

EvaNet-Positionen 5/2004

Dr. Reinhard Stockmann,

Professor für Soziologie an der Universität des Saarlandes

Leiter des Centrums für Evaluation (CEval)

Evaluationsforschung – Ansatz und Methoden

Auszug aus dem Band

Stockmann, Reinhard (Hrsg.) (2004). *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. 2., überarbeitete und aktualisierte Auflage. Opladen: Leske + Budrich.

Inhalt:

1. Evaluation – Ansatz und Vorgehen.....	2
• Evaluations- und Grundlagenforschung	2
• Gegenstände und Definitionen	3
• Analyseperspektiven.....	3
• Formative und summative Evaluationen	4
• Funktionen von Evaluationen.....	5
• Verhältnis von sozialwissenschaftlicher Forschung und Evaluationsforschung.....	7
2. Theorien und Methoden der Evaluationsforschung.....	7
• Theoretische Ansätze und Paradigmen.....	7
• Kalter Krieg der Paradigmen zu Gunsten der Nützlichkeit beendet.....	9
• Standards für die Evaluationsforschung	9
3. Literatur	10
4. Über den Autor	15

1. Evaluation – Ansatz und Vorgehen

Auch wenn den meisten Menschen der Begriff „Evaluierung“ oder „Evaluation“ wenig vertraut sein dürfte¹, ist die Tätigkeit, die damit gemeint ist, so alt wie die Menschheit selbst. Wenn jemand z. B. überprüft, ob gebratenes Fleisch besser schmeckt und bekömmlicher ist als rohes oder wenn jemand ausprobiert, ob bestimmte Pilze oder Pflanzen essbar sind oder nicht, dann führt er eine Evaluation durch. Evaluationen stellen ein wichtiges Instrument zur Generierung von Erfahrungswissen dar. Sie werden durchgeführt, in dem Informationen gesammelt und anschließend bewertet werden, um letztendlich Entscheidungen zu treffen. Die dabei verwendeten Bewertungskriterien können sehr verschieden sein, orientieren sich jedoch sehr oft an dem Nutzen eines Gegenstandes, Sachverhaltes oder Entwicklungsprozesses für bestimmte Personen oder Gruppen. Je nach Kriterienauswahl kann die Nutzenbewertung durch einzelne Personen oder Gruppen dementsprechend sehr unterschiedlich ausfallen.

Deshalb kommt es sehr darauf an, wer eine Evaluation in Auftrag gibt, wer sie durchführt, welche Ziele von wem damit verfolgt werden, welche Vorgehensweisen gewählt und welche Methoden angewendet werden. Damit sind schon einige wichtige Fragen umrissen, mit denen sich Evaluationen auseinandersetzen müssen und die in diesem Band behandelt werden.

Wissenschaftliche Evaluationen unterscheiden sich von Alltagsevaluationen lediglich durch die Anwendung empirischer Forschungsmethoden. Diese umfassen dabei die gesamte Bandbreite sozialwissenschaftlicher Forschungsparadigmen. Systematisch sind sie insofern, als die grundlegenden Regeln für die Sammlung valider und relevanter Daten gelten.²

• Evaluations- und Grundlagenforschung

Evaluationsforschung kann als ein Teilgebiet anwendungsbezogener Forschung verstanden werden, die sich von der Grundlagenforschung in einigen Aspekten unterscheidet. Evert Vedung hebt hervor: „the basic difference between evaluation research and fundamental research is that the former is intended for use“. Während Grundlagenforschung relativ zweckfrei nach Erkenntnissen streben kann, hat Evaluationsforschung einen Auftraggeber, der damit bestimmte Absichten verfolgt. Der Evaluator muss sich deshalb an den Zielbestimmungen seines Auftraggebers orientieren. Ein weiterer wesentlicher Unterschied zur Grundlagenforschung besteht darin, dass Evaluationen immer mit einer Wertung verbunden sind. Während sich Grundlagenforschung weitgehend normativer Urteile enthalten kann, wird bei einer Evaluation immer eine Bewertung verlangt, sie ist Teil des Forschungsauftrags. Dabei kann der Evaluator die Beurteilungskriterien des Auftraggebers bzw. der Zielgruppen, die evaluiert werden übernehmen (z. B. Zielerreichung, Erwartungen der Zielgruppen), oder aber eigene Bewertungskriterien festlegen (z. B. Bedürfnisse der Zielgruppen, Beitrag zum Abbau sozialer oder gesellschaftlicher Ungleichheit).

Keine prinzipiellen Unterschiede sind hingegen zwischen Evaluations- und Grundlagenforschung im Hinblick auf die Auswahl des Untersuchungsgegenstandes sowie die Verwendung von Datenerhebungs- und Analysemethoden zur Identifizierung von Wirkungen und der Bearbeitung der Kausalitätsfrage (Ursache/Wirkungszusammenhänge) zu erkennen.

¹ Die Begriffe „Evaluierung“, „Evaluation“ und „Evaluationsforschung“ werden hier synonym verwendet. Zu definitiven Unterschieden vgl. Wottawa & Thierau (1990: 9).

² vgl. Rossi u.a. (1988: 1ff.), Wottawa/Thierau (1990: 9f.), Will u.a. (1987: 12ff.)

- **Gegenstände und Definitionen**

Gegenstände einer Evaluation können z. B. Personen, Organisationen, Produkte, Reformen, Gesetze, Maßnahmen, Projekte, Programme oder gar Evaluationen selbst sein. Dementsprechend gibt es eine Vielfalt unterschiedlicher Definitionen.

Eine Definition, die sowohl auf die verschiedenen Evaluationsgegenstände passt, als auch auf die besondere Aufgaben- und Zielbestimmung der Evaluations- im Unterschied zur Grundlagenforschung hinweist, schlägt Donna Mertens vor:

„Evaluation is the systematic investigation of the merit or worth of an object (program) for the purpose of reducing uncertainty in decision making.“ (Mertens 1998: 219)

Eine ebenfalls weithin akzeptierte Begriffsbestimmung ist die von Peter Rossi & Howard Freeman, die Evaluationsforschung definieren

„als systematische Anwendung sozialwissenschaftlicher Forschungsmethoden zur Beurteilung der Konzeption, Ausgestaltung, Umsetzung und des Nutzens sozialer Interventionsprogramme. Evaluationsforschung bezeichnet den gezielten Einsatz sozialwissenschaftlicher Forschungsmethoden zur Verbesserung der Planung und laufenden Überwachung sowie zur Bestimmung der Effektivität und Effizienz von (...) sozialen Interventionsmaßnahmen.“ (Rossi u.a. 1988: 3)³

- **Analyseperspektiven**

Die Evaluationsforschung kann sich mit drei Phasen des politischen Prozesses beschäftigen und dabei drei verschiedene Analyseperspektiven einnehmen (vgl. Schaubild 1):

- (1.) Die Evaluationsforschung kann die Phase der Programmentwicklung behandeln, die die Konzeptualisierung und Ausarbeitung einer geplanten Intervention mit einschließt. Dabei hat die Evaluationsforschung die Aufgabe, „die materiellen, personellen, institutionellen, finanziellen, theoretischen Rahmen- bzw. Eingangsbedingungen eines Programms“ zu untersuchen, um zur Erstellung eines Programmdesigns beizutragen (vgl. Brandtstädter 1990b: 217). Dabei sollen möglichst schon frühzeitig negative Effekte eines Programms oder einer Maßnahme abgeschätzt werden (vgl. Hellstern/Wollmann 1980a: 13, 1984: 24; Staudt u.a. 1988: 40f.). Solche Untersuchungen werden „ex-ante evaluations“, „input-evaluations“ oder „preformative evaluations“ (Scriven 1991: 169) genannt.
- (2.) Während der Implementationsphase übernimmt die Evaluationsforschung vor allem Kontroll- und Beratungsfunktionen. Indem Informationen über den Programmverlauf und die Programmergebnisse gesammelt und bewertet werden, sollen Entscheidungshilfen für die Steuerung der Durchführung des Programms gegeben und durch eine frühzeitige Korrekturmöglichkeit Veränderungen des Programmdesigns ermöglicht werden (vgl. Staudt u.a. 1988: 27; Rossi u.a. 1988: 12, 31 u. 63; Wottawa & Thierau 1990: 54). Solche Evaluationen zur Überwachung der Umsetzung und Ausführung von laufenden Programmen („on-going“) werden als „Begleitforschung“ (Rossi u.a. 1988: 11) oder als „formative Evaluationen“ bezeichnet (vgl. Scriven 1980, 1991). Sie

³ Sinngemäße Definitionen finden sich z.B. bei Wittmann (1985:17); Wottawa/Thierau (1990: 9).

beschäftigen sich mit derselben Phase des politischen Prozesses wie die Implementationsforschung⁴ und verfolgen dabei ähnliche Zielsetzungen.

- (3.) Nach Abschluss der Implementation eines Programms kommt der Evaluationsforschung die Aufgabe zu, den vollen Umfang der Effekte, die durch ein Programm ausgelöst wurden, zu erfassen und zu bewerten sowie Zusammenhänge aufzudecken (vgl. Wottawa/Thierau 1990: 55; Scriven 1991: 340). Solche Untersuchungen werden als „summative Evaluationen“ bezeichnet, und in der Regel ex-post durchgeführt.⁵

Demnach kann das Erkenntnisinteresse in der Evaluationsforschung (1.) mehr darauf gerichtet sein, Programme und den Prozess ihrer Durchführung und Wirkungsweise zu verbessern, oder (2.) mehr darauf, die Resultate vollzogener policies zu analysieren. Folgende von Jann (1994: 311) für die Politikfeldanalyse getroffene Unterscheidung, lässt sich deshalb auch auf die Evaluationsforschung übertragen:

„Während auf der einen Seite ‚*analysis for policy*‘ im Vordergrund steht, d.h. die Verbesserung zukünftiger *policies* durch die Anwendung wissenschaftlicher Erkenntnisse und Methoden (*science for action*), ist die zweite Perspektive eher ‚*analysis of policy*‘, mit Erklärung und Verallgemeinerung als vorrangige Zielsetzungen (*science for knowledge*).“⁶

• **Formative und summative Evaluationen**

Evaluationen können mehr *formativ*, d.h. aktiv-gestaltend, prozessorientiert, konstruktiv und kommunikationsfördernd angelegt sein, oder mehr *summativ*, d.h. zusammenfassend, bilanzierend und ergebnisorientiert. Prinzipiell können beide Evaluations-Perspektiven, bei allen *Phasen des politischen Prozesses* eingenommen werden. Da es in der Planungs- und Designphase eines Programms jedoch kaum Ansatzpunkte für eine summative Evaluation gibt, kann sie in der Durchführung nur formativen Charakter haben. Während der Durchführungsphase sind sowohl formative als auch summative Evaluationen möglich. Ex-post Analysen sind in der Regel summative Evaluationen, da der Gestaltungsaspekt entfällt. Durch entsprechende informationelle Rückkopplungsschleifen für Folgeprojekte können sie jedoch auch formative Bedeutung gewinnen (vgl. Schaubild 1).

⁴ Zur Implementationsforschung vgl. u.a. Mayntz (1977, 1980, 1983); Windhoff-Héritier (1980, 1983, 1987, 1993); Wollmann (1980, 1994).

⁵ Die terminologische Unterscheidung zwischen formativer und summativer Evaluation geht auf Scriven (1967, 1972a u. b, 1980, 1983, 1991) zurück. Eine aufschlussreiche Gegenüberstellung der Merkmale von formativer und summativer Evaluation geben Wottawa & Thierau (1990: 56). Vgl. auch Rossi u.a. (1988), Will u.a. (1987: 20ff.).

⁶ Hervorhebungen durch Jann (1994: 311).

Schaubild 1: Dimensionen der Evaluationsforschung

Phasen des politischen Prozesses	Analyseperspektive	Erkenntnisinteresse	Evaluationskonzepte
Programmformulierung / Planungsphase	ex-ante	„analysis for policy“ „science for action“	preformativ/ formativ: aktiv gestaltend, prozessorientiert, konstruktiv
Implementationsphase	on-going	beides möglich	formativ/summativ: beides möglich
Wirkungsphase	ex-post	„analysis of policy“ „science for knowledge“	summativ: zusammenfassend, bilanzierend, ergebnisorientiert

- **Funktionen von Evaluationen**

Generell können Evaluationen vier Funktionen zugeschrieben werden (vgl. Schaubild 2):

1. Erkenntnisfunktion

Mit Hilfe von Evaluationen werden entscheidungsrelevante Daten gesammelt. Es soll u.a. festgestellt werden, ob die Maßnahmen die Zielgruppe erreichen, welche Bedarfe die Zielgruppe hat, wie es mit der Akzeptanz des Programmes bestellt ist, ob die Träger in der Lage sind, das Programm effektiv und effizient umzusetzen, wie sich die Rahmenbedingungen verändert haben, welche Kausalbeziehungen bestehen etc. Die durch Evaluationen gesammelten Informationen sollen dann für Steuerungsentscheidungen genutzt werden.

2. Kontrollfunktion

Bei der Beobachtung eines Programms und seiner Wirkungen steht zwar in erster Linie das Interesse im Vordergrund, Defizite zu erkennen, um möglichst rasch steuernd eingreifen zu können. Gleichzeitig gewinnt man natürlich aber auch Informationen, die erkennen lassen, ob alle Beteiligten ihre Aufgaben erfüllen, den eingegangenen Verpflichtungen nachkommen etc., so dass direkt oder indirekt auch eine Form von Kontrolle mit Evaluationen verbunden ist.

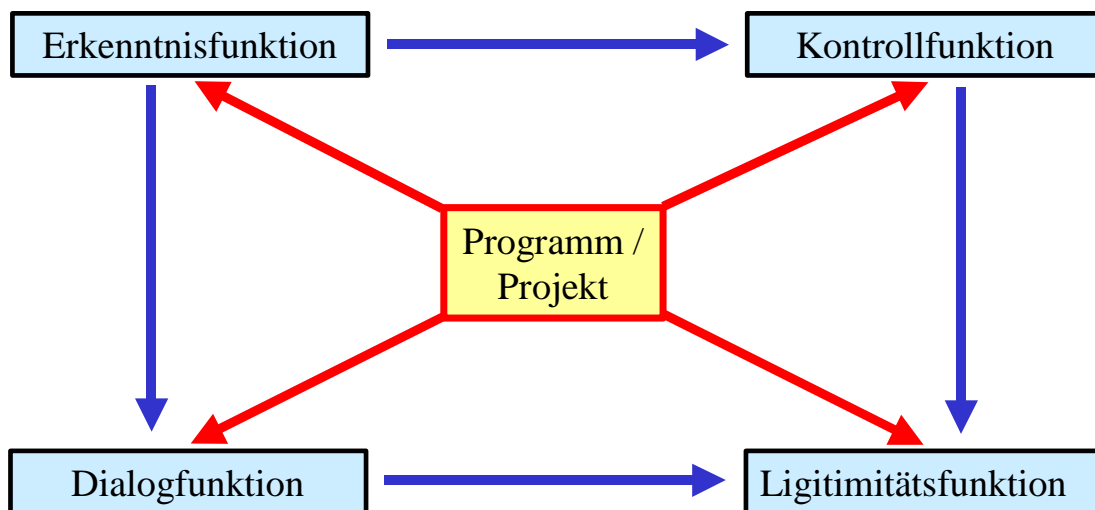
3. Dialogfunktion

Durch Evaluationen werden Informationen bereitgestellt die den Dialog zwischen verschiedenen ‚Stakeholdern‘ (Mittelgeber, Durchführungsorganisation, Zielgruppen, sonstige Beteiligte und Betroffene) auf eine solide Grundlage stellen. Auf der Basis der ermittelten Ergebnisse kann gemeinsam und für alle transparent bilanziert werden wie erfolgreich die Zusammenarbeit verlief und wo Defizite auftraten, um daraus Konsequenzen für die Gestaltung der weiteren Zusammenarbeit zu ziehen.

4. Legitimitätsfunktion

Sehr oft werden Evaluationen auch „taktische“ Funktionen zugeschrieben. Davon wird dann gesprochen, wenn die Ergebnisse von Evaluationen nur dazu verwendet werden sollen, um lediglich bestimmte politische Entscheidungen (manchmal sogar nachträglich) zu legitimieren, z. B. weil ein Programm weitergeführt oder im Gegenteil eingestellt werden soll. Mittlerweile ist es für Politiker auch ‚schick‘ geworden „to use evaluations as baubles or as bolsters“ (Pollitt 1998: 223), als dekorative Symbole für eine moderne Politik, ohne die Ergebnisse von Evaluationen ernsthaft nutzen zu wollen. Diese Art von ‚taktischer‘ Funktion lässt sich jedoch kaum mit dem eigentlichen Zweck von Evaluationen vereinbaren und stellt eher ihre pathologische Seite dar. Deshalb wurde sie hier auch nicht in das Aufgabenfeld von Evaluationen mit aufgenommen (vgl. Schaubild 2).

Schaubild 2: Funktionen von Evaluationen



© Stockmann 2000

- **Verhältnis von sozialwissenschaftlicher Forschung und Evaluationsforschung**

Da Evaluationen unterschiedliche Analyseperspektiven einnehmen und unterschiedliche Erkenntnisinteressen verfolgen können, werden auch verschiedene theoretische Ansätze, methodologische Paradigmen und Erhebungsmethoden verwendet. Diese unterscheiden sich prinzipiell nicht von den in der sozialwissenschaftlichen Forschung üblichen Ansätzen und Verfahren. Allerdings gelten *spezielle Anwendungsbedingungen*, da die Evaluationsforschung in einem Kontext konkreten Handelns agiert und ihrer Intention nach nutzungs- und handlungsorientiert ausgerichtet ist (vgl. Weiss 1974: 7 u. 25; Rossi u.a. 1988: 3; Vedung 1999: 12). Wenn sie darauf abzielt, die Entwicklung von Programminitiativen zu unterstützen, die Planung und Umsetzung sozialer Interventionen zu beeinflussen und ihre Durchführung zu verbessern, ist sie „in einem weiten Sinn auch eine politische Tätigkeit“ (Rossi u.a. 1988: 183). Indem sie Gestaltungs-, Kontroll-, Steuerungs- und Bewertungsfunktionen innerhalb von Handlungsfeldern der Politik wahrnimmt und ihre Ergebnisse in den Verwaltungs- und Politikprozess zurückmeldet, wird sie selbst zu einem Politikum und bewegt „sich notwendig in einem Minenfeld politischer, administrativer und gesellschaftlicher Interessen“ (Hellstern/Wollmann 1980b: 61). Evaluationsforscher müssen deshalb die soziale Ökologie ihres Arbeitsumfeldes berücksichtigen. Verschiedene Interessengruppen sind direkt oder indirekt an den Evaluierungen beteiligt und können die Durchführung behindern oder fördern. Solche ‚Stakeholders‘ können politische Entscheidungsträger, die Auftraggeber der Evaluation, Durchführungsorganisationen, Implementationsträger, Programmteilnehmer, Zielgruppen, Projektmitarbeiter, Programmkonkurrenten und andere sein.

Aufgrund dieser „Dualität“ der Evaluationsforschung, die sich darin ausdrückt, dass sie einerseits Teil der empirischen Sozialforschung ist und sich ihrer Theorien und Methoden bedient, aber andererseits auch Teil des politischen Prozesses ist, den sie selbst mit ihren Ergebnissen beeinflusst und umgekehrt als Instrument zur Entscheidungsfindung für die politische Steuerung wissenschaftsfremden Anforderungen ausgesetzt ist, haben sich im Laufe der Entwicklung der Evaluationsforschung unterschiedliche methodologische Paradigmen herausgebildet. Diese orientieren sich entweder stärker an wissenschaftlichen Standards oder an den Informationsbedürfnissen der Auftraggeber bzw. der Zielgruppen.

2. Theorien und Methoden der Evaluationsforschung

Da sich die Evaluationsforschung nicht losgelöst von der allgemeinen wissenschaftlichen Forschung entwickelt, fanden und finden in ihr auch die dort geführten Debatten mit nicht geringerer Heftigkeit ihre Fortsetzung.

- **Theoretische Ansätze und Paradigmen**

Vor allem in den Anfangsjahren der Evaluationsforschung herrschte ein *positivistisch bestimmter methodologischer Rigorismus* vor. Experimentellen Designs wurde der Vorzug gegeben um „die wahren Zusammenhänge zwischen kausalen Kräften“ (Cook/Matt 1990: 20) aufzudecken, damit politische Entscheidungen über Interventionsprogramme nicht auf falschen Behauptungen hinsichtlich dessen beruhen, was ein Programm leistet und was nicht. Donald Campbell, der sich die Welt als ein Labor für Sozialexperimente vorstellte (vgl. Campbell 1969: 409ff.), misst deshalb der „internen Validität“ die größte Bedeutung bei Eva-

lationen zu, d.h. der Sicherheit, dass die Verbindung zwischen zwei Variablen eine kausale Beziehung einer antezedenten Variablen und ihrer Konsequenz reflektiert (vgl. Cook /Campbell 1979; Scriven 1972, 1980, 1991).

Nicht zuletzt als Reaktion auf diesen methodologischen Rigorismus formierte sich ein Lager von Evaluatoren, das vor allem die instrumentelle Funktion der Evaluationsforschung für die Entscheidungsfindung im politischen Prozess in den Vordergrund stellte: Evaluation wird eher als politischer, denn als wissenschaftlicher Akt gesehen. Aus Gründen der Ethik und wegen der politischen Nützlichkeit sollten Evaluationen vor allem den Interessen der Betroffenen dienen. Die Evaluatoren sollen dabei die Rolle von Lehrern übernehmen und allen beteiligten Parteien bei der Evaluation beratend zur Seite stehen. Evaluation sollte zu einer Kunst entwickelt werden (vgl. Cronbach u.a. 1981; Cronbach 1982; Wottawa/Thierau 1990: 33).

Das *handlungstheoretische Paradigma* führt diese Gedanken weiter und postuliert, dass Evaluationen nicht nur die Kontrolle der Qualität von Innovationen beinhalten darf „sondern gleichzeitig die Konstruktion, Optimierung und die Legitimierung der Modellmaßnahmen zu beinhalten habe“ (Lange 1983: 256). Daraus ergeben sich folgende, dem konventionellen sozialwissenschaftlichen Forschungsparadigma zuwiderlaufende, methodische Konsequenzen:

- Nicht die Falsifikation von Theorien oder Hypothesen ist primäres Erkenntnisziel, sondern die Angabe von Handlungsalternativen zur Lösung auftretender Probleme.
- Die Trennung zwischen Evaluatoren und den Evaluationsobjekten wird aufgehoben. Die Wissenschaftler geben ihre distanzierte Position zum Untersuchungsgegenstand auf und werden zu gleichberechtigten Partnern von den unmittelbar an der Evaluation Beteiligten und den davon Betroffenen (Evaluationsforschung als Aktionsforschung).
- Nicht die Forschungsfragen des Evaluators stehen im Mittelpunkt des Interesses, sondern die Informationsbedarfe der Zielgruppen.
- Nicht Wertneutralität in den Aussagen wird angestrebt, sondern im Gegenteil es werden stellungsbeziehende Wertungen verlangt.
- Gütekriterien der Evaluation sind nicht mehr primär Validität, Reliabilität und Objektivität, sondern Kommunikation, Intervention, Transparenz und Relevanz. (vgl. Gruschka 1976: 142-151; Weiss 1972: 6f.; Rein 1984: 179; Lachenmann 1987: 320; Staudt 1988: 27f.; Gage 1990: 45ff.; Schneider-Barthold 1992: 379ff.)

Der positivistisch unterfütterte methodologische Rigorismus wurde zudem vom *konstruktivistischen, interpretativen Paradigma* in Frage gestellt. Das Vorhandensein einer einzigen ‚wahren‘ Realität wird bestritten. Statt dessen wird angenommen, dass Realität aus verschiedenen Perspektiven sozial konstruiert ist, die in Konflikten zueinander stehen können. Deshalb verwenden die Anhänger dieses Ansatzes auch andere, mehr qualitativ ausgerichtete Forschungsmethoden, um die Ergebnisse besser in ihrem partikularen Kontext interpretieren zu können. Evaluierungsergebnisse werden deshalb zentral bestimmt durch das spezifische Programm das gerade jetzt (zu einem bestimmten Zeitpunkt), unter spezifischen Umständen und mit ausgewählten Betroffenen (stakeholdern) evaluiert wird. Generalisierungen sind deshalb nur sehr begrenzt möglich. (vgl. vor allem Guba/Lincoln 1989; Patton 1987; Stake 1983)

Als eine Weiterführung des handlungstheoretischen und des konstruktivistischen Paradigmas kann das *transformative/emanzipatorische Paradigma* betrachtet werden, das die unterschiedliche Machtstellung einzelner Stakeholder-Gruppen hervorhebt, die zur Überbetonung bestimmter Interessen führen kann (vgl. Mertens 1998 in diesem Band⁷).

⁷ gemeint ist Stockmann, Reinhard (Hrsg.) (2004). *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. 2., überarbeitete und aktualisierte Auflage. Opladen: Leske + Budrich.

- **Kalter Krieg der Paradigmen zu Gunsten der Nützlichkeit beendet**

Mittlerweile gilt der ‚kalte Krieg‘ der Paradigmen zwar keineswegs als endgültig beendet. Unterschiedliche Sichtweisen bleiben weiterhin bestehen. Doch es herrscht weitgehend Konsens darüber, dass Evaluationen die Perspektiven und Bedürfnisse der Stakeholder zu berücksichtigen haben, und dass quantitative und qualitative Methoden zu verwenden sind (Multi-Methodenansätze).

Von vielen wird mittlerweile die Auffassung geteilt, „dass eine Untersuchung gleichzeitig strengen wissenschaftlichen Anforderungen genügen und für den Auftraggeber und andere Interessengruppen von maximalem Nutzen sein kann“ (Rossi u.a. 1988: 10). Sie folgen dabei der kritisch-rationalen Forschungslogik und halten prinzipiell alle bekannten empirischen Forschungstechniken für einsetzbar.

Chelimsky (1995: 6) beschreibt die neuere Entwicklung so:

„We think less today about the absolute merits of one method versus another, and more about whether and how using them in concert could result in more conclusive findings.“ Die eigentlichen Evaluationsfragen rücken dabei mehr in den Vordergrund: „We have learned that the choice of methods (and measures and instruments and data) depends much more on the type of question being asked than on the qualities of any particular method.“ (ebenda)

Über alle theoretischen und methodischen Fragen hinweg besteht zudem Einigkeit darin, dass es vor allem darauf ankommt, dass Evaluationsergebnisse im politischen Prozess genutzt werden (vgl. Chelimsky 1995: 8). Damit rückt das Kriterium der Nützlichkeit von Evaluationsergebnissen für die Auftraggeber und die unterschiedlichen Stakeholder-Gruppen in den Vordergrund. Seit Mitte der 80er Jahre ist dies ein Schwerpunkt in der Evaluationsdiskussion (vgl. vor allem Patton 1997).

Die Qualität von Evaluationen kann demnach nicht allein an einem Gütekriterium gemessen werden. Neben der Wissenschaftlichkeit der Ergebnisse, kommt es eben auch darauf an, dass Evaluationen nützlich sind, denn nur dann werden sie auf politische und soziale Veränderungsprozesse erwirken können.

- **Standards für die Evaluationsforschung⁸**

Im Zuge der Professionalisierung der Evaluationsforschung haben Ende der 70er Jahre verschiedene Organisationen in den USA eine Reihe von Kriterien entwickelt mit denen die Qualität von Evaluationen erfasst werden soll. Am weitesten verbreitet haben sich die ursprünglich vom „Joint Committee on Standards for Educational Evaluation“ vorgelegten „Standards for Evaluation“, die postulieren, dass Evaluationen

- nützlich sein sollen, d.h. an den Informationsbedürfnissen der Nutzer ausgerichtet sind (*Nützlichkeit*),
- realistisch, gut durchdacht, diplomatisch und kostenbewusst durchgeführt werden sollen (*Durchführbarkeit*),
- rechtlich und ethisch korrekt ablaufen und dem Wohlergehen der in die Evaluation einbezogenen und von den Ergebnissen betroffenen Personen Aufmerksamkeit schenken sollen (*Korrektheit*) und

⁸ Mit der *Qualität von Evaluationen* befasst sich der Beitrag von Thomas Widmer, der auch auf die Entstehung, Struktur und den Inhalt der Standards sowie ihre Anwendung in der Evaluationspraxis eingeht. (in: Stockmann, Reinhard (Hrsg.) (2004) *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. 2., überarbeitete und aktualisierte Auflage. Opladen: Leske + Budrich.)

- über die Güte und/oder die Verwendbarkeit eines evaluierten Programms fachlich angemessene Informationen hervorbringen und vermitteln sollen (*Genauigkeit*).

3. Literatur

- Alber, Jens (1989): Der Sozialstaat in der Bundesrepublik Deutschland 1950-1983. Frankfurt, New York: Campus.
- Argyris, Chris u.a. (1999): Die lernende Organisation. Stuttgart: Kett-Kotta.
- Aurin, Kurt/Stolz, Gerd E. (1990): Erfahrungen aus der Aufarbeitung von Evaluationsvorhaben am Beispiel der Tätigkeit der Projektgruppe „Gesamtschule“ der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung. In: Zeitschrift für Pädagogische Psychologie. Jg. 4, H. 4, S. 269-282.
- Beywl, Wolfgang (1999): Handbuch der Evaluationsstandards. Die Standards des „Joint Committee for Educational Evaluation“. In: Klaus Künzel (Hg.): Internationales Jahrbuch der Erwachsenenbildung. Vol. 27, Evaluation der Weiterbildung. Köln, Weimar, Wien: Böhlau. S.: 269-271.
- Bortz, Jürgen/Döring, Nicola (1995): Forschungsmethoden und Evaluation. 2. vollständig überarb. und aktualisierte Aufl.. Berlin, Heidelberg, New York: Springer.
- Brandstädter, Jochen (1990b): Evaluationsforschung: Probleme der wissenschaftlichen Bewertung von Interventions- und Reformprojekten. In: Zeitschrift für Pädagogische Psychologie. Jg. 4, H. 4, S. 215-228.
- Bundesrechnungshof, Präsident des (1989): Erfolgskontrolle finanzwirksamer Maßnahmen in der öffentlichen Verwaltung. Stuttgart, Berlin, Köln: Kohlhammer.
- Bussmann, Werner u.a. (Hg.) (1997): Einführung in die Politikevaluation. Basel, Frankfurt a. Main: Helbing & Lichtenhahn.
- Campbell, Donald T. (1969): Reform as Experiments. In: American Psychologist. Jg. 24, H. 4, S. 409-429.
- Chen, Huey (1990): Theory-Driven Evaluations. Newbury Park: Sage.
- Chen, Huey/Rossi, Peter H. (1983): Evaluating With Sense: The Theory-Driven Approach. In: Evaluation Review. Jg. 7, S. 283-302.
- Chen, Huey/Rossi, Peter H. (1980): The Multi-Goal, Theory-driven Approach to Evaluation: A Model Linking Basic and Applied Social Science. In: Social Forces. Jg. 59, S. 106-122.
- Chelimsky, Eleanor (1995): New dimensions in evaluation. In: World Bank Operations Evaluations Department (OED): Evaluation and Development: proceedings of the 1994 World Bank Conference Washington D.C. S. 3-11.
- Cook, T. D./Matt, G. E. (1990): Theorien der Programmevaluation. In: Koch, Uwe/Wittman, Werner W.: Evaluationsforschung: Bewertungsgrundlage von Sozial- und Gesundheitsprogrammen. Berlin u.a.: Springer.
- Cook, T. D./Campbell, Donald T. (1979): Quasi-Experimentation. Design and Analysis Issues for Field Settings. Chicago: Rand McNally.
- Cronbach, Lee J. u.a. (1981): Toward Reform of Program Evaluation. San Francisco u.a.: Jossey-Bass.

- Cronbach, Lee J. (1982): *Designing Evaluations of Educational and Social Programs*. San Francisco u.a.: Jossey-Bass.
- Derlien, Hans-Ulrich (1997): Die Entwicklung von Evaluation im internationalen Kontext. In: Bussmann, Werner u.a. (Hg.) (1997): *Einführung in die Politikevaluation*. Basel; Frankfurt a. Main: Helbing & Lichtenhahn. S. 4-13
- Derlien, Hans-Ulrich (1994): Evaluation zwischen Programm und Budget. In: Hofmeister, Albert (Hg.): *Möglichkeiten und Grenzen der Programmsteuerung: Controlling und Evaluation*. *Verwaltungspraxis in Ost und West in Zeiten des Wandels*. Schriftenreihe der Schweizerischen Gesellschaft für Verwaltungswissenschaft. Band 21. S.: 43-61.
- Derlien, Hans-Ulrich (1990b): Genesis and Structure of Evaluation Efforts in Comparative Perspective. In: Rist, Ray C.: *Program Evaluation and the Mangement of Government*. New Brunswick/London.
- Deutscher, Irvin/Ostrander, Susan A. (1985): Sociology and Evaluation Research: Some Past and Future Links. In: *History of Sociology*. Jg. 6, S. 11-32.
- Fend, Helmut (1982): *Gesamtschule im Vergleich. Bilanz der Ergebnisse des Gesamtschulversuchs*. Weinheim, Basel: Beltz.
- Gagel, Dieter (1990): *Aktionsforschung - Methoden partizipativer Handwerksförderung*. In: Boehm, Ullrich/Kappel, Robert: *Kleinbetriebe des informellen Sektors und Ausbildung im sub-saharischen Afrika*. Hamburg.
- Gruschka, A. (Hg.) (1976): *Ein Schulversuch wird überprüft. Das Evaluationsdesign für die Kollegstufe NW als Konzept handlungsorientierter Begleitforschung*. Kronberg.
- Guba, Egon G./Lincoln, Yvonna S. (1989): *Fourth Generation Evaluation*. Newbury Park, London, New Delhi: Sage.
- Habersam, Michael (1997): *Controlling als Evaluation – Potentiale eines Perspektivenwechsels*. München und Mering: Rainer Ham
- Hellstern, Gerd-Michael (Hg.) (1984): *Handbuch zur Evaluierungsforschung : Band 1*. Opladen : Westdt. Verl.
- Hellstern, Gerd-Michael/Wollmann, Hellmut (1984): *Bilanz – Reformexperimente, wissenschaftliche Begleitung und politische Realität*. In: dies.: *Handbuch zur Evaluierungsforschung*. Bd. 1. Opladen: Westdeutscher Verlag.
- Hellstern, Gerd-Michael/Wollmann, Hellmut (1980): *Evaluierung in der öffentlichen Verwaltung – Zweck und Anwendungsfelder*. In: *Verwaltung und Fortbildung*. S. 61 ff.
- Héritier, Adrienne (Hg.) (1993): *Policy-Analyse: Kritik und Neuorientierung*. Opladen: Westdt. Verl.
- Horváth, Péter (1996): *Controlling*. 6. Aufl., München: Vahlen.
- Jann, Werner (1994): *Politikfeldanalyse*. In: Nohlen, Dieter: *Lexikon der Politik*. Bd. 2: *Politikwissenschaftliche Methoden*. (hrsg. von Jürgen Kuz/Dieter Nohlen/Rainer-Olaf Schulze). München: Beck.
- Joint Committee on Standards for Educational Evaluation, James R. S. (Hg.) (2000): *Handbuch der Evaluationsstandards. Die Standards des „Joint Committee on Standards for Educational Evaluation“*. Aus dem Amerikanischen übersetzt von Wolfgang Beywl /Thomas Widmer. Deutsche Ausgabe bearbeitet und ergänzt von Wolfgang Beywl/Thomas Widmer/James R. Sanders. 2. Auflage. Opladen: Leske + Budrich.

- Koch, Uwe/Wittmann, Werner W. (Hg.) (1990): Evaluationsforschung. Bewertungsgrundlage von Sozial- und Gesundheitsprogrammen. Berlin u.a.: Springer-Verlag.
- Kohler-Koch, Beate (1991): Inselillusion und Interdependenz: Nationales Regieren unter den Bedingungen von „International Governance“. In: Blanke, Bernhard/W. H. (Hg.): Die Alte Bundesrepublik. Opladen. S. 45-67.
- Lachenmann, Gudrun (1987): Soziale Implikationen und Auswirkungen der Basisgesundheitspolitik. In: Schwefel, D.: Soziale Wirkungen von Projekten in der Dritten Welt. Baden-Baden: Nomos.
- Lachenmann, Gudrun (1977): Evaluierungsforschung: Historische Hintergründe, sozialpolitische Zusammenhänge und wissenschaftliche Einordnung. In: Kantowsky, Detlef: Evaluierungsforschung und -praxis in der Entwicklungshilfe. Zürich: Verlag der Fachvereine.
- Lange, Elmar (1983): Zur Entwicklung und Methodik der Evaluationsforschung in der Bundesrepublik Deutschland. In: Zeitschrift für Soziologie. Jg. 12, H. 3, S. 253-270.
- Leeuw, Frans L. u.a. (2000): Can Governments learn? Comparative Perspectives on Evaluation and Organizational learning. New Brunswick: Transaction Pub.
- Lewin, K. (1951): Field Theory in social science. New York: Harper. Deutsch: Feldtheorie in den Sozialwissenschaften. Bern: Huber, 1963.
- Lippitt, R. (1940): An Experimental Study of Authoritarian and Democratic Group Atmospheres. Univ. Iowa Stud. Child. Welf., 16, S. 45-195.
- Mayer, E. u.a. (Hg.) (1990): Handbuch Controlling. Stuttgart: Poeschel.
- Mayntz, Renate (1977): Die Implementation politischer Programme: Theoretische Überlegungen zu einem neuen Forschungsgebiet. In: Die Verwaltung. S. 51 ff.
- Mayntz, Renate (1980c): Die Entwicklung des analytischen Paradigmas der Implementationsforschung. In: Mayntz, Renate: Implementation politischer Programme. Königsstein: Athenäum.
- Mayntz, Renate (1980e): Die Implementation politischer Programme: Theoretische Überlegungen zu einem neuen Forschungsgebiet. In: Mayntz, Renate: Implementation politischer Programme. Königsstein: Athenäum.
- Mayntz, Renate (Hg.) (1980d): Implementation politischer Programme. Königsstein: Athenäum.
- Mayntz, Renate (1983): Implementation politischer Programme II. Ansätze zur Theoriebildung. Opladen.
- Mayntz, Renate (1997): Soziologie in der öffentlichen Verwaltung. Heidelberg: C.F.Müller.
- Mertens, Donna M. (1998): Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches. Thousand Oaks, CA: Sage.
- Palumbo, Dennis J./Oliverio, Annamarie (1989): Implementation Theory and the Theory-Driven Approach to Validity. In: Evaluation and Program Planning. Jg. 12, S. 337-344.
- Patton, Michael Q. (1997): Utilization – Focused Evaluation: The New Century Text 3. Aufl.. Thousand Oaks, London, New Delhi: Sage.
- Patton, M. Q. (1987): Evaluation's Political Inherency: Practical Implications for Design and Use. In: Palumbo, D. J. (Hg.): The Politics of Program Theory. Thousand Oaks, CA: Sage. S. 100-145.

- Pollitt, Christopher (1998): Evaluation in Europe: Boom or Bubble? In: Evaluation. Jg. 4(2), H. 2, S. 214-224.
- Rein, M. (1984): Umfassende Programmevaluierungen. In: Hellstern, Gerd-Michael/Wollmann, Helmut: Handbuch zur Evaluierungsforschung. Bd. 1. Opladen: Westdeutscher Verlag.
- Roethlisberger, F. J./Dickson, W. J. (1934): Management and the Worker. Cambridge.
- Rossi, Peter H./ Freeman, Howard E. (1999): Evaluation. A Systematic Approach. 6 Aufl.. Thousand Oaks u.a.: Sage.
- Rossi, Peter H./Freeman, Howard E./Hofmann, Gerhard (1988): Programm Evaluation: Einführung in die Methoden angewandter Sozialforschung. Stuttgart: Enke.
- Schneider-Barthold, Wolfgang (1992): Zur Angemessenheit von quantitativen und qualitativen Erhebungsmethoden in der Entwicklungsländerforschung. Vorzüge und Probleme der Aktionsforschung. In: Reichert, Ch./Scheuch, Erwin K./Seibel, H. D.: Empirische Sozialforschung über Entwicklungsländer. Methodenprobleme und Praxisbezug. Saarbrücken: Breitenbach.
- Schröter, Eckhard (1998): New Public Management. In: Bandemer, Stephan (Hg.): Handbuch Zur Verwaltungsreform. Opladen. S. 59-69.
- Scriven, Michael (1972a): Pros and Cons About Goal-Free Evaluation. In: Evaluation Comment. S. 1-4.
- Scriven, Michael (1972b): Die Methodologie der Evaluation. In: Wulf, Christoph: Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München: Piper.
- Scriven, Michael (1980): The Logic of Evaluation. California: Edgepress.
- Scriven, Michael (1991): Evaluation Thesaurus. Newbury Park u.a.: Sage.
- Scriven, Michael (1967): The Methodology of Evaluation. In: Stake, R. E.: AERA Monograph Series on Curriculum Evaluation. Vol. 1 Chikago: Rand McNally.
- Scriven, Michael (1983): Evaluation Ideologies. In: Madaus, G. F./Scriven, M./Stufflebeam, D. L.: Evaluation Models: Viewpoints on Educational and Human Services Evaluation. Boston: Kluwe-Nijhoff.
- Stake, R. E. (1983): The Case Study Method in Social Inquiry. In: Madaus, G. F. u.a. (Hg.): Evaluation Models. Boston: Kluwer-Nijhoff. S. 279-86
- Staudt, Erich/Hefkesbrink, Joachim/Treichel, Heinz-Reiner (1988): Forschungsmanagement durch Evaluation: Das Beispiel Arbeitsschwerpunkt Druckindustrie. Frankfurt a.M.: Campus.
- Stouffer, S. A. u.a. (1949): The American Soldier. Vol II: Combat and ist Aftermath. Princeton, NJ: Princeton University Press.
- Vedung, Evert (1999): Evaluation im öffentlichen Sektor. Wien, Köln, Graz: Böhlau.
- Weiss, Carol H. (1972): Evaluation Research. Englewood Cliffs, NJ: Prentice Hall.
- Weiss, Carol H. (1974): Evaluierungsforschung. Opladen: Westdeutscher Verlag.
- White, R./Lippitt, R. (1953): Leader Behaviour and Menber Reaction in Three „Social Climates“. In: Cartwright, D./Zander, A. (Hg.): Group Dynamics, Research and Theory. Evanston, Ill.: Row, Peterson and Company, 1953 (1st ed.).

- Will, Hermann/Blickhan, Claus (1987): Evaluation als Intervention. In: Will, Hermann/ Winteler, Adolf/ Krapp, Andreas: Evaluation in der beruflichen Aus- und Weiterbildung. Heidelberg: Sauer.
- Will, Hermann/Winteler, Adolf/Krapp, Andreas (Hg.) (1987): Evaluation in der beruflichen Aus- und Weiterbildung. Heidelberg: Sauer.
- Will, Hermann/Winteler, Adolf/Krapp, Andreas (1987): Von der Erfolgskontrolle zur Evaluation. In: Will, Hermann/Winteler, Adolf/Krapp, Andreas: Evaluation in der beruflichen Aus- und Weiterbildung. Heidelberg: Sauer.
- Windhoff-Héritier, Adrienne (1980): Politikimplementation. Königsstein.
- Windhoff-Héritier, Adrienne (1983): Policy-Analyse. Eine Einführung. Frankfurt a.M., New York.
- Windhoff-Héritier, Adrienne (1987): Policy-Analyse. Frankfurt a.M., New York.
- Wittmann, Werner W. (1990): Aufgaben und Möglichkeiten der Evaluationsforschung in der Bundesrepublik Deutschland. In: Koch, Uwe/Wittmann, Werner W.: Evaluationsforschung: Bewertungsgrundlage von Sozial- und Gesundheitsprogrammen. Berlin u.a.: Springer.
- Wittmann, Werner (1985): Evaluationsforschung. Aufgaben, Probleme und Anwendungen. Berlin u.a.: Springer.
- Wollmann, Hellmut (1999): Politik- und Verwaltungsmodernisierung in den Kommunen: Zwischen Managementlehre und Demokratiegebot. In: Die Verwaltung. (Schwerpunktheft 3).
- Wollmann, Hellmut (1998): Evaluation research and politics: Between a science-driven and a pluralist controversy-responsive policy-making model. Potential and limitations. Diskussionspapier. Humboldt-Universität zu Berlin.
- Wollmann, Hellmut (1998): Modernisierung der kommunalen Politik- und Verwaltungswelt – Zwischen Demokratie und Managementschub. In: Grunow, Dieter/Wollmann, Hellmut: Lokale Verwaltungsreform in Aktion: Fortschritte und Fallstricke. Basel u.a.: Birkhäuser. S. 400-439.
- Wollmann, Hellmut (1998): Kommunale Verwaltungsmodernisierung in Ostdeutschland. Zwischen Worten und Taten. Diskussionspapier. Humboldt-Universität zu Berlin.
- Wollmann, Hellmut (1997): Evaluation in Germany. In: European Evaluation Society. Newsletter (3). S. 4-5.
- Wollmann, Hellmut (1994): Evaluierungsansätze und -institutionen in Kommunalpolitik und -verwaltung. Stationen der Planungs- und Steuerungsdiskussion. In: Schulze-Böing, Matthias/Johrendt, Norbert: Wirkungen kommunaler Beschäftigungsprogramme. Methoden, Instrumente und Ergebnisse der Evaluation kommunaler Arbeitsmarktpolitik. Basel, Boston, Berlin: Birkhäuser. S. 79-110.
- Wollmann, Hellmut (1994): Implementationsforschung/Evaluationsforschung. In: Nohlen, Dieter: Lexikon der Politik. Bd. 2: Politikwissenschaftliche Methoden (hrsg. von Kuz, Jürgen/Nohlen, Dieter/Schulze, Rainer-Olaf). München: Beck.
- Wollmann, Hellmut (1980): Implementationsforschung - eine Chance für kritische Verwaltungsforschung? In: Wollmann, Hellmut: Politik im Dickicht der Bürokratie. Beiträge zur Implementationsforschung. (Leviathan Sonderheft 3).

Wottawa, Heinrich T. H. (1998): Lehrbuch Evaluation. 2. Auflage.

4. Über den Autor

Prof. Dr. Reinhard Stockmann

Lehrstuhl für Soziologie an der Universität des Saarlandes. Leiter des Centrums für Evaluation (CEval).

Gründungsmitglied der Deutschen Gesellschaft für Evaluation (DeGEval), Leiter des Arbeitskreises „Evaluation und Entwicklungspolitik“, Herausgeber der Reihe „Evaluationsforschung“. Geschäftsführender Herausgeber der „Zeitschrift für Evaluation“ (ZfEv).

Arbeitsschwerpunkte: Methoden der empirischen Sozialforschung, insbesondere Evaluationsforschung, Entwicklungs-, Bildungs- und Umweltsoziologie, Politikberatung und Evaluation in der Entwicklungszusammenarbeit und im Umweltbereich.

Kontakt:

Universität des Saarlandes

Lehrstuhl für Soziologie FR 5.2

Postfach 15 11 50

66041 Saarbrücken

Tel.: 0681/302 33 72

Fax: 0681/302 38 99

E-Mail: <mailto:r.stockmann@mx.uni-saarland.de>

Internet: www.uni-saarland.de/fak5/stockmann
www.ceval.de

Dieser Text steht unter:

<http://evanet.his.de/evanet/forum/pdf-position/StockmannPosition.pdf>