

Was ist Evaluation?

Meyer, Wolfgang

Preprint / Preprint

Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

SSG Sozialwissenschaften, USB Köln

Empfohlene Zitierung / Suggested Citation:

Meyer, W. (2002). *Was ist Evaluation?* (CEval-Arbeitspapier, 5). Saarbrücken: Universität des Saarlandes, Fak. 05 Empirische Humanwissenschaften, CEval - Centrum für Evaluation. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-113436>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Was ist Evaluation?

Wolfgang Meyer

2002

Meyer, Wolfgang: Was ist Evaluation?
Saarbrücken: Centrum für Evaluation, 2002
(CEval-Arbeitspapiere; 5)

NICHT IM BUCHHANDEL ERHÄLTlich

SCHUTZGEBÜHR: 5 €

BEZUG: Centrum für Evaluation (CEval)
Universität des Saarlandes
Postfach 15 11 50
D-66041 Saarbrücken
info@ceval.de



oder kostenfrei zum Download:
<http://www.ceval.de>

LAYOUT + SATZ: Alexandra Caspari

INHALT

EINLEITUNG.....	2
1. BEGRIFFSBESTIMMUNG.....	2
2 HISTORISCHE ENTWICKLUNG UND INSTITUTIONEN.....	3
3 EVALUATIONSSTANDARDS UND DURCHFÜHRUNGSPRINZIPIEN.....	5
4 AUSBLICK – DIE GRÜNDUNG DES CENTRUMS FÜR EVALUATION	8
LITERATUR	10

EINLEITUNG¹

Die wissenschaftliche Evaluationsforschung gehört als Teilgebiet sozialwissenschaftlicher Forschungen weltweit seit vielen Jahren zu den am stärksten expandierenden Tätigkeitsfeldern von Akademikern. Immer mehr staatliche Institutionen in einer ebenfalls ständig wachsenden Zahl von Ländern greifen zur Vorbereitung, Durchführung und Aufbereitung der Ergebnisse ihrer politischen Programme auf die Unterstützung externer Experten – Evaluatoren – zurück. In einigen Anwendungsbereichen konnten sich Evaluationsroutinen herausbilden, und in manchen Ländern ist teilweise die Durchführung von Wirkungskontrollen sogar gesetzlich vorgeschrieben.

Über den gegenwärtigen Stand der Institutionalisierung von Evaluationen sowie den historischen Verlauf dieser Wissenschaftssparte in den westlichen Demokratien möchte dieser Beitrag informieren. Eine kurze Begriffsklärung ist diesen Ausführungen vorangestellt; die schließlich auch eine Beschreibung internationaler Standards und methodische Hinweise zur Durchführung von Evaluationen enthalten.

Ziel dieses Beitrags ist es, eine kurze Übersicht über den Stand der internationalen Evaluationsforschung und –praxis zu liefern und gleichzeitig das neue Centrum für Evaluation (CEval) der Universität des Saarlandes und seine Forschungsarbeiten vorzustellen.

1. BEGRIFFSBESTIMMUNG

Der Begriff Evaluation stammt von dem lateinischen Wort „valor“, d. h. „Wert“, und der Vorsilbe e / ex, d. h.: aus, und bedeutet übersetzt „Bewertung“, „einen Wert aus etwas ziehen“. In seinem weitesten Wortsinn beschreibt der Begriff den „Prozess der Beurteilung des Wertes eines Produktes, Prozesses oder eines Programmes, was nicht notwendigerweise systematische Verfahren oder datengestützte Beweise zur Untermauerung einer Beurteilung erfordert“². Somit ist z.B. bereits der morgendliche Blick aus dem Fenster zur Begutachtung des Wetters eine „Evaluation“. Obwohl der Begriff „Evaluation“ den meisten Menschen unbekannt sein dürfte, ist der damit beschriebene Vorgang so alt wie die Menschheit selbst³.

Für die Bezeichnung der wissenschaftlichen Herangehensweise wurde das Wort „Evaluation“ gerade wegen seiner vergleichsweise geringen Verbreitung gewählt, um sich von dem alltagssprachlichen Verständnis von „Bewertung“ oder „Beurteilung“ abzuheben. Die allgemein anerkannten sozialwissenschaftlichen Definitionen von „Evaluation“ unterstreichen die explizite Verwendung wissenschaftlicher Forschungsmethoden⁴ und verstehen unter „Evalu-

¹ Dieser Text wurde in lateinischer Sprache in der Zeitschrift VOX Latina publiziert und basiert teilweise auf einem zweisprachigen (deutsch-russisch) Arbeitspapier des Bundesinstituts für Berufliche Bildung (BiBB), welches in gekürzter Fassung in einer russischen Berufsbildungszeitschrift veröffentlicht wurde (vgl. Meyer & Höhns 2002). Gegenüber dieser Vorlage wurden die spezifischen Ausführungen zur beruflichen Bildung weggelassen und dafür zusätzlich Informationen zur aktuellen Entwicklung der Evaluationsforschung an der Universität des Saarlandes eingefügt. Der Autor dankt Frau Gabriela Höhns (BiBB) für die Zustimmung zur Publikation dieser Fassung und Frau Dr. Sigrid Albert (Universität des Saarlandes) für die Übersetzung in die lateinische Sprache.

² Wottawa & Thierau 1990: 9

³ vgl. Stockmann 2000: 11

⁴ vgl. z.B. Bortz & Döring 2002: 102; Hadley & Mitchell 1995: 48; Mertens 2000: 41; Shadish 1994: 352

ationsforschung“ die „*systematische Anwendung sozialwissenschaftlicher Forschungsmethoden zur Beurteilung der Konzeption, Ausgestaltung, Umsetzung und des Nutzens sozialer Interventionsprogramme*“⁵.

Damit unterscheiden sich wissenschaftliche Evaluationen von Alltagsevaluationen ausschließlich durch die Nutzung empirischer Forschungsmethoden und die Einhaltung der hiermit verbundenen wissenschaftlichen Standards.⁶ Diese Differenzierung lässt sich nicht nur für sozialwissenschaftliche Fragestellungen vornehmen: im Unterschied zur eigenen Begutachtung der Wetterlage basiert das Urteil des Meteorologen auf wissenschaftlichen Analysen und dem Einsatz geeigneter standardisierter Messinstrumente. Dies bedeutet nicht zwangsläufig, dass sein Urteil zutreffender als die persönliche Einschätzung sein muss. Durch die wissenschaftliche Fundierung des Urteils werden aber die Bewertungskriterien offengelegt und dadurch die logischen Ableitungen von Bewertungen für Dritte nachvollziehbar. Wissenschaftliche Evaluationen leisten somit einen wesentlichen Beitrag zur Versachlichung von Diskussionen und zur Objektivierung des Urteilsprozesses.

2 HISTORISCHE ENTWICKLUNG UND INSTITUTIONEN

Die Ursprünge wissenschaftlicher Evaluationen lassen sich in den USA bis in das achtzehnte Jahrhundert zurückverfolgen, als die Bundesregierung zum ersten Mal externe Inspektoren mit der Bewertung öffentlicher Programme beauftragte. Zu den damals untersuchten staatlichen Einrichtungen gehörten Waisenhäuser, Schulen, Krankenhäuser und Gefängnisse⁷. Eine erste Expansionsphase von Evaluationen lässt sich in den Vereinigten Staaten für die dreißiger und vierziger Jahre des zwanzigsten Jahrhunderts festhalten. Im Rahmen des „New Deal“ wurden z.B. Reformprogramme zur Verminderung der Arbeitslosigkeit und zur Verbesserung der sozialen Sicherheit von Evaluationsstudien begleitet⁸. Die meisten Autoren bestimmen allerdings erst die frühen sechziger Jahre als Ausgangspunkt professioneller Evaluationen, weil in der Folge umfangreicher staatlicher Sozial-, Bildungs-, Gesundheits-, Ernährungs- und Infrastrukturprogramme nicht nur die Anzahl der Evaluationsstudien sprunghaft anstieg, sondern auch ein Institutionalisierungsprozess einsetzte⁹.

Ein wesentliches Kennzeichen der fortschreitenden Professionalisierung der Evaluationstätigkeiten in den USA war die Gründung einer Reihe einschlägiger Fachzeitschriften, unter denen wohl das „American Journal of Evaluation“ und die „Evaluation Review“ zu den international am meisten rezipierten gehören dürften. 1986 formierte sich durch den Zusammenschluss zweier kleinerer Evaluationsverbände die American Evaluation Association (AEA), die gegenwärtig mehr als 3.000 Mitglieder vertritt. Ziel dieser Organisation ist die Verbesserung von Evaluationsmethoden und ihres praktischen Einsatzes sowie die Förderung wissenschaftlicher Evaluationsforschung in der amerikanischen Öffentlichkeit (zu den Aufgaben und Zielen der AEA siehe die Website der Organisation „www.eval.org“). Durch die Verbrei-

⁵ Rossi, Freeman & Hofmann 1988: 3

⁶ vgl. Stockmann 2000: 12

⁷ vgl. Madaus, Stufflebeam & Scriven 1983

⁸ vgl. Stockmann 2000: 21f.

⁹ vgl. Mertens 2000: 42ff.

tung der Anfang der neunziger Jahre entwickelten Evaluationsstandards des „Joint Committee“ sowie der „Guiding Principles for Evaluators“ hat die AEA einen wesentlichen Beitrag zur Qualitätssicherung von Evaluationen geleistet.

Nach Schätzungen wurden in den USA bereits 1976 etwa 600 Mio. \$ für die Evaluation von sozialen Dienstleistungsprogrammen aufgewendet¹⁰. In vielen Ressorts sind Evaluationen gesetzlich vorgeschrieben und eigene Haushaltsmittel werden dafür bereitgestellt. In den achtziger Jahren wurden so genannte „Sunset-Gesetze“ erlassen, die bei fehlendem Nachweis der Wirksamkeit von Programmen deren automatische Beendigung nach einer festgelegten Zeitspanne vorsahen. 1993 verabschiedete der US-Congress den „Government Performance and Results Act“, welcher die regelmäßige Durchführung von Evaluationen für alle Bundesprogramme gesetzlich verankerte¹¹. Insgesamt ist somit festzuhalten, dass in den USA wissenschaftliche Evaluationen zu einem zentralen Steuerungsinstrument politischen Handelns geworden sind und sich entsprechende Institutionen zur Qualitätssicherung der Evaluationstätigkeit herausgebildet haben.

Im Vergleich zu den USA setzte in Europa eine Professionalisierung der Evaluationsforschung mit zehnjähriger Verspätung Ende der sechziger Jahre ein. Zu den Vorreitern dieser Entwicklung gehörten Schweden, Großbritannien und die Bundesrepublik Deutschland. Ähnlich wie in den USA wurden Evaluationen zunächst hauptsächlich zum Nachweis der Effektivität von Programmen in der Bildungs-, Gesundheits- und Sozialpolitik eingesetzt, um dann Ende der siebziger Jahre mit einer Prioritätsverlagerung auf Kosten-Nutzen-Erwägungen konfrontiert zu werden. Im Unterschied zu den USA stagnierte aber die Entwicklung der Evaluationsforschung in Folge der Einsparungen öffentlicher Haushalte und die noch in den Anfängen befindlichen Institutionalisierungsversuche scheiterten weitgehend. Erst Anfang der 90er Jahre ergab sich für die Evaluationsforschung im Zuge der Reform- und Modernisierungsdebatte staatlicher Institutionen, die unter dem Stichwort „New Public Management“ zuerst in Großbritannien, Holland und den skandinavischen Staaten Konturen gewann, ein neuer Aufschwung¹².

Beginnend in den späten 80er Jahren in Großbritannien werden besonders seit Mitte der 90er Jahre in zunehmend mehr Ländern - und auf europäischer Ebene mit Formierung der European Evaluation Society (EES) – Evaluationsgesellschaften ähnlich der American Evaluation Association gegründet (1997 z.B. die Deutsche Gesellschaft für Evaluation). Gegenwärtig existieren Evaluationsgesellschaften in Belgien, Deutschland, Finnland, Frankreich, Großbritannien, Italien, den Niederlanden, Schweden und der Schweiz¹³. In weiteren Ländern (z.B. in Österreich und Spanien) wurden vergleichbare Initiativen gestartet.

Trotz erheblicher Entwicklungsrückstände lässt sich jedoch auch für Westeuropa feststellen, dass sich die Evaluationsforschung etabliert hat und derzeit weiterhin stark expandiert. Die Bemühungen zur Professionalisierung befinden sich verglichen mit den USA noch in einer Anfangsphase, es lassen sich allerdings in vielen Ländern rasche Fortschritte erkennen. Ohne systematische wissenschaftliche Beobachtungen und Beurteilungen der Wirkungen politischen Handelns kommen moderne Demokratien nicht mehr aus. Ein aktuelles Beispiel für

¹⁰ vgl. Stockmann 2000: 21

¹¹ vgl. Mertens 2000: 49

¹² vgl. Stockmann 2000: 23f.

¹³ vgl. Leeuw 2000: 59f.; als aktueller weltweiter Überblick siehe Meyer 2002

die politischen Konsequenzen wissenschaftlicher Evaluationen stellt die PISA-Studie zur vergleichenden Evaluierung von Schulleistungen in einer Vielzahl von Ländern dar:¹⁴ die Ergebnisse für die Bundesrepublik Deutschland wiesen zum einen aufgrund des schwachen Abschneidens der Schüler auf einen akuten und seit Jahren vernachlässigten Handlungsbedarf hin und zum anderen konnten einige langjährige ideologische Grabenkämpfe der deutschen Bildungspolitik durch Bereitstellung sachlicher Informationen beendet werden.

Generell können in Deutschland eine Reihe von Bereichen aufgezählt werden, für die bereits regelmäßig und standardisiert Evaluationen durchgeführt werden. In der Entwicklungszusammenarbeit gehören wissenschaftliche Evaluationen schon seit mehr als zwanzig Jahren zu den Routinen der Projektdurchführung, wobei sich in letzter Zeit vor allem die Zahl der ex-post Evaluationen stark erhöht hat¹⁵. Die Diskussionen über Formen der Hochschulevaluation sind in einigen Bundesländern bereits soweit fortgeschritten, dass es zur Etablierung standardisierter Evaluationsinstitutionen (z.B. der Zentralen Akkreditierungsstelle ZEVA in Hannover oder des CHE in Gütersloh) gekommen ist. Auch im Arbeitsmarktbereich, in der beruflichen Bildung, bei sozialen Einrichtungen, in der Schulpolitik, im Bereich der Forschungs- und Technologieförderung, in der Umweltpolitik sowie im Rahmen der Verwaltungsmodernisierung sind entsprechende Ansätze erkennbar¹⁶.

3 EVALUATIONSSTANDARDS UND DURCHFÜHRUNGSPRINZIPIEN

Wie werden nun wissenschaftliche Evaluationen durchgeführt? Hierfür wurden im Jahr 1994 von einem eigens eingerichteten „Joint Committee on Standards of Educational Evaluation“ in den USA Standards entwickelt, die in vier Hauptgruppen zusammengefasst wurden¹⁷. Dabei sollen *Nützlichkeitsstandards* sicherstellen, dass sich eine Evaluation an den Informationsbedürfnissen der vorgesehenen Evaluationsnutzern ausrichtet. Zu ihnen gehört z.B. die Offenlegung der eingesetzten Verfahren und Vorgehensweisen gegenüber den Auftraggebern und gegenüber den Evaluierten. Mit *Durchführbarkeitsstandards* soll gewährleistet werden, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst ausgeführt wird. Ein wesentlicher Aspekt ist z.B. die angemessene Berücksichtigung der Interessen und Werte der an einer Evaluation beteiligten Interessengruppen und Personen. Über die Festlegung von *Korrektheitsstandards* wird geregelt, dass eine Evaluation rechtlich und ethisch korrekt durchgeführt sowie dem Wohlergehen der in die Evaluation einbezogenen und durch die Ergebnisse betroffenen Personen gebührende Aufmerksamkeit gewidmet wird. Zu den wichtigsten Regeln ist hier die Offenlegung der Ergebnisse gegenüber allen an der Evaluation beteiligten Personen zu zählen. Mit den *Genauigkeitsstandards* schließlich soll garantiert werden, dass eine Evaluation fachlich angemessene Informationen über den Nutzen des evaluierten Programms hervorbringt und vermittelt. Insbesondere die Verpflichtung auf die

¹⁴ vgl. Baumert et al. 2000, 2001

¹⁵ vgl. Stockmann 1997; 2000a

¹⁶ vgl. als Überblick DeGEval 2001, DIW 2000

¹⁷ ausführlich siehe Joint Committee on Standards of Educational Education 2000; eine Kurzzusammenfassung ist bei Widmer 2000 zu finden

adäquate Nutzung allgemein anerkannter Verfahren zur Informationsgewinnung ist dabei als wichtiger Gesichtspunkt hervorzuheben.

Die Standards wurden mittlerweile von Evaluationsgesellschaften in vielen verschiedenen Ländern (u.a. in Deutschland und der Schweiz) übersetzt und in leicht modifizierter Form übernommen. Aufgrund ihrer weltweiten Verbreitung und der großen Zustimmung unterschiedlichster Akteure können die „Joint Committee“-Standards als allgemein anerkannter Konsens der Evaluationsforschung zur Sicherung der Qualität von Evaluationen bezeichnet werden. Auch auf der Ebene der EU, die ebenfalls Evaluationen in den verschiedensten Politikbereichen durchführt bzw. in Auftrag gibt, wird gegenwärtig an einer Anpassung der DEval-Standards auf Evaluationen und Wirkungsforschungen speziell im europäischen Berufsbildungswesen gearbeitet. In den USA wurden diese Standards von der American Evaluation Association (AEA) durch Leitlinien für Evaluatoren¹⁸ ergänzt. Die „Guiding Principles for Evaluators“ fordern von den Evaluatoren

- die systematische, datenbasierte Untersuchung des zu evaluierenden Gegenstands,
- den Beleg ausreichender Fachkompetenz zur Durchführung von Evaluationsstudien,
- die Gewährleistung eines fairen und integren Evaluationsprozesses,
- den angemessenen Respekt gegenüber den Persönlichkeitsrechten aller Beteiligten sowie
- ein allgemeines Verantwortungsgefühl für die durch Programme beeinflusste öffentliche Wohlfahrt.

Den Auftraggebern wird empfohlen, Aufträge nur an Evaluatoren zu vergeben, die sich diesen Prinzipien verpflichtet haben.

In den europäischen Evaluationsgesellschaften ist es bisher noch nicht gelungen, einen vergleichbaren Konsens zur Formulierung von Leitlinien zu erzielen. Es gibt dabei allerdings keine grundsätzlichen Widerstände gegen die von der AEA vorgelegten und hier kurz vorgestellten Prinzipien.

Das wesentliche Problem einer einheitlichen Festlegung von Leitlinien für Evaluatoren liegt vielmehr in der Vielfalt der bearbeiteten Aufgaben und Themenfeldern begründet. Es würde hier zu weit führen, auf die Vielzahl der bearbeiteten Themenstellungen (zum Beispiel in den bereits oben erwähnten Politikbereichen Gesundheitswesen, Sozialwesen, Bildungswesen und vielen anderen mehr) im Einzelnen einzugehen und ihre Besonderheiten herauszustellen. Die vielfältigen Anforderungen an die Durchführung einer Evaluation hängen beispielsweise ab vom politischen Prozess eines Programms, der sehr grob in drei verschiedene Phasen unterteilt werden kann¹⁹. Während der Programmentwicklung werden zur Untersuchung der Ausgangs- und Rahmenbedingungen *ex-ante Evaluationen* (Machbarkeitsstudien) mit dem Ziel einer Einschätzung möglicher Durchführungsprobleme sowie der Erfolgchancen geplanter Maßnahmen in Auftrag gegeben. Der Charakter dieser prognostisch orientierten Studien unterscheidet sich stark von dem *formativer Evaluationen*, die als Begleitforschung möglichst aktuelle Informationen über den Programmverlauf sammeln und ihre Bewertungen als Entscheidungshilfen für die Steuerung zur Verfügung stellen möchten. Sum-

¹⁸ Shadish et al. 1995

¹⁹ vgl. Stockmann 2000: 13ff.

mative oder *ex-post Evaluationen* schließlich erfassen und beurteilen die langfristigen Programmwirkungen, wobei stärker als bei den anderen Vorgehensweisen die Identifikation kausaler Zusammenhänge im Vordergrund steht.

Als ein Beispiel kann das von Reinhard Stockmann entwickelte Evaluationskonzept angeführt werden, welches derzeit u.a. vom Bundesinstitut für Berufsbildung (BIBB) in Zusammenarbeit mit dem Institut für die Entwicklung der Berufsbildung (IRPO) zur *ex-post* Evaluation von ausgewählten deutsch-russischen Modellprojekten aus dem TRANSFORM-Programm genutzt wird²⁰. Es handelt sich hierbei um ein theoriegeleitetes Konzept mit einem integrativen methodischen Ansatz, dessen Grundlagen wie folgt kurz zu skizzieren sind²¹:

Zur Bestimmung der Analysefelder werden drei unterschiedliche theoretische Modelle zusammengeführt.

1. Durch Nutzung von Erkenntnissen der Lebensverlaufsforschung (das zu evaluierende Projekt wird, grob vergleichbar mit dem menschlichen Lebensverlauf, in eine Entstehungsphase, eine Verlaufsphase und eine Phase der Nachwirkungen aufgeteilt) ergibt sich eine zeitliche Perspektive, die die historische Projektentwicklung und die jeweiligen Anforderungen zueinander in Beziehung setzt und so für die einzelnen Projektphasen eine Bewertung der Entscheidungen zur Projektplanung und -steuerung ermöglicht.

2. Ein organisationstheoretisch gestütztes Wirkungsmodell lenkt den Blick auf die Leistungsfähigkeit der jeweiligen Trägerorganisation eines Projekts (z.B. eine berufliche Lehranstalt, an der eine Neuerung des Berufsbildungssystems modellhaft erprobt wurde) und hilft die Frage zu beantworten, welche Auswirkungen im weitesten Sinn ein Projekt auf die Trägerorganisation hatte, was sich im Zusammenhang mit dem Projekt im unmittelbaren Umfeld veränderte und welchen Einfluss wiederum die Trägerorganisation sowie ihre soziale Umwelt auf den Projektverlauf hatte.

3. Die Leistungsfähigkeit der Trägerorganisation ist auch für die Verbreitung der eingeführten Neuerungen über die Grenzen der Organisation hinaus von entscheidender Bedeutung. In der beruflichen Bildung sollen beispielsweise von Modellprojekten Wirkungen sowohl innerhalb des Berufsbildungssystems (z.B. durch die Verbreitung des Modellversuchs) als auch innerhalb des Erwerbssystems (z.B. durch die Verbesserung der Anpassung von Qualifikationen an die Erfordernisse des Arbeitsmarktes) erzielt werden. Derartige Verbreitung oder, um den Fachterminus zu gebrauchen, Diffusion von Projektwirkungen, ist in dem Evaluationskonzept von Stockmann ebenfalls Gegenstand der Untersuchungen.

Bei der methodischen Vorgehensweise berücksichtigt das Konzept von Stockmann einige Besonderheiten von Evaluationen. So werden durch den partizipativen Ansatz die Fachkenntnisse und Erfahrungen von Evaluationsexperten mit dem Insiderwissen der Projektbeteiligten zusammengeführt. Grundlage ist der offene Austausch von Informationen und die gegenseitige faire Behandlung während der Evaluierung.

Besondere methodische Schwierigkeiten könnten sich durch die Mängel einzelner Datenerhebungsmethoden ergeben, die während einer Evaluation nicht wie im Experiment durch entsprechende Kontrollmaßnahmen ergänzt werden können. Um die Schwächen der einzel-

²⁰ Als weitere Anwendungsbeispiele siehe die Evaluationsstudien zu Berufsbildungsprojekten in China (Stockmann u.a. 2000) oder zu Umweltberatungsprogrammen in Ostdeutschland (Stockmann u.a. 2001).

²¹ Ausführlich hierzu siehe: Stockmann 1997.

nen Verfahren auszugleichen und dabei ihre Stärken möglichst optimal zu nutzen, werden in dem Evaluationskonzept unterschiedliche Datenerhebungstechniken in einem Multimethodenansatz zusammengefasst.

Obwohl auf diesem Weg wesentliche Verbesserungen der Datenqualität erzielt werden können, bleibt ein weiteres Problem davon unberührt: in der Regel werden bei einer Evaluation einige wenige Projekte untersucht, zu denen dann sehr viele Informationen vorliegen, während die Evaluation letztendlich eine Vielzahl unterschiedlicher Forschungsfragen beantworten soll. Letztlich lässt sich nur durch geeignete Vergleiche bestimmen, inwieweit die in einem Einzelfall gefundenen Zusammenhänge rein zufällig entstanden und damit einmalig oder die Folge systematischer und dementsprechend regelmäßig wieder auftretender kausaler Ursachen sind. Die Integration eines komparativen Ansatzes ist dementsprechend zur Identifikation von Problemfeldern und zur Erarbeitung von Lösungsvorschlägen unabdingbar notwendig.

4 AUSBLICK – DIE GRÜNDUNG DES CENTRUMS FÜR EVALUATION

Wie aus dem kurzen historischen Abriss dieses Beitrags deutlich wurde, hat sich die wissenschaftliche Evaluation in allen modernen westlichen Ländern als wichtiges Instrument zur rationalen Steuerung politischen Handelns und zur öffentlichen Legitimierung von Entscheidungen, zur Überprüfung der Wirkungen staatlicher Programme und zur Kontrolle der Effektivität eingesetzter finanzieller und sonstiger Ressourcen bewährt. Entscheidend ist hierbei die Sicherstellung der Professionalität der Evaluationsdurchführung, wozu die Einhaltung internationaler Standards ebenso gehört wie die Vermeidung von Beeinflussung der Evaluationsergebnisse durch einzelne Interessengruppen. Der Einsatz wissenschaftlicher Evaluationen nach internationalem Standard und die rückhaltlose Offenlegung der Ergebnisse, auch wenn diese möglicherweise nicht den Wünschen und Erwartungen entsprechen, trägt letzten Endes zu einem Klima der Offenheit, des Vertrauens des Dialogs in der Gesellschaft bei, von dem alle profitieren.

Damit diese wichtigen Funktionen dauerhaft erfüllt werden können, bedarf es einer systematischen Weiterentwicklung von Evaluationsverfahren und ihrer gewissenhaften Erprobung in der Praxis. Ferner müssen die gewonnenen Erkenntnisse bei der Durchführung von Evaluationen kontinuierlich zur Überprüfung der theoretischen und methodischen Grundlagen nutzbar gemacht und den unterschiedlichen Gruppen der an Evaluationen Beteiligten (Auftraggeber, Evaluatoren, Evaluierte, Grundlagenforscher, interessierte Öffentlichkeit etc.) angemessen zur Verfügung gestellt werden. Diese Aufgaben sind im Rahmen der Tätigkeiten kommerzieller Consultings oder einer reinen Drittmittelforschung nicht zu leisten, da die Auftraggeber von Evaluationen in der Regel nicht an der Förderung und Finanzierung von Grundlagenforschung interessiert sind. Dementsprechend konnten sich in den USA bereits frühzeitig sozialwissenschaftliche Forschungsinstitute an den Universitäten etablieren, die sich dieser Aufgabe gestellt haben.

In der Bundesrepublik Deutschland fehlte bisher ein solches Forschungsinstitut. Diese Lücke soll nun das 2002 an der Universität des Saarlandes des Saarlandes neu gegründete Centrum für Evaluation (CEval) schließen. Als wissenschaftliches Institut, welches der Pro-

fessur für Soziologie angegliedert ist und eng mit der Fakultät für empirische Humanwissenschaften zusammenarbeitet, leistet es einen Beitrag zur Weiterentwicklung der Evaluationsforschung und zur verstärkten Durchführung und Nutzung von Evaluationen. Neben Grundlagenforschung und eigener Evaluationstätigkeit bereitet das CEval unterschiedliche Nutzer- und Anwenderzielgruppen auf die vielfältigen Evaluationsaufgaben in der Forschung, in Bundes- und Landesministerien, Behörden, bei freien Trägern (Kirchen, Parteien, Wohlfahrtsverbänden, Stiftungen, Nicht-Regierungsorganisationen) und in der Consulting-Wirtschaft vor. Auf Initiative von Professor Reinhard Stockmann wurde außerdem im letzten Jahr die erste deutschsprachige Fachzeitschrift für Evaluationsforschung, die „Zeitschrift für Evaluation“, ins Leben gerufen, deren Redaktion beim CEval angesiedelt ist (für weitere Informationen zur Zeitschrift für Evaluation siehe deren Webseite <http://www.zfev.de>).

Bei der Theorie- und Methodenentwicklung sowie der Beratung und Durchführung von Evaluationen konzentriert sich das CEval auf drei fachliche Arbeitsschwerpunkte: Umweltforschung/Umweltkommunikation, Entwicklungszusammenarbeit sowie allgemeine und berufliche Bildung. Durch seine Evaluationskompetenz in diesen Arbeitsfeldern möchte das Institut dazu beitragen, entscheidungsrelevante Informationen für die Steuerung von Programmen bereitzustellen und die Wirksamkeit und Nachhaltigkeit von Politik offen zu legen, um dadurch die Chancen für eine nachhaltige Entwicklung zu erhöhen.

Neben eigener Forschungstätigkeit bietet das CEval in Kooperation mit anderen Einrichtungen Fort- und Weiterbildungsveranstaltungen an. Z.B. werden in Zusammenarbeit mit der InWent GmbH in Mannheim regelmäßig Schulungen für Praktiker und Praktikerinnen aus Entwicklungsländern im Bereich der beruflichen Bildung durchgeführt. Außerdem legt das CEval gemeinsam mit der Arbeitsgemeinschaft Entwicklungspolitischer Gutachter (AGEG) im Jahr 2003 erstmals ein Weiterbildungsprogramm zu Evaluationsmethoden für Fachkräfte der Entwicklungszusammenarbeit auf. Weitere Kursangebote sind derzeit in Vorbereitung. Nähere Informationen zu den Arbeitsschwerpunkten und den Leistungen des CEval finden Sie auf der Internetseite des Instituts (<http://www.ceval.de>).

LITERATUR

- Baumert, J., W. Bos & R. Lehmann (Hrsg., 2000), TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn (2 Bde.), Opladen: Leske+Budrich.
- Baumert, J., E. Klieme & M. Neubrand (Hrsg., 2001), PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich, Opladen: Leske+Budrich.
- Bortz, J. & N. Döring (2002), Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, Berlin u.a.: Springer (3. Auflage).
- Deutsche Gesellschaft für Evaluation DeGEval (Hrsg., 2001), Evaluation - Reformmotor oder Reformbremse? Köln: Eigenverlag.
- Deutsches Institut für Wirtschaftsforschung (Hrsg., 2001), Evaluation im Spannungsfeld zwischen Wissenschaft und Politik, Vierteljahreshefte zur Wirtschaftsforschung Jg. 69, H. 3.
- Hadley, R.G. & L.K. Mitchell (1995), Counseling research and program evaluation, Pacific Grove, CA: Brooks/Cole.
- Joint Committee on Standards for Education Evaluation, J.R. Sanders (eds., 2000), Handbuch der Evaluationsstandards. Die Standards des 'Joint Committee on Standards for Educational Evaluation'. Aus dem Amerikanischen übersetzt von Wolfgang Beywl und Thomas Widmer. Deutsche Ausgabe bearbeitet und ergänzt von Wolfgang Beywl, Thomas Widmer und James R. Sanders, Opladen: Leske+Budrich.
- Leeuw, F.L. (2000), Evaluation in Europe, in: Stockmann, R. (Hrsg.), Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder, Opladen: Leske+Budrich, S. 57-76.
- Maddaus, G. F., D. L. Stufflebeam & M. S. Scriven (1983), Program evaluation: A historical overview, in: Maddaus, G. F., M. S. Scriven & D. L. Stufflebeam (eds.), Evaluation Models, Boston: Kluwer-Nijhoff, pp. 3-22.
- Mertens, D. M. (2000), Institutionalizing Evaluation in the United States of America, in: Stockmann, R. (Hrsg.), Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder, Opladen: Leske+Budrich, S. 41-56.
- Meyer, W. (2002), Evaluationsgesellschaften und ihre Themenschwerpunkte. Ein weltweiter Überblick. In: Zeitschrift für Evaluation 2/2002, S. 134-142 (im Erscheinen).
- Meyer, W. & Höhns, G. (2002), Was ist Evaluation? Bonn: Bundesinstitut für Berufliche Bildung (Wissenschaftliche Diskussionspapiere Heft 59 – zweisprachig deutsch-russisch)
- Rossi, P. H., H. E. Freeman & G. Hofmann (1988), Programm Evaluation. Einführung in die Methoden angewandter Sozialforschung, Stuttgart: Enke. Neueste Englische Auflage: Rossi, P. H., H. E. Freeman & M. W. Lipsey (1999), Evaluation. A Systematic Approach, Thousand Oaks et al.: Sage (6th ed.).
- Shadish, W. R. (1994), Need-based evaluation: good evaluation and what you need to know about it, in: Evaluation Practice, 15 (3), pp. 347-358.
- Shadish, W. R., D. Newman, M. A. Scheirer & C. Wye (eds., 1995), The American Evaluation Association's Guiding Principles, San Francisco: Jossey-Bass.

-
- Stockmann, R. (1997), *The Sustainability of Development Cooperation*, Baden-Baden: Nomos.
- Stockmann, R. (2000), *Evaluation in Deutschland*, in: ders. (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*, Opladen: Leske+Budrich, S. 11-40.
- Stockmann, R. (2000a), *Evaluation staatlicher Entwicklungspolitik*, in: ders. (Hrsg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*, Opladen: Leske+Budrich, S. 375-408.
- Stockmann, R., Meyer, W., Krapp, S. & Köhne, G. (2000), *Wirksamkeit deutscher Berufsbildungszusammenarbeit. Ein Vergleich zwischen staatlichen und nicht-staatlichen Programmen in der Volksrepublik China*, Opladen: Westdeutscher Verlag.
- Stockmann, R., Meyer, W., Gaus, H., Urbahn, J. & Kohlmann, U. (2001), *Nachhaltige Umweltberatung. Evaluation eines Förderprogramms der Deutschen Bundesstiftung Umwelt*, Opladen: Leske+Budrich.
- Widmer, T. (2000), *Qualität der Evaluation – Wenn Wissenschaft zur praktischen Kunst wird*, in: Stockmann, R. (Hrsg., 2000), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*, Opladen: Leske+Budrich, S. 77-102.
- Wottawa, H. & H. Thierau (1990), *Lehrbuch Evaluation*, Bern u.a.: Huber Verlag.

CEval-ARBEITSPAPIERE:

- Nr. 0 Selbstdarstellung des CEval (Reinhard Stockmann)
- Nr. 1 Evaluation als integriertes Lehr- und Forschungsprogramm (Reinhard Stockmann)
- Nr. 2 Soziologie im Abwärtstrend – Eine empirische Untersuchung zur Situation der Soziologie an den bundesdeutschen Hochschulen (Thomas Knoll, Wolfgang Meyer, Reinhard Stockmann)
- Nr. 3 Qualitätsmanagement und Evaluation – Konkurrierende oder sich ergänzende Konzepte? (Reinhard Stockmann)
- Nr. 4 Konzept zur Evaluation von E-Learning Angeboten im Rahmen von VISU (Virtuelle Saar-Universität) (Reinhard Stockmann, Erik Schäffer)
- Nr. 5 Was ist Evaluation? (Wolfgang Meyer)
- Nr. 6 Sociological Theory and Evaluation Research. An Application and its Usability for Evaluating Sustainable Development (Wolfgang Meyer)
- Nr. 7 Evaluation der Umweltberatungsprojekte des Bundesumweltministeriums und des Umweltbundesamtes / Evaluation of Environmental Consulting Projects sponsored by BMU and UBA (Wolfgang Meyer, Klaus-Peter Jacoby, Reinhard Stockmann)
- Nr. 8 Measuring the Impact of Vocational Training Projects – An Evaluation Method and its Theoretical Foundations (Reinhard Stockmann, Wolfgang Meyer)

Centrum für Evaluation (CEval)

Universität des Saarlandes
Philosophische Fakultät III
Empirische Humanwissenschaften
Lehrstuhl für Soziologie

Postfach 15 11 50
66041 Saarbrücken

info@ceval.de
www.ceval.de

LEITUNG:

Prof. Dr. Reinhard Stockmann
Tel.: +49 (0)681 – 302 3372
Fax: +49 (0)681 – 302 3899
Email: r.stockmann@mx.uni-saarland.de

BEREICHSKOORDINATION**ENTWICKLUNGSZUSAMMENARBEIT + BILDUNG:**

Dipl.-Soz. Stefanie Krapp
Tel.: +49 (0)681 – 302 4509
Fax: +49 (0)681 – 302 3899
Email: s.krapp@ceval.de

BEREICHSKOORDINATION UMWELT:

Dr. Wolfgang Meyer
Tel.: +49 (0)681 – 302 4358
Fax: +49 (0)681 – 302 3899
Email: w.meyer@mx.uni-saarland.de