschiedene Publikationswege gewählt, um die Ergebnisse interessierten Fachpersonen und Eltern von Kindergartenkindern zugänglich zu machen (FREY, SCHULZ & WOLF, 1995, FREY & WOLF, 1996, FREY, 1997). Zudem wurde für die Öffentlichkeit ein Tageszeitungsbeitrag veröffentlicht).

Um eine *geschlossene Evaluation* handelt es sich, wenn die Ergebnisse nur einem bestimmten Personen- bzw. Instanzenkreis zugänglich gemacht werden. Dabei kann es sich z.B. um eine interne Berichtslegung gegenüber den Auftraggebern handeln, die auf Grund ihres geschlossenen Charakters eine besonders objektive, Schwachstellen aufzeigende Analyse ermöglicht. Wissenschaftliche Ergebnisse sind prinzipiell zu veröffentlichen, doch können z.B. Konkurrenzverhältnisse zwischen Firmen oder Sicherheitsgesichtspunkte im öffentlichen Bereich diesem Ansinnen zuwider laufen.

Evaluationsmodell

Die Zielrichtung eines Evaluationsprojektes kann im Wesentlichen unter drei Perspektiven erfolgen.

Bei einer *praxisorientierten Evaluation* rückt ein konkretes Projekt in den Mittelpunkt des Interesses. Die Ergebnisse dieses Projektes werden zum Evaluationsziel.

Die Untersuchung von FISCHER und KRIEGER (1998) zielte z.B. auf die Bewertung von durchgeführten Maßnahmen im Bereich der Suchtprävention ab.

Dienen Ergebnisse einer Evaluation als Grundlage von Verbesserungen ähnlicher Maßnahmen in der Zukunft, so spricht man von einer *entwicklungsorientierten Evaluation*. Hierbei wird es wichtig, eine Vergleichsgrundlage für die Gestaltung späterer Maßnahmen zu schaffen.

DLUGOSCH und KRIEGER (1999) berichten, dass es in ihrem Evaluationsprojekt mehrere Modellphasen gibt, in denen Gesundheitsfördermaßnahmen selbst über die Ergebnisse optimiert und weiterentwickelt werden.

Werden hingegen die Güte und Bewährung verschiedener Theorien anhand der Ergebnisse der Evaluation bewertet, so spricht man von *theorieorientierter Evaluation*.

SIEPMANN (1993) analysiert Forschungsergebnisse zu Einflussgrößen für den Erfolg von EDV-Schulungsmaßnahmen und integriert sie in ein Gesamtmodell. EDV-Dozenten und Teilnehmer von EDV-Schulungen werden daraufhin zu Komponenten des Modells befragt und um Angaben über ihre Einstellungen bzw. Erwartungen bezüglich möglicher Einflussgrößen auf Lernerfolg in EDV-Schulungsmaßnahmen gebeten.

Ausrichtung der Evaluation

Der Schwerpunkt der Fragestellung bestimmt das Evaluationsvorhaben. Interessieren nur Details wie z.B. die konkrete Gestaltung von Unterrichtsmaterialien, spricht man von *Mikroevaluation*.

JÄGER, ARBINGER, BANNERT und LISSMANN (1993) berichten von einer Mikroevaluation, in der sie die Computerarbeit von Jugendlichen auf Video aufnehmen ließen und bestimmte Details der Arbeit auswerteten.

Bilden hingegen komplexe Systeme den Evaluationsgegenstand, so handelt es sich um eine *Makroevaluation*.

KRIEGER und DLUGOSCH (1998) berichten von einer Makroevaluation, die ein siebentägiges Schulungskompaktangebot zur Förderung eines gesundheitsdienlichen Lebensstils zum Gegenstand hatte, wobei ein Schwerpunkt auf psychologischen Themengebieten (z.B. Stress/bewältigung, Entspannung, Kommunikation) lag. Die Makroevaluation bezog sich dabei insbesondere auf Veränderungen im Gesundheitsverhalten und -zustand.

Eine Unterscheidung durch die Fragestellung ganz anderer Art liegt dann vor, wenn Eigenschaften bestimmter Ressourcen oder Materialien für eine bestimmte Maßnahme untersucht werden. Hier handelt es sich um eine *Inputevaluation*.

GÖTZ, KLEBSATTEL und REICHERT (1993) legten eine Studie vor, die u.a. der Frage nachging, welchen Einfluss die Qualität der Kursunterlagen auf den Erfolg von EDV-Qualifizierungsmaßnahmen hat.

Wenn man andererseits den Ertrag eines Systems zum Ziel seiner Evaluation macht, so spricht man von *Outputevaluation*.

In der Untersuchung von HARDT, ZAIB, KLEINBECK und METZ-GÖCKEL (1996) wurde u.a. untersucht, inwieweit im Kontext des kaufmännischen Unterrichts die Unterrichtsziele (Vermittlung von Wissen) bei den Schülerinnen und Schülern erreicht werden konnten.

In der *Complianceevaluation* wird das Verhalten von Betroffenen während einer Maßnahme oder die Einhaltung von Interventionsbedingungen evaluiert.

OSER, PATRY, ELSÄSSER, SARASIN und WAGNER (1997) untersuchten auf der Grundlage eines theoretischen Basismodells zur Planung und Strukturierung von Unterricht per Videoanalyse die Handlungs- und Lernschritte sowie das Verhalten von Schülerinnen und Schülern in diesem Lernkontext.

Zeitpunkt der Evaluation

Untersucht und bewertet man ein Vorhaben vor seinem eigentlichen Beginn hinsichtlich der Planung, des Aufbaus und möglicher Ergebnisse, so spricht man von *prognostischer Evaluation*. Es geht dann zum einen darum, auf Fehler und Fallen im Vorfeld aufmerksam zu machen, um das eigentliche Vorhaben dann noch besser durchführen zu können; zum anderen bietet sich hier die Gelegenheit, zwischen mehreren Möglichkeiten zu wählen, wenn der reale Test beider Situationen nicht möglich ist, wenn mit erheblichen negativen Konsequenzen für die Betroffenen zu rechnen wäre oder wenn die eingeleitete Maßnahme kaum oder nur mit erheblichen finanziellen Aufwendungen rückgängig gemacht werden könnte.

INGENKAMP (1991) entwickelte ein Beurteilungsinstrument für Erzieherinnen zur Diagnose der Schulfähigkeit (BEDS). Mit dem BEDS werden Vorschulkinder innerhalb der Dimensionen ‚sprachlich kognitive Leistungen‘, ‚Sozial- und Arbeitsverhalten‘ sowie ‚allgemeine Schulfähigkeit‘ getestet. Die prognostische Aussage über den allgemeinen Schulerfolg eines Kindes wird durch die Integration der Ergebnisse aller drei Skalen getroffen.

Solch ein Vorgehen erfordert eine andere Herangehensweise als eine nachträgliche Bewertung. Sorgfältige Literatarbeit sowie Sekundäranalysen gewinnen hier an Bedeutung. Eine Erweiterung findet dieser Evaluationsbegriff, wenn die Entscheidung im Vorfeld tatsächlich in reduzierter Form empirisch von statten gehen kann. Hier spricht man von *antizipatorischer Evaluation*.

Bevor über die schweizerische Reform der kaufmännischen Grundausbildung durch die Eidgenössische Vernehmlassung entschieden wird, wurde im Sommer 1999 ein Pilotversuch mit

dem Ziel gestartet, die ausprobierten Neuerungen in Bezug auf ihre Machbarkeit und Wirksamkeit hin zu überprüfen (AMBÜHL, 1999). Erst nach der Evaluation wird entschieden, welche Neuerungen in der folgenden Reform zum Tragen kommen.

Bei der *Prozessevaluation* (auch *dynamische Evaluation*) wird ein Vorhaben während der Durchführung kontinuierlich begleitet. Gegebenenfalls können Ergebnisse der Evaluation auch für Kurskorrekturen eingesetzt werden.

ABHAUER und HANEWINKEL (1998) berichten von einem Projekt, bei dem Lehrerinnen und Lehrer nach jedem der zwanzig Unterrichtsabschnitte der Unterrichtseinheit „Fit und stark fürs Leben“ auf verschiedenen Dimensionen das Unterrichtsgeschehen beurteilen sollten. Anregungen flossen sofort in die Neugestaltung der Materialien ein.

Bei der *Ergebnisevaluation* rückt der Ertrag einer Maßnahme bzw. die Wirksamkeit einer Intervention im Nachhinein in den Mittelpunkt des Interesses. Dieser Begriff wird synonym mit dem Begriff Outputevaluation verwendet.

Mögliche Durchführungsformen

Von einer *extrinsischen Evaluation* spricht man, wenn eine Maßnahme auf Grund ihrer Auswirkung bei den Teilnehmern dieser Maßnahme evaluiert wird.

HOLLING (1999) überprüfte die längerfristigen Effekte eines Stressbewältigungstrainings bei den teilnehmenden Polizeibeamten des Landes Nordrhein-Westfalen.

Bei einer *intrinsischen Evaluation* spielen personenbezogene Auswirkungen keine Rolle. Anstelle dessen stehen theoriebezogene Konzepte im Mittelpunkt des Interesses.

HENDRIX (1999) untersuchte, ob die im Berufsbild des Datenverarbeitungskaufmanns vorgesehenen Aufgabenbereiche mit den gegenwärtig üblichen Arbeitsabläufen noch übereinstimmen. Aus den Ergebnissen wurden vier neue Ausbildungsberufe entwickelt.

Von *interner Evaluation* ist die Rede, wenn Tätigkeiten die Maßnahme betreffend und die Evaluation selbst in einer Hand liegen. Dieser Begriff wird oft synonym als Selbstevaluation bezeichnet.

Eine *externe Evaluation* trennt diese Aufgaben personell. Diese prinzipiell wünschenswerte Trennung scheidet allerdings oft am finanziellen Rahmen solcher Projekte. Dieser Begriff wird oft synonym mit Fremdevaluation verwendet.

Aus der *summativen Evaluation* ergibt sich eine globale, zusammenfassende Bewertung zum Ende der Maßnahme,

FREY (1999) untersuchte am Ende der schulischen Ausbildung die Effektivität der Bildungsmaßnahme mit Hilfe von Einschätzungen von Schülerinnen und Schülern sowie Lehrerinnen und Lehrern. Am Ende wurde eine globale und zusammenfassende Endbewertung der Ausbildungsmaßnahme vorgenommen.

in der *formativen Evaluation* wird eine Intervention begleitet und mit ihren Zwischenergebnissen wird gegebenenfalls auch Einfluss auf die Intervention genommen.

JÄGER und BEHRENS (1994) entwickelten ein Modell für die künftige Lehrerbildung in Rheinland-Pfalz. Hierzu wurde neben empirischen Untersuchungen auch eine zweistufige Delphi-Befragung durchgeführt, die in eine Expertentagung einmündete. Dieser methodische Zugang kann als eine Evaluation über einen mehrstufigen Prozess betrachtet werden.

Werden mehrere Alternativen miteinander verglichen, spricht man von *vergleichender Evaluation*.

Im Rahmen der Überprüfung des Systems TOAM wurde der Frage nachgegangen, ob der Einsatz dieses Systems erfolgreicher ist als wenn lediglich herkömmlicher Unterricht zur Vermittlung von Mathematikkenntnissen durchgeführt wird (vgl. PROJEKTGRUPPE TOAM, 1992).

Werden hingegen normative Standards oder zuvor definierte Ziele als Bewertungsmaßstab zu Grunde gelegt, handelt es sich um eine *nicht-vergleichende Evaluation*.

Innerhalb der Evaluation des Projektes ‚Orientierungsschule‘ wurde von BÄTZ, OSER und KLAGHOFER (1995) die Orientierungsschule nicht mit ihrer Vorgängerin verglichen, sondern an den Zielen gemessen, die im Reformprojekt definiert waren.

Metaevaluation

Eine Evaluation auf höherem Abstraktionsniveau findet statt, wenn man Ergebnisse einer oder mehrerer Evaluationsstudien zum Zwecke der Verallgemeinerung der Ergebnisse zu einer Globalaussage zusammenführt. Dies wird *Metaevaluation* genannt.

ROSENSTIL (1992) führt die Ergebnisse von ca. 160 Studien über das Lehr- und Lernverhalten von Männern und Frauen aus dem Bereich von Bildungs- und Erziehungsinstitutionen zusammen, um Aussagen über mögliche Verhaltensunterschiede zwischen Männern und Frauen zu machen.

Eine Erfahrungszusammenfassung und Überprüfung zum Zwecke einer verbesserten zukünftigen Durchführung hinsichtlich Planung, Gestaltung, verwendeter Methoden, Zielsetzungen etc. wird als *Program-Design Evaluation* bezeichnet.

BEHRENS (1999) führte eine theoriegeleitete Analyse von Videos durch, welche im Zusammenhang mit Business-TV entstanden und gesendet wurden. Hierbei ging es um die Frage, ob und in welchem Ausmaß die Konstruktion der Lerneinheiten optimiert werden müssen. Das hierbei entwickelte Qualitätsraster dient als Ausgangspunkt einer empirischen Analyse. In der Konsequenz werden Planung, Methoden und Zielsetzungen einer Analyse und Bewertung unterzogen (vgl. JÄGER, BEHRENS & LEIDIG, 1999).

ZUSAMMENFASSUNG

Wie aus der vorhergehenden systematischen Darstellung hervorgeht, sind die Evaluationsprozesse ebenso wie die Evaluationsbegriffe von den verschiedenen Anwendungsbereichen abhängig und lassen sich somit unterschiedlichen Formen, Methoden und Modellen von Evaluationsvorhaben zuordnen.

Wichtig erscheint in diesem Zusammenhang allerdings, dass alle Evaluationsarten innerhalb ihres Evaluationsprozesses einige strukturelle Gemeinsamkeiten aufzeigen sollten:

1. Definition der zu evaluierenden Maßnahme, Grundlage der Evaluation
2. Zielsetzung der Evaluation
3. Planung des Evaluationsprojektes
4. Datenerhebung und Datenauswertung
5. Berichtlegung
6. Bewertung und weitergehende Nutzung

Diese Stadien des Evaluationsprozesses bilden den ‚roter Faden‘ des Evaluationsprojektes und können den zeitlichen und inhaltlichen Arbeitsstand der Evaluationsmaßnahme aufzeigen. Die Güte einer Evaluation kann demnach nicht nur an ihren Produkten, sondern auch an der internen Struktur ihres Prozesses gemessen werden.

Des Weiteren ist die Vielzahl der dargestellten Evaluationsbegriffe und der dazu aufgeführten praktischen Beispiele auch als Hinweis darauf zu sehen, dass Anwendungsbereiche sehr vielseitig sein können, selbst wenn nur ein Kontext wie „Aus- und Fortbildung“ betrachtet wird und manche Begriffe sich nur schwerlich voneinander trennen lassen. Zudem wird deutlich, dass es Evaluati-

onsarten gibt, bei denen in manchen Bereichen eine eindeutige Zuordnung zu bestimmten Evaluationsbegriffen kaum möglich ist. Aus diesem Grund erscheint es wichtig, dass für jedes Evaluationsprojekt von Seiten der Evaluatoren eine begriffliche und eine damit verbundene inhaltliche Einbettung des Gesamtvorhabens vorgenommen wird, damit sowohl für interne (beispielsweise die am Evaluationsprozess Beteiligten) als auch für externe Personen (beispielsweise die Leserschaft von Fachzeitschriften) deutlicher wird, um welche Art von Evaluation es sich handelt und warum eine bestimmte Evaluationsart einem bestimmten Begriff und Inhaltsbereich zugeordnet werden kann.

LITERATUR

- AMBÜHL, G. (1999). Evaluation der Pilotversuche. In Bundesamt für Berufsbildung und Technologie (Hrsg.), *Reform der kaufmännischen Grundausbildung* (S. 3). Bern: BBT.
- ABHAUER, M. & HANEWINKEL, R. (1998). Prozessevaluation eines Lebenskompetenztrainings in der Grundschule. *Empirische Pädagogik*, 12(4), 327 - 345.
- BÄTZ, U., OSER, F. & KLAGHOFER, R. (1995). *Evaluation Orientierungsschule Basel: 1. Schuljahr, 1994/1995*. Basel: Erziehungsdepartement.
- BALZER, L. (1998). *Der Erfolg von EDV-Seminaren: Eine Evaluation von Computerkursen einer Volkshochschule*. Unveröffentlichte Diplomarbeit. Universität Koblenz-Landau.
- BANNERT, M. (1996). *Gestaltung und Evaluation von EDV-Schulungsmaßnahmen. Eine empirische Studie zur Effektivität und Akzeptanz*. Landau: Verlag Empirische Pädagogik.
- BEHRENS, U. (1999). *Teleteaching is easy!? Pädagogisch-psychologische Qualitätskriterien und Methoden der Qualitätskontrolle im Rahmen von Teleteaching-Projekten*. Landau: Verlag Empirische Pädagogik.
- BORTZ, J. & DÖRING, N. (1995). *Forschungsmethoden und Evaluation*. Berlin: Springer.
- DLUGOSCH, G.E. & KRIEGER, W. (1999). "Wege zum Wohlbefinden" - Ein Bericht zur praxisnahen Evaluation eines Gesundheitsförderungsseminars. *Zeitschrift für Gesundheitspsychologie*, 7, 27-43.
- DLUGOSCH, G.E. & WOTTAWA, H. (1994). Evaluation in der Gesundheitspsychologie. In P. SCHWENKMEZGER & L.R. SCHMIDT (Hrsg.), *Lehrbuch der Gesundheitspsychologie* (S. 149-168). Stuttgart: Enke.
- FISCHER, U.CH. & KRIEGER, W. (1998). „Suchtprävention an der Gemeng“. *Entwicklung, Durchführung und Evaluation eines Modells zur gemeindeorientierten Suchtprävention*. Luxembourg: Centre de prévention des toxicomanies.
- FREY, A., SCHULZ, S. & WOLF, B. (1995). Fortbildung von Erzieherinnen im deutsch-deutschen Vergleich. *Zeitschrift klein & groß*, 48, 20-23.
- FREY, A. & WOLF, B. (1996). Aus- und Fortbildungen von Erzieherinnen und ihre pädagogische Arbeit im Kindergarten. In R. MÖLLER, J. ABEL, G. NEUBAUER & K. TREUMANN (Hrsg.), *Kindheit, Familie und Jugend* (S. 35-45). Münster & New York: Waxmann.
- FREY, A. (1997). Konfliktbewältigungsverhalten von Erzieherinnen. In B. DIPPELHOFFER-STIEM & B. WOLF (Hrsg.), *Ökologie des Kindergartens* (S. 136-161). Weinheim und München: Juventa.

- FREY, A. (1999). *Erzieherinnenausbildung gestern - heute - morgen. Konzepte und Modelle zur Ausbildungs-evaluation*. Landau: Verlag Empirische Pädagogik.
- GÖTZ, K., KLEBSATTEL, A. & REICHERT, H.-J. (1993) Qualitätssteigerung in der DV-Qualifizierung. *Empirische Pädagogik*, 7, 149-173.
- HARDT, B., ZAIB, V., KLEINBECK, U. & METZ-GÖCKEL, H. (1996). Unterrichtsziele und Lernmotivation im Kontext kaufmännischen Unterrichts. *Unterrichtswissenschaft*, 3, 235-249.
- HENDRIX, W. (1999). Evaluation des Berufsbildes des Datenverarbeitungskaufmanns. In H. HOLLING & G. GEDIGA (Hrsg.), *Evaluationsforschung* (S. 127-178). Göttingen: Hogrefe.
- HOLLING; H. (1999). Evaluation eines Stressbewältigungstrainings. In H. HOLLING & G. GEDIGA (Hrsg.), *Evaluationsforschung* (S. 35-57). Göttingen: Hogrefe.
- INGENKAMP, K. (1991). *Beurteilungsbogen für Erzieherinnen zur Diagnose der Schulfähigkeit (BETS)*. (2. ergänzte Aufl.). Weinheim: Beltz.
- JÄGER, R.S., ARBINGER, R., BANNERT, M. & LISSMANN, U. (1993). Ergebnisse der wissenschaftlichen Begleitung. In A. RISSBERGER (Hrsg.), *Computerunterstütztes Lernen an allgemeinbildenden Schulen Teil II: Abschlussbericht des Modellversuchs CULAS* (S. 193-251). Mainz: v. Hase und Koehler.
- JÄGER, R.S. & BEHRENS, U. (1994). *Weiterentwicklung der Lehrerbildung*. Mainz: v. Haase & Koehler.
- JÄGER, R.S., BEHRENS, U. & LEIDIG, U. (1999). Evaluationskonzepte für Business TV: Zur Notwendigkeit von Qualitätssicherung. In W. JÄGER (Hrsg.), *Unternehmenskommunikation durch Business TV. Strategien, Technikkonzepte, Praxisbeispiele* (S. 37-53). Wiesbaden: Gabler.
- KRIEGER, W. & DLUGOSCH, G.E. (1998). Die „Gesundheitswoche“ - Eine Kompaktmaßnahme zur Gesundheitsförderung im stationären Bereich. In G. AMANN & R. WIPPLINGER (Hrsg.), *Gesundheitsförderung. Ein multidimensionales Tätigkeitsfeld* (S. 451-470). Tübingen: dgvt-Verlag.
- OSER, F., PATRY, J.-L., ELSÄSSER, T., SARASIN, S. & WAGNER, B. (1997). *Che-reographien unterrichtlichen Lernens. Schlußbericht an den Schweizerischen Nationalfonds zur Förderung der wissenschaftlichen Forschung*. Freiburg/CH: Pädagogisches Institut der Universität.
- PROJEKTGRUPPE TOAM (1992). Erfahrungsbericht der Projektgruppe TOAM. In R.S. JÄGER (Hrsg.), *Computerunterstütztes Lernen an allgemeinbildenden Schulen. Einsatz und Erprobung des TOAM-Systems* (S. 121-127). Mainz: v. Haase & Koehler.

- ROSENSTIL, L. (1992). *Grundlagen der Organisationspsychologie*. Stuttgart: Schäffer-Poeschel Verlag.
- SIEPMANN, B. (1993). *Effektives EDV-Training (Bde. Empirische Schul- und Unterrichtsforschung)*. Frankfurt am Main: Peter Lang Verlag.
- TIETZE, W. (1999). Zur Relevanz einer Externen Empirischen Evaluation. In P. BECKER, S. CONRAD & B. WOLF (Hrsg.), *Kindersituationen im Diskurs* (S. 29-37). Landau: Verlag Empirische Pädagogik.
- WOLF, B., BECKER, P. & CONRAD, S. (1999). *Der Situationsansatz in der Evaluation*. Landau: Verlag Empirische Pädagogik.
- WOTTAWA, H. & THIERAU, H. (1998). *Lehrbuch Evaluation*. (2., vollst. überarb. Aufl.). Bern: Verlag Hans Huber.

BALZER, L., FREY, A. & NENNIGER, P. (1999). Was ist und wie funktioniert Evaluation? *Empirische Pädagogik*, 13(4), 393-413.

Die Beurteilung und Beschreibung von Bildungsmaßnahmen ist besonders in der aktuellen Diskussion um Qualitätssicherung und –steigerung ein Feld von großem Interesse. Es wurde eine Vielzahl von Modellen und Instrumenten für den Markt entwickelt, erprobt und validiert, um über abgesicherte Vorgehensweisen zu guten Evaluationsergebnissen zu gelangen.

Der vorliegende Artikel beschreibt aus diesem Grund zunächst exemplarisch einen klassischen Evaluationsprozeß und erläutert daraufhin mit Hilfe zentraler Evaluationsbegriffe verschiedene Arten von Evaluationen, die mit Beispielen aus der gegenwärtigen Bildungspraxis veranschaulicht werden.

BALZER, L., FREY, A. & NENNIGER, P. (1999). Evaluation: What it is and how it works. *Empirische Pädagogik*, 13(4), 393-413.

In the last decade questions on how to assure and to stimulate quality have also become a prominent issue in education. In this field an important number of theoretical models and practical instruments has been subject to development, experiment and validation. This was done in order to achieve accurate results in evaluation over the secure way of professional and state-of-the-art procedures.

For this reason the present article starts with an evaluation procedure that will be followed by an outline of various types of evaluation in education which are illustrated by the use of real cases of the field.