

## Von der Evaluation zur Evaluationsforschung - Perspektiven der qualitativen Sozialforschung

Arnold, Maik

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

### Empfohlene Zitierung / Suggested Citation:

Arnold, M. (2005). Von der Evaluation zur Evaluationsforschung - Perspektiven der qualitativen Sozialforschung. *SIETAR-Journal*, 11(2), 11-14. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-337465>

### 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.

# Von der Evaluation zur Evaluationsforschung

## Perspektiven der qualitativen Sozialforschung

Maik Arnold

### Auf dem Weg zu einer Theorie qualitativer Evaluation<sup>1</sup>

Gegenwärtige Diskussionen im Forschungsfeld interkultureller Kommunikation und Kompetenz beschäftigen sich u.a. mit der Einschätzung der Wirksamkeit und Qualität von interkulturellen Trainings und anderen Weiterbildungsmaßnahmen, welche maßgeblich von der zugrunde gelegten theoretischen Konzeption und methodischen Vorgehensweise abhängig sind (vgl. u.a. Mendenhall *et al.* 2004; O' Reilly/Arnold 2005). In Erweiterung und Ergänzung bisheriger Konzeptionen von *Evaluation* wird im Folgenden ein möglicher Zugang zur Evaluationsforschung aus Sicht der qualitativen Sozialforschung beschrieben.<sup>2</sup>

In der einschlägigen Literatur gibt es mittlerweile eine fast unüberschaubare Vielfalt an Definitionen des Begriffs *Evaluation*. Patton (2002) bezeichnet die Evaluationsforschung – im Gegensatz zur Grundlagenforschung – als *Anwendungsforschung*, weil sie an Ziele und Rahmenbedingungen der Auftraggeber gebunden ist und unter Ergebniszwang steht.<sup>3</sup> Eine genaue Begriffsbestimmung wird dadurch erschwert, dass der Begriff *Evaluation* häufig synonym mit anderen Begriffen wie Erfolgs- oder Wirkungskontrolle, Controlling oder Begleitforschung verwendet wird (vgl. u.a. Beywl/Schobert 2000; Rossi/Freeman/Hofmann 1988). Hinzu kommt, dass trotz verstärkter interdisziplinärer Forschungsbemühungen (vgl. u.a. Scriven 1991) eine grundlegende Theorie und Methodologie der *Evaluation* fehlt (vgl. u.a. Lee in Stockmann 2000; Guba 1969; Shadish/Cook/Leviton 1991).

Seit den 1980er Jahren lassen sich verstärkt Bemühungen mit dem Ziel einer theoretischen Fundierung des Konzeptes *Evaluation* erkennen. Gegenwärtige Forschungsschwerpunkte bestätigen die Interdisziplinarität dieses Konzeptes:<sup>4</sup>

- angemessene Forschungsdesigns (z.B. qualitative Forschungsdesigns, vgl. Shaw 1999; Patton 2002 bzw. (quasi-)experimentelle Designs, vgl. Campbell 1964, 1996; Wittmann 1985; Wottawa/Thierau 1998);
- Steigerung der Aussagekraft der Ergebnisse einer Evaluation unter vorrangiger Berücksichtigung der internen Validität (vgl. u.a. Campbell 1964, 1996; Scriven 1986)

- oder externen Validität (vgl. Cronbach 1980);
- Beteiligung von unterschiedlichen Interessengruppen und Betroffenen an der Evaluation (vgl. Fetterman 1994, 2000; Stake 1994);
- Rollendefinition der Beteiligten an Evaluation (vgl. u.a. Bryk 1983);
- Evaluation durch Explikation des politischen Kontextes (vgl. u.a. Weiss 1987, 1998);
- Auseinandersetzung mit Bewertungskriterien und der Wertfrage (vgl. House 1980; Shadish/Cook/Leviton 1991; Cronbach 1980; Weiss 1973; Stake 1975) und
- die Einführung von Gütekriterien und Standards zur Beurteilung der Qualität einer Evaluation (vgl. u.a. Widmer in Stockmann 2000; Joint Committee 2000).

Das Konzept einer *qualitativen Evaluation* ist schließlich in Form einer empirisch gestützten Beschreibung und Analyse von Erlebnis- und Handlungsfeldern der beteiligten Akteure denkbar, wenn damit das Ziel verfolgt wird, nicht nur die institutionelle Praxis des Evaluierens selbst zu untersuchen, sondern auch die alltäglichen Praktiken, die subjektiven Theorien und Diskurse der Beteiligten und Adressaten in die Evaluation einzubeziehen. Dabei kommen »qualitative, prozess-, kommunikations- und beteiligungsorientierte Verfahren« zum Einsatz (v. Kardorff 2000, 242), ohne dass die Ansprüche der Effizienzquantifizierung und summarischen Evaluation damit ersetzt werden, sondern letztere durch Einbezug der Interessenlagen und Meinungen der Betroffenen und durch eine weniger ergebnisbezogene, dafür aber ablaufbegleitete formative Evaluation (vgl. v. Kardorff 1988) ergänzt und erweitert werden.

### Kriterien und Merkmale für eine qualitative Evaluationsforschung

Die qualitative Evaluationsforschung kann anhand folgender Merkmale und Kriterien beschrieben werden: Prinzipien, Methodologie, Interpretation und Gütekriterien qualitativer Evaluationsforschung.

### Prinzipien qualitativer Evaluationsforschung

Die grundlegenden Prinzipien einer qualitativen Evaluationsforschung stützen sich im Wesentlichen auf die Prinzi-

Aus Wissenschaft und Forschung

1 Der vorliegende Artikel ist Ergebnis von Vorarbeiten zu einem umfangreichen Forschungsbericht, der im kommenden Jahr erscheinen soll. Aktuelle Informationen über die qualitative Evaluationsforschung und die Evaluation interkultureller Trainings sowie eine ausführliche Literaturliste finden sich auf der Website des Verfassers ([www.qualitative-evaluation.de](http://www.qualitative-evaluation.de)).

2 Verwandte Überlegungen zur Beschreibung einer qualitativen Evaluationsforschung finden sich u.a. bei v. Kardorff 2000 und Guba/Lincoln 1989.

3 Vielleicht liegt es daran, dass in der Praxis der Evaluation häufig »Abkürzungsstrategien« gewählt werden und bisher nur selten aufwändigere Methoden qualitativer Sozialforschung zum Einsatz kamen (vgl. u.a. Flick 2000a).

4 Dies ist nur eine kleine Auswahl der gegenwärtigen Diskussion in der Literatur.

### Aus Wissenschaft und Forschung

prien qualitativer Sozialforschung (vgl. u.a. Lamnek 1995; Flick/v. Kardorff/Steinke 2000).<sup>5</sup> Evaluationsforschung ist unter normativer Perspektive wertgebundene Forschung (vgl. u.a. Lee in Stockmann 2000). Guba und Lincoln (1989) nennen demokratische Werte wie Transparenz, Partizipation, freier Wille, soziale Verantwortung, Aufrichtigkeit und Fairness als maßgebliche Beurteilungskriterien für die Planung, Durchführung und Auswertung der Evaluationsprozesse. Evaluation wird als kommunikativer Aushandlungsprozess verstanden, wobei der Evaluationsforschung selbst eine aktive Rolle in Veränderungsprozessen zukommt. Ihre Funktion als »change-agent« sollte stets reflektiert werden (Freundlieb/Wolff 1999). Die Ergebnisse des beschriebenen Evaluationsvorgehens nehmen die Form reflexiven und orientierenden statt technisch-instrumentellen Wissens an (vgl. Beck/BonB 1989).

### Methodologie einer qualitativ-empirischen Evaluationsforschung

Überlegungen zu einer qualitativen Evaluationsstrategie lassen sich in folgender Globalfrage zusammenfassen: Wer soll für wen was warum auf welche Weise evaluieren (vgl. Froschauer/Lueger 1996)? Um diese Globalfrage zu beantworten, wurden in der empirischen Evaluationsforschung bisher überwiegend Untersuchungsdesigns nach dem deduktiv-nomologischen Erklärungsmodell gewählt (vgl. v. Kardorff 2000 u.a.), obwohl zunehmend – in sinnvoller Ergänzung des Methodenrepertoires – auch aus qualitativer Perspektive theoriegenerierende Forschungsmethoden eingesetzt werden (vgl. u.a. Patton 2002). Der Einsatz letzterer Methoden erscheint besonders deshalb sinnvoll, da Wissensbestände den an einer Evaluation beteiligten Akteuren nicht ohne Weiteres reflexiv zugänglich sind (vgl. u.a. Zielke 2004). Außerdem gestatten sie es, den alltagsweltlichen Konstrukten in einer möglichst weitgehenden methodischen Offenheit bzgl. des Untersuchungsgegenstandes, der Untersuchungsperson und der Untersuchungssituation zu begegnen (vgl. Bohnsack 2003; Straub 1998, 2004; Hoffmann-Riem 1980).

An die Stelle von Kriterien, die statistische Repräsentativitäten gewährleisten, tritt in der qualitativen Evaluationsforschung eine hypothesengeleitete Auswahl von Unter-

suchungseinheiten nach dem Konzept des »theoretical sampling« (Glaser/Strauss 1967). Die Konzeption einer *qualitativen Evaluationsforschung* verfolgt eine Prozessorientierung (formative Evaluation) anstatt einer Ergebnisorientierung (summative Evaluation), wobei der Prozessnutzen – also Lernprozesse, Akzeptanzgewinnung, Analyse von Fehlschlägen – zur Weiterentwicklung der betrachteten Maßnahme führen können (vgl. Patton 2002; v. Kardorff 2000). Dies ermöglicht schließlich rekursive, durch Rückkopplungsschleifen gekennzeichnete Lehr- und Lernprozesse (vgl. u.a. Guba/Lincoln 1989), zumal auch unerwartete Nebenwirkungen in den Wahrnehmungs- und Deutungsmustern der Interaktanten enthalten sind.<sup>6</sup> Die Ergebnisse einer qualitativen Evaluation enthalten eher kontext- und projektbezogene bzw. spezifische, lokal begrenzte Aussagen anstatt Generalisierungen von Sachverhalten. Kontextfaktoren wie Tradition, Historie, Milieus, Netzwerke der Projektmitarbeiter und Macht der Evaluationsbeteiligten haben damit eine höhere Bedeutung.

### Interpretation, Aussagekraft und Gütekriterien einer qualitativen Evaluationsforschung

Die Interpretation und analytische Auswertung der erhobenen qualitativen Daten kann anhand verschiedener Datenauswertungsmethoden erfolgen, wie z.B. mit Hilfe der *Grounded Theory* (vgl. Glaser/Strauss 1967; Glaser 1978; Strauss 1998; Strauss/Corbin 1993), der Biographie- und Erzählanalyse (vgl. u.a. Schütze 1983), der objektiven (bzw. strukturalen) Hermeneutik (vgl. Oevermann/Allert/Konau/Krambeck 1979; Reichertz 1986, 1991), mit konversationsanalytischen Verfahren (vgl. u.a. Sacks 1990; Sacks/Schegloff/Jefferson 1974; Schegloff/Sacks 1973) und der qualitativen Inhaltsanalyse (vgl. u.a. Mayring 1985, 2003; Merten 1983).

In der Auseinandersetzung um die Güte und Wissenschaftlichkeit qualitativer Sozialforschung lassen sich verschiedene Forschungspositionen erkennen (vgl. Appelsmeyer/Kochinka/Straub 2000): Auf der einen Seite bemüht man sich um die Weiterentwicklung und Ausweitung der Gültigkeit traditioneller Gütekriterien quantitativer Sozialforschung – wie z. B. Objektivität, Reliabilität, Validität – auf die qualitative Forschung (vgl. u.a. Kirk/Miller 1986; Silver-

5 Zu den Prinzipien qualitativer Sozialforschung zählen u.a. Offenheit, Kommunikation, Prozesscharakter, Reflexivität, Zurückhaltung und Alltagspassung. Zur Diskussion dieser Kriterien vgl. neben den oben genannten Publikationen die einst wegweisende Arbeit von Hoffmann-Riem (1980).

6 Ein analoger Gedanke findet sich in der Auseinandersetzung um die »goalfree« Evaluation bei Scriven (1972) im Kontrast zur zielorientierten Evaluation von Tyler (1949), bei der neben positiven und negativen auch nicht-intendierte Effekte einer sozialen Interventionsmaßnahme in die Analyse einbezogen werden.

mann 1993). Andererseits wird versucht, die Aussagekraft qualitativer Daten nach anderen als den traditionellen Kriterien, z.B. nach spezifischen methodologischen und methodischen Standards, zu beurteilen (vgl. Flick 1995; Küchler 1983).

Bezogen auf den Prozess der Kommunikation und Interaktion der Beteiligten an einer Evaluation kann man mit Hilfe der *kommunikativen Validierung* (vgl. u.a. Kvale 1995; Groeben/Scheele 1977) die Aussagekraft der Ergebnisse erhöhen, wenn man diese den Forschungspartnern vorlegt und gemeinsam mit ihnen nach ihrer Reichweite, Vollständigkeit und Breite der Argumentation diskutiert. Dies wird z.B. durch externe Audits von Experten und durch die Einbeziehung unterschiedlicher Interessengruppen bzw. so genannter *stakeholders* ermöglicht (vgl. v. Kardorff 2000; Guba/Lincoln 1989). Zusätzlich erscheint die multimethodale und -perspektivische Strategie der *Triangulation* sinnvoll (vgl. u.a. Denzin 1989; Denzin/Lincoln 1994; Marotzki 1995). Dabei werden verschiedenen Datenquellen (wie z.B. Selbst- und Fremdbenachrichtigungen der Akteure und reflektierte eigene Erfahrungen des Evaluators), Untersuchungsmethoden (wie z.B. offene Interviews, Gruppendiskussionen, moderierte Zielfindungs- und Ergebnisgespräche, Dokumentenanalysen und Verfahren der ethnographischen Feldforschung) sowie Theorien der beteiligten Beobachter und Interpreten bei der Datenerhebung und -auswertung miteinander kombiniert (vgl. u.a. Flick 2000b).

Als allgemeines Kriterium der Wissenschaftlichkeit qualitativer Evaluationsforschung gilt die *intersubjektive Nachvollziehbarkeit* (vgl. u.a. Appelmeyer/Kochinka/Straub 2000). Dazu zählt die Offenlegung der forschungsleitenden Interessen und Vorverständnisse, Vereinbarungen und Begründungen über die Wahl der Methoden und die Beschreibung bzw. Analyse der Ergebnisse ebenso wie die Darstellung der einzelnen Interpretationsschritte.

In der Evaluationsforschung gibt es in Bezug auf die Qualitätssicherung und die Einführung von *Standards der Evaluation* eine gesonderte Diskussion und Anfrage an die Wissenschaftlichkeit derartiger Bemühungen. Cronbachs (1983) 95 Thesen und Kriterien für Programmevaluationen gelten als hilfreiche Orientierung (vgl. u.a. House 1980; Sanders 1999).

Ausgehend vom noch unbestimmten Begriff ›Qualität‹ (Kromrey in Stockmann 2000) zählt Widmer (in Stockmann 2000) verschiedene weitere Beurteilungskriterien für die differenzierende Betrachtung und Beurteilung qualitativer Evaluationsforschung auf: die Evaluationsgegenstände (Programm- vs. Ergebnisevaluation), die Verortung einer Evaluation (interne vs. externe; formative vs. summative; Selbst- vs. Fremdevaluation), Methodenwahl (experimentelles vs. konstruktivistisches Paradigma), die Distanz von Evaluation und Evaluandum, der Umfang einer Evaluation (zeitliche, personelle, rechtliche und finanzielle Ressourcen), Erwartungen an die Evaluation seitens der Interessengruppen Politik, Wissenschaft, Öffentlichkeit.

Im Zuge der Professionalisierung der Evaluation haben sich in den 1970er Jahren – zunächst in den USA – verschiedene Initiativen gegründet, um Kriterien zur Einschätzung der Güte und Qualität einer Evaluation zu entwickeln. So kam es zur Formulierung von Evaluationsstandards durch das *Joint Committee on Standards for Educational Evaluation*, wonach u.a. ein Nebeneinander zwischen den qualitativen und klassischen, (quasi-)experimentellen Methoden ermöglicht wird. Im deutschen Sprachraum blieben diese Standards lange Zeit unbeachtet. Erst Ende der 1990er Jahre kam es zur Übersetzung dieser ›Standards‹ (vgl. Joint Committee 2000).<sup>7</sup> Das Regelwerk beinhaltet vier heterogene, z.T. schwer fassbare Konzepte: Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit der Evaluation. In der Evaluationspraxis ist es meist kaum möglich, stets alle Standards einzuhalten, da die mit den einzelnen Konzepten verbundenen Ziele schwer kombinierbar oder integrierbar sind (vgl. u.a. Widmer in Stockmann 2000).

### Praxis der Evaluation

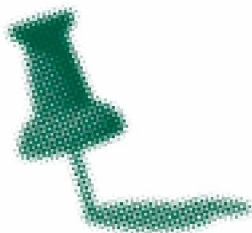
Bei der Darstellung der Ergebnisse steht der Evaluator häufig vor einem nicht zu unterschätzenden Problem, da die Forschungsergebnisse u.U. politisch oder gesellschaftlich hoch sensible Felder berühren können (vgl. v. Kardorff 2000 u.a.). Regelmäßig muss dann zwischen den Interessen der Auftraggeber, dem Datenschutz, der Nachvollziehbarkeit, der wissenschaftlichen Redlichkeit, der Verpflichtung zur Information der Öffentlichkeit und der Schaffung einer für die Evaluation fruchtbaren Vertrauensbasis verhandelt wer-

7 Die Deutsche Gesellschaft für Evaluation e.V. (DeGEval), die sich aus Institutionen und Personen zur Förderung der Professionalisierung und zum Informationsaustausch rund um das Thema Evaluation zusammengeschlossen hat, übernimmt in Deutschland die Herausgabe und Entwicklung von Standards ohne rechtsverbindlichen Charakter. Wie die ISO-Normen haben diese Standards der DeGEval keine rechtsverbindliche Wirkung, es sei denn, dass dies zwischen Auftraggeber und -nehmer der Evaluation vertraglich vereinbart wurde.

### Aus Wissenschaft und Forschung

den. House (1993) empfiehlt, die beteiligten Interessengruppen (u.a. die Auftraggeber) in den Prozess der Ergebnisdarstellung einzubeziehen, um sicherzustellen, dass die Erkenntnisse aus der Evaluation tatsächlich zu Verbesserungsmaßnahmen und individuellen wie institutionellen Lernprozessen führen können (vgl. u.a. Torres/Preskill/Piontek 1996).

Die Evaluationspraxis ist wesentlich pragmatischer in ihrer Herangehensweise. Der Evaluationsmarkt wird weitgehend von Controlling- und Marktforschungsunternehmen dominiert, die nicht in erster Linie an der wissenschaftlichen Diskussion und an der Kumulation von Wissen interessiert sind, sondern marktwirtschaftliche Entscheidungen treffen (Stichwort: gewinnorientierte Auftragsabwicklung; vgl. Stockmann 2000). Die Professionalisierung im Berufsbild *Evaluator* zeichnet sich durch eine stark disziplinäre Segmentierung und institutionelle Zersplitterung aus, wobei immer noch Fragen der Aus- und Weiterbildung – zumindest in Deutschland – ungeklärt sind. Die Umsetzung der Evaluation in Form eines wissenschaftlich geleiteten Vorgehens und die zunehmenden Erwartungen der Öffentlichkeit an die Ergebnisse der Evaluation stellen häufig weitere Probleme dar (vgl. u.a. Freundlieb/Wolff 1999). Diskussionen um die Professionalisierung von Evaluation werden auch vor dem Hintergrund der Verortung des Standpunktes des Evaluators geführt, wobei häufig dessen Unabhängigkeit, Neutralität oder Externalität betont wird. Der Evaluator steht damit in einer hohen Verantwortung, denn sowohl das Forschungsdesign als auch die Einzelinteressen der Beteiligten haben maßgeblichen Einfluss auf die Qualität der Evaluation der untersuchten Maßnahme.



### Literatur

Aus Platzgründen werden nur mehrfach zitierte bzw. übersichtliche Titel aufgeführt. Eine vollständige Literaturliste ist abrufbar auf der Website des Verfassers.

Appelsmeyer, Heide/Alexander Kochinka/  
Jürgen Straub (2000):  
*Qualitative Methoden*. In: Jürgen Straub/Wilhelm Kempff/Hans Werbik (Hrsg.): *Psychologie. Eine Einführung. Grundlagen, Methoden, Perspektiven*. 3. Auflage, München: Deutscher Taschenbuch Verlag, S. 709–742.

Freundlieb, Andreas/Stephan Wolff (1999):  
*Evaluation*. In: Jürgen M. Pelikan/Stephan Wolff (Hrsg.): *Das gesundheitsfördernde Krankenhaus*. Weinheim/München: Juventa, S. 80–91.

Guba, Egon G./Yvonna S. Lincoln (1989):  
*Fourth Generation Evaluation*. Newbury Park, CA u.a.: Sage.

House, Ernest R. (1993):  
*Professional Evaluation. Social Impact and Political Consequences*. Newbury Park, CA u.a.: Sage.

Joint Committee on Standards  
for Educational Evaluation (Hrsg.) (2000):  
*Handbuch der Evaluationsstandards*.  
Opladen: Leske + Budrich.

v. Kardorff, Ernst (2000):  
*Qualitative Evaluationsforschung*. In: Uwe Flick/Ernst v. Kardorff/Ines Steinke (Hrsg.): *Qualitative Methoden*. Ein Handbuch. Reinbek bei Hamburg: Rowohlt, S. 238–250.

Mendenhall, Mark E./Günter K. Stahl/Ina Ehnert/Gary Oddou/Joyce S. Osland/Torsten M. Kühlmann (2004):  
*Evaluation studies of Cross-Cultural Training Programmes: A review of the literature from 1988 to 2000*. In: Dan Landis/Janet M. Bennett/Milton J. Bennett (Hrsg.): *Handbook of Intercultural Training*. 3. Auflage, Thousand Oaks u.a.: Sage Publications, S. 129–143.

Shadish, William/Thomas D. Cook/  
Laura C. Leviton (1991):  
*Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage.

Stockmann, Reinhard (Hrsg.) (2000):  
*Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. Opladen: Leske + Budrich.

Weiss, Carol (1998):  
*Evaluation: Methods for Studying Policies and Programs*. New York: Prentice Hall.

Wottawa, Heinrich/Heike Thierau (1998):  
*Lehrbuch Evaluation*. 2. Auflage, Bern/Göttingen: Huber.

### Autor

Maik Arnold, Dipl.-Kfm.  
Wissenschaftlicher Mitarbeiter an der Juniorprofessur  
Interkulturelles Training, Technische Universität Chemnitz

### Kontakt

Juniorprofessur Interkulturelles Training  
TU Chemnitz, 09107 Chemnitz  
Tel. 0371–531 29 70, Fax 0371–531 29 33  
maik.arnold@phil.tu-chemnitz.de  
www.qualitative-evaluation.de