

Standards für Evaluation

Hochschulen

Verwaltungs-
handeln

Umweltpolitik

Schulen

Gesundheitswesen

Soziale
Dienstleistungen

Betriebliche und
berufliche Bildung

Entwicklungs-
zusammenarbeit

Forschungs-,
Technologie- und
Innovationspolitik

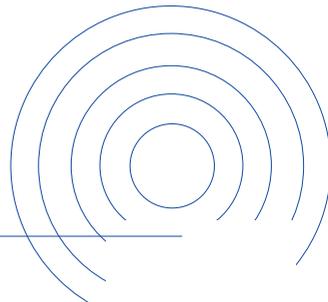
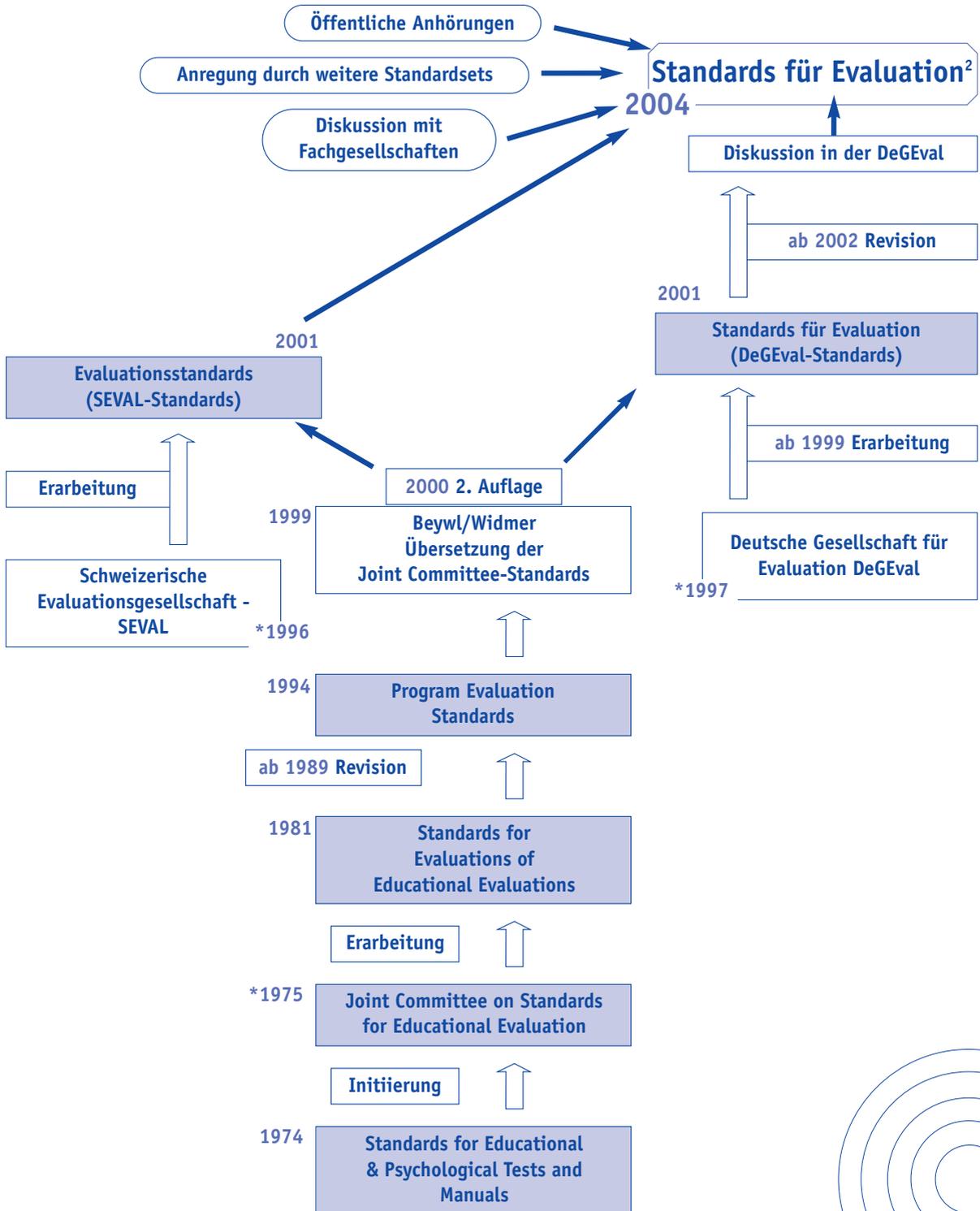
A decorative graphic in the bottom-left corner consisting of several concentric white circles of varying diameters, centered on the page.

Kommunikation
und Medien

Aus- und Weiterbildung

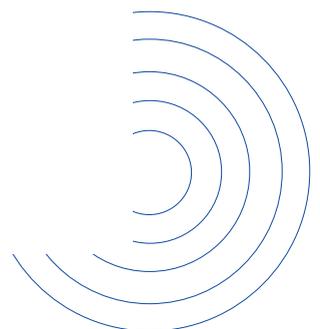
Stadt- und Regional-
entwicklung

ENTSTEHUNG UND PERSPEKTIVE DER DEGEVAL-STANDARDS



Deutsche Gesellschaft für Evaluation

Standards für Evaluation



Die Deutsche Bibliothek – CIP Einheitsaufnahme

**Deutsche Gesellschaft für Evaluation e.V.
Standards für Evaluation
Köln, 2002, ISBN 3-00-009022-3**

Herausgeber: Deutsche Gesellschaft für Evaluation e.V.

Redaktion: Dr. Wolfgang Beywl

Redaktionsassistentin: Anke Frey

Vertrieb:

Geschäftsstelle DeGEval

Postfach 26 01 29

50514 Köln

Das Werk einschließlich seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Herausgebers unzulässig und strafbar. Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Die Vervielfältigung der Kurzfassung der Standards ist ausdrücklich erwünscht und ist ohne Rücksprache zulässig. Von der Veröffentlichung sind der DeGEval zwei Belegexemplare zur Verfügung zu stellen.

ISBN 3-00-009022-3

Satz: Wolf Bee/Silke Schaufuß

Einbandentwurf: Wolf Bee/Silke Schaufuß

Druck: Zimmermann-Medien, Köln

Printed in Germany

INHALTS VERZEICHNIS

■ Vorwort	6
■ Preface	7
■ Standards für Evaluation	8
■ Einleitung	12
Zielsetzung der Standards	12
Definition und Formen der Evaluation	13
Anwendung der Standards	16
Entstehung der Standards	17
Weiteres Vorgehen	18
■ Erläuterungen zu den Standards für Evaluation	20
Nützlichkeit	21
Durchführbarkeit	26
Fairness	28
Genauigkeit	31
■ Funktionales Inhaltsverzeichnis	37
■ Transformationstabelle	41
■ Literatur	43
■ Mitarbeiter und Mitarbeiterinnen an der Standardentwicklung	46
Die Mitglieder der Kommission „Standards für Evaluation“	46
Kommentatoren und Kommentatorinnen	46
■ Summary of Evaluation Standards	47



VORWORT

Im Oktober 2001 hat die Mitgliederversammlung der Deutschen Gesellschaft für Evaluation (DeGEval) auf ihrer 4. Jahrestagung einstimmig die „Standards für Evaluation“ verabschiedet. In einer Zeit wachsenden öffentlichen Interesses an Evaluation stellt die DeGEval mit dieser Broschüre eine fachliche Grundlage bereit, die durch die Gesellschaft formuliert, entwickelt und schließlich verabschiedet worden ist. Sie erhofft sich davon eine Verbreitung methodischen und professionellen Wissens über Evaluation und eine Verstärkung des Dialogs über Qualität und Nutzen von Evaluationen. Die vorliegende Broschüre enthält eine deutsche und eine englische Kurzfassung der Standards sowie ausführliche Erläuterungen in deutscher Sprache.

■ Wie in den Erläuterungen ausgeführt, beanspruchen die Standards Gültigkeit für verschiedenste Ansätze der Evaluation, für unterschiedliche Evaluationszwecke sowie eine Vielzahl von Evaluationsfeldern.

■ Die DeGEval-Standards sollen die Qualität von Evaluationen sichern und entwickeln helfen und den öffentlichen und professionellen Dialog darüber fördern. Die Standards geben auch konkrete Hinweise für die Planung und Durchführung von Evaluationen. Darüber hinaus können sie als Orientierungsgrundlage für die Aus- und Weiterbildung in Evaluation genutzt werden und Bezugsrahmen sein für die Evaluation von Evaluationen (Meta-Evaluation). Zur Unterstützung dieser Anwendungen sind in der Umschlagtasche am Ende der Broschüre Einsteckkarten mit einer Checkliste sowie die Zusammenfassung der Standards als Handkarte beigelegt.

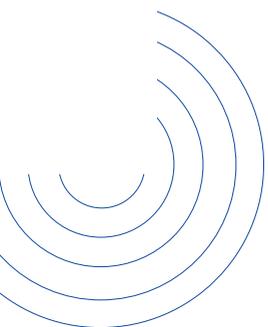
■ Aus Sicht der DeGEval ist es wünschenswert, die Standards für Evaluation selbst zum Gegenstand der kritischen Bewertung zu machen. Institutionen und Personen, die Evaluationen beauftragen, an ihrer Durchführung mitwirken und ihre Ergebnisse nutzen, soll Gelegenheit gegeben werden, zu den DeGEval-Standards Stellung zu nehmen. Die bis 2004 angelegte Revision soll die Passung der Standards auf die Breite der Evaluationsfelder überprüfen sowie die Klarheit der Formulierungen und die Verständlichkeit für die verschiedenen Nutzergruppen optimieren. Dies wird um so besser gelingen, je engagierter sich Politik, Verwaltungen, Wissenschaft, Fachkräfte aus den verschiedensten Evaluationsfeldern und ihre Verbände sowie diejenigen, die als Evaluatoren und Evaluatorinnen tätig sind, aktiv an dieser Diskussion beteiligen. In dieser Broschüre und auf den Internet-Seiten der DeGEval (www.degeval.de) werden Zugänge für eine solche aktive Beteiligung eröffnet, z. B. über elektronische Diskussionsforen, durch schriftliche Stellungnahmen oder Teilnahme an öffentlichen Veranstaltungen zu den Standards.

Wir wünschen uns, dass dieses Diskussionsangebot der DeGEval intensiv genutzt wird und möchten die Leserinnen und Leser dieser Broschüre nachdrücklich um Stellungnahmen bitten.

Für den Vorstand der DeGEval

Prof. Dr. Hans-Dieter Daniel
(1. Vorsitzender)

Dr. Wolfgang Beywl
(Koordinator DeGEval-Standards)



PREFACE

At the Forth Annual Meeting of the German Society for Evaluation (DeGEval) in October of 2001, the general assembly unanimously passed the „Standards for Evaluation“. At a time of growing public interest in evaluation, the present DeGEval booklet provides a technical basis for discussion which has been formulated, developed and finally approved by the members of the society. DeGEval hopes that the Standards will contribute to the methodological and professional knowledge about evaluation and a permanent dialogue about quality and use of evaluations.

■ The booklet contains an English and a German abstract of the Standards, as well as detailed explanations in German language.

■ As explicated in the text, the Standards claim to be applicable to a diverse set of evaluation approaches, various evaluation purposes, and a variety of evaluation fields.

■ The DeGEval Standards are intended to safeguard and develop the quality of evaluations and to promote the public and professional dialogue about this important topic. The Standards give concrete advice on the planning and implementation of evaluations. Moreover, they can be used to guide evaluation training and the evaluation of evaluations (meta-evaluation). On the inside cover at the end of the booklet you may find a checklist of all the 25 Standards. Also enclosed is a removable short version of the Standards. These practical aids should support the above mentioned applications.

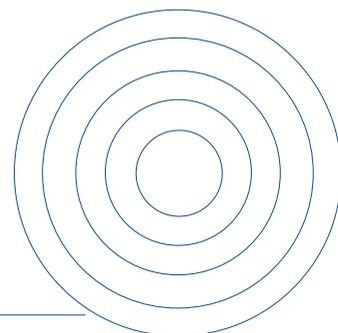
■ DeGEval intends to make the Standards themselves subject to critical assessment. Institutions and individuals who contract evaluations, participate in their conduct or use their results, are most welcome to take position on the DeGEval Standards. The three-year revision process (2002-2004) will allow for an optimization of the expressive clarity, interdisciplinary applicability and comprehensibility of the Standards. The ultimate success of this process will depend upon the active engagement of politicians, administrators, scientists and professionals, including professional associations, from all the fields where evaluation is performed, as well as practicing evaluators. This booklet and the DeGEval website (www.degeval.de) offer a number of venues for such an active participation, e.g., via electronic discussion, written statements, or participation in public events related to the Standards.

We hope that this opportunity for discussion and dialogue initiated by DeGEval will be used intensively. We greatly appreciate feedback from all readers.

For DeGEval board of directors

Prof. Dr. Hans-Dieter Daniel
(President)

Dr. Wolfgang Beywl
(‘Standards’ coordinator)



STANDARDS FÜR EVALUATION

der Deutschen Gesellschaft für Evaluation (DeGEval-Standards)*

Evaluationen sollen vier grundlegende Eigenschaften aufweisen:

Nützlichkeit - Durchführbarkeit - Fairness - Genauigkeit.

NÜTZLICHKEIT **N** Die Nützlichkeitsstandards sollen sicherstellen, dass die Evaluation sich an den geklärten Evaluationszwecken sowie am Informationsbedarf der vorgesehenen Nutzer und Nutzerinnen ausrichtet.

N1 Identifizierung der Beteiligten und Betroffenen

Die am Evaluationsgegenstand beteiligten oder von ihm betroffenen Personen bzw. Personengruppen sollen identifiziert werden, damit deren Interessen geklärt und so weit wie möglich bei der Anlage der Evaluation berücksichtigt werden können.

N2 Klärung der Evaluationszwecke

Es soll deutlich bestimmt sein, welche Zwecke mit der Evaluation verfolgt werden, so dass die Beteiligten und Betroffenen Position dazu beziehen können und das Evaluationsteam einen klaren Arbeitsauftrag verfolgen kann.

N3 Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin

Wer Evaluationen durchführt, soll persönlich glaubwürdig sowie methodisch und fachlich kompetent sein, damit bei den Evaluationsergebnissen ein Höchstmaß an Glaubwürdigkeit und Akzeptanz erreicht wird.

N4 Auswahl und Umfang der Informationen

Auswahl und Umfang der erfassten Informationen sollen die Behandlung der zu untersuchenden Fragestellungen zum Evaluationsgegenstand ermöglichen und gleichzeitig den Informationsbedarf des Auftraggebers und anderer Adressaten und Adressatinnen berücksichtigen.

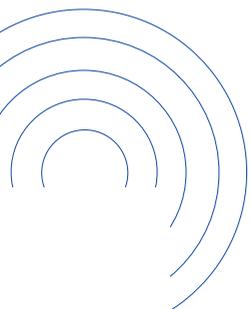
N5 Transparenz von Werten

Die Perspektiven und Annahmen der Beteiligten und Betroffenen, auf denen die Evaluation und die Interpretation der Ergebnisse beruhen, sollen so beschrieben werden, dass die Grundlagen der Bewertungen klar ersichtlich sind.

N6 Vollständigkeit und Klarheit der Berichterstattung

Evaluationsberichte sollen alle wesentlichen Informationen zur Verfügung stellen, leicht zu verstehen und nachvollziehbar sein.

*Eine Überarbeitung der Standards ist bis spätestens Ende 2004 vorgesehen.



N7 Rechtzeitigkeit der Evaluation

Evaluationsvorhaben sollen so rechtzeitig begonnen und abgeschlossen werden, dass ihre Ergebnisse in anstehende Entscheidungsprozesse bzw. Verbesserungsprozesse einfließen können.

N8 Nutzung und Nutzen der Evaluation

Planung, Durchführung und Berichterstattung einer Evaluation sollen die Beteiligten und Betroffenen dazu ermuntern, die Evaluation aufmerksam zur Kenntnis zu nehmen und ihre Ergebnisse zu nutzen.

DURCHFÜHRBARKEIT **D** Die Durchführbarkeitsstandards sollen sicherstellen, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst geplant und ausgeführt wird.

D1 Angemessene Verfahren

Evaluationsverfahren, einschließlich der Verfahren zur Beschaffung notwendiger Informationen, sollen so gewählt werden, dass Belastungen des Evaluationsgegenstandes bzw. der Beteiligten und Betroffenen in einem angemessenen Verhältnis zum erwarteten Nutzen der Evaluation stehen.

D2 Diplomatisches Vorgehen

Evaluationen sollen so geplant und durchgeführt werden, dass eine möglichst hohe Akzeptanz der verschiedenen Beteiligten und Betroffenen in bezug auf Vorgehen und Ergebnisse der Evaluation erreicht werden kann.

D3 Effizienz von Evaluation

Der Aufwand für Evaluation soll in einem angemessenen Verhältnis zum Nutzen der Evaluation stehen.

FAIRNESS **F** Die Fairnessstandards sollen sicherstellen, dass in einer Evaluation respektvoll und fair mit den betroffenen Personen und Gruppen umgegangen wird.

F1 Formale Vereinbarungen

Die Pflichten der Vertragsparteien einer Evaluation (was, wie, von wem, wann getan werden soll) sollen schriftlich festgehalten werden, damit die Parteien verpflichtet sind, alle Bedingungen dieser Vereinbarung zu erfüllen oder aber diese neu auszuhandeln.

F2 Schutz individueller Rechte

Evaluationen sollen so geplant und durchgeführt werden, dass Sicherheit, Würde und Rechte der in eine Evaluation einbezogenen Personen geschützt werden.



F3 Vollständige und faire Überprüfung

Evaluationen sollen die Stärken und die Schwächen des Evaluationsgegenstandes möglichst vollständig und fair überprüfen und darstellen, so dass die Stärken weiter ausgebaut und die Schwachpunkte behandelt werden können.

F4 Unparteiische Durchführung und Berichterstattung

Die Evaluation soll unterschiedliche Sichtweisen von Beteiligten und Betroffenen auf Gegenstand und Ergebnisse der Evaluation in Rechnung stellen. Berichte sollen ebenso wie der gesamte Evaluationsprozess die unparteiische Position des Evaluationsteams erkennen lassen. Bewertungen sollen fair und möglichst frei von persönlichen Gefühlen getroffen werden.

F5 Offenlegung der Ergebnisse

Die Evaluationsergebnisse sollen allen Beteiligten und Betroffenen soweit wie möglich zugänglich gemacht werden.

GENAUIGKEIT **G** Die Genauigkeitsstandards sollen sicherstellen, dass eine Evaluation gültige Informationen und Ergebnisse zu dem jeweiligen Evaluationsgegenstand und den Evaluationsfragestellungen hervorbringt und vermittelt.

G1 Beschreibung des Evaluationsgegenstandes

Der Evaluationsgegenstand soll klar und genau beschrieben und dokumentiert werden, so dass er eindeutig identifiziert werden kann.

G2 Kontextanalyse

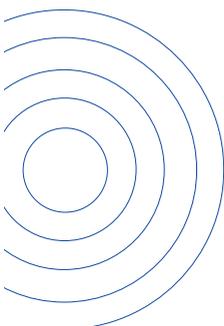
Der Kontext des Evaluationsgegenstandes soll ausreichend detailliert untersucht und analysiert werden.

G3 Beschreibung von Zwecken und Vorgehen

Gegenstand, Zwecke, Fragestellungen und Vorgehen der Evaluation, einschließlich der angewandten Methoden, sollen genau dokumentiert und beschrieben werden, so dass sie identifiziert und eingeschätzt werden können.

G4 Angabe von Informationsquellen

Die im Rahmen einer Evaluation genutzten Informationsquellen sollen hinreichend genau dokumentiert werden, damit die Verlässlichkeit und Angemessenheit der Informationen eingeschätzt werden kann.



G5 Valide und reliable Informationen

Die Verfahren zur Gewinnung von Daten sollen so gewählt oder entwickelt und dann eingesetzt werden, dass die Zuverlässigkeit der gewonnenen Daten und ihre Gültigkeit bezogen auf die Beantwortung der Evaluationsfragestellungen nach fachlichen Maßstäben sichergestellt sind. Die fachlichen Maßstäbe sollen sich an den Gütekriterien quantitativer und qualitativer Sozialforschung orientieren.

G6 Systematische Fehlerprüfung

Die in einer Evaluation gesammelten, aufbereiteten, analysierten und präsentierten Informationen sollen systematisch auf Fehler geprüft werden.

G7 Analyse qualitativer und quantitativer Informationen

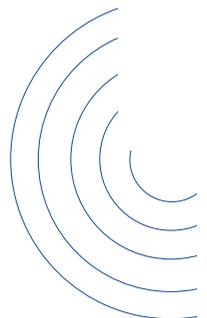
Qualitative und quantitative Informationen einer Evaluation sollen nach fachlichen Maßstäben angemessen und systematisch analysiert werden, damit die Fragestellungen der Evaluation effektiv beantwortet werden können.

G8 Begründete Schlussfolgerungen

Die in einer Evaluation gezogenen Folgerungen sollen ausdrücklich begründet werden, damit die Adressaten und Adressatinnen diese einschätzen können.

G9 Meta-Evaluation

Um Meta-Evaluationen zu ermöglichen, sollen Evaluationen in geeigneter Form dokumentiert und archiviert werden.



EINLEITUNG

Das vorliegende Papier besteht aus einer Einleitung, den nach vier Gruppen geordneten "Standards für Evaluation" mit Erläuterungen und einem Anhang.

■ In der Einleitung dieses Papiers wird zunächst die Zielsetzung der „Standards für Evaluation“ beschrieben. Anschließend wird ein Überblick gegeben über Definition und Formen der Evaluation. Ein kurzer Abschnitt skizziert die Art und Weise, wie die Standards angewendet werden sollen. Abschließend wird der Entstehungsprozess der Standards dargestellt und ein Ausblick auf ihre gewünschte Weiterentwicklung gegeben.

■ Evaluationen sollen folgende vier grundlegende Eigenschaften aufweisen: Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit. Den so benannten vier Gruppen sind die insgesamt 25 Einzelstandards zugeordnet.

■ Im Anhang findet sich zunächst ein funktionales Inhaltsverzeichnis, das die 25 Einzelstandards zehn zentralen Aufgaben zuordnet, die im Rahmen einer Evaluation zu leisten sind. Dieses Verzeichnis kann auch zur Planung und Ablaufkontrolle von Evaluationen genutzt werden. Anschließend wird durch eine „Transformationstabelle“ ein Vergleich der DeGEval-Standards mit zwei benachbarten Standard-Sets aus den USA und der Schweiz ermöglicht. Am Schluss findet sich ein Literaturverzeichnis.

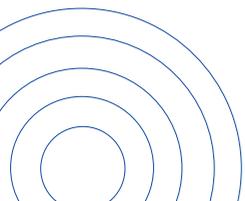
1. ZIELSETZUNG der Standards

Evaluation hat sich auch in Deutschland in den letzten Jahren stark entwickelt. Sie gewinnt in den meisten Bereichen der Gesellschaft, der Politik, der Wirtschaft und der Wissenschaft weiter an Bedeutung.¹

■ Vor diesem Hintergrund hat die Deutsche Gesellschaft für Evaluation (DeGEval) die „Standards für Evaluation“ formuliert. Die 25 Standards bestehen aus einem Standard-Namen und einer Standard-Formulierung. Diese besteht aus bis zu drei Sollensaussagen. Sie richten sich sowohl an Evaluatoren und Evaluatorinnen als auch an Personen und Einrichtungen, die Evaluationen in Auftrag geben, sowie an Beteiligte und Betroffene im Bereich des Evaluationsgegenstandes. Die Standards sollen die Qualität von Evaluationen sichern und entwickeln helfen. Sie sollen als Dialoginstrument und fachlicher Bezugspunkt für einen Austausch über die Qualität von professionellen Evaluationen dienen. Außerdem sollen sie Orientierung bei Planung und Durchführung von Evaluationen geben, Anknüpfungspunkte bieten für die Aus- und Weiterbildung in Evaluation und für die Evaluation von Evaluationen (Meta-Evaluation) sowie Transparenz über Evaluation als professionelle Praxis gegenüber einer breiteren Öffentlichkeit schaffen. Die Standards können bei der Kommunikation von Evaluatoren und Evaluatorinnen mit den Auftraggebern, Adressaten und Adressatinnen sowie einem weiten Kreis von Beteiligten und Betroffenen nützlich sein. Dies gilt für Zeitpunkte vor, während und nach der Durchführung konkreter Evaluationen.

■ Mit dieser Fassung der Standards wird ausdrücklich keine verbindliche Grundlage für Akkreditierung oder Zertifizierung von Personen oder Einrichtungen, die Evaluationen anbieten oder durchführen, oder auch von Evaluationsfortbildungen angestrebt.

12 ¹Als Überblicke zu unterschiedlichen Zeitpunkten vgl. Hellstern/Wollmann (1994) und Stockmann (2000); für die Perspektive des Bundesrechnungshofes vgl. v. Wedel (1998); für die Schweiz vgl. Bussmann/Klöti/ Knoepfel (1997); einen umfassenden Überblick über die Entwicklung in einem einzelnen Evaluationsfeld gibt Grohmann (1997).



2. DEFINITION und Formen der Evaluation

Evaluation ist die systematische Untersuchung des Nutzens oder Wertes eines Gegenstandes. Solche Evaluationsgegenstände können z. B. Programme, Projekte, Produkte, Maßnahmen, Leistungen, Organisationen, Politik, Technologien oder Forschung sein. Die erzielten Ergebnisse, Schlussfolgerungen oder Empfehlungen müssen nachvollziehbar auf empirisch gewonnenen qualitativen und/oder quantitativen Daten beruhen.

■ Aufgrund der Vielfalt von Evaluationen, ihrer Anwendungsbereiche, Aufgaben und zugrunde liegenden Konzepte, gibt es darüber hinaus zahlreiche Definitionen, die sich jeweils in einzelnen Aspekten unterscheiden. Der Begriff „Evaluation“ wird zunehmend auch in der Alltagssprache verwendet, jedoch häufig nicht in diesem professionellen Sinn, sondern es wird damit lediglich ausgedrückt, dass etwas in irgendeiner Weise geprüft, bewertet oder beurteilt wurde (Kromrey 2001). Professionelle Evaluation zielt auf eine nachvollziehbare Bewertung ihres jeweiligen Gegenstandes, so dass dessen Güte und/oder Nutzen² möglichst genau bestimmt werden kann. Indem sie den Gegenstand systematisch beschreibt, bereitet sie dessen Bewertung vor. Diese kann dann entweder von den Evaluatoren oder Evaluatorinnen selbst, von denjenigen, die Evaluationen in Auftrag geben, von den für den Evaluationsgegenstand Verantwortlichen (z. B. Führungskräfte oder Mitarbeiter und Mitarbeiterinnen in Organisationen oder Programmen) oder auch von anderen Beteiligten und Betroffenen vorgenommen werden. Auch das gemeinsame Bewerten durch verschiedene Gruppierungen ist möglich.

■ Evaluation erfolgt systematisch, d. h. eine Evaluation wird auf bestimmte Evaluationszwecke³ hin ausgerichtet und so geplant, dass ihre Durchführung durch die gesetzten Zwecke angeleitet wird. Die notwendigen Informationen und Daten sollen mit angemessenem Aufwand erhoben und die Bewertungsgrundlagen so herausgearbeitet werden, dass Bewertungen darauf aufbauen können.

■ Das für den jeweiligen Gegenstandsbereich gültige Fachwissen und die Erkenntnisse der wissenschaftlichen Forschung sind bei der Planung einer Evaluation und bei der Gewinnung und Aufbereitung der Daten angemessen zu berücksichtigen.

■ Evaluation ist datengestützt und arbeitet mit einer Bandbreite empirisch-wissenschaftlicher Methoden. Dies sind insbesondere die quantitativen und qualitativen Methoden der empirischen Sozialforschung.

² Mit ‚Güte und Nutzen‘ wird Bezug genommen auf das Begriffspaar ‚merit‘ und ‚worth‘, das auch in der Definition der Standards des Joint Committee enthalten ist. Es geht zurück auf eine Unterscheidung von Guba/Lincoln (1981), die unter Güte die ‚intrinsische‘ Qualität eines Evaluationsgegenstandes verstehen (z. B. Stringenz und fachwissenschaftliche Absicherung eines Konzeptes), unter Nutzen hingegen den Gebrauchswert für bestimmte Anwender und Anwenderinnen, in bestimmten Situationen, zu bestimmten Zeitpunkten. Güte ist in diesem Verständnis zeitlich und räumlich überdauernd, Nutzen hingegen nur situativ bestimmbar. Konsequenterweise ist ‚Nutzen‘ im Plural zu gebrauchen, also mehrere Nutzen (evtl. konkurrierende) für verschiedene Anwender und Anwenderinnen.

³ Die intendierten Verwendungen von Evaluationen bzw. ihrer Ergebnisse werden in diesem Text als „Evaluationszwecke“ bezeichnet. Die Kommission hat sich bewusst für den Terminus „Evaluationszweck“ entschieden, um eine deutliche Unterscheidung gegenüber „Zielen“ vorzunehmen, die typischerweise im Bereich des Evaluationsgegenstandes formuliert sind (z. B. als Programmziele) und z. B. für „zielgeführte“ Evaluationen wichtige Orientierungspunkte bzw. Steuerungsfaktoren darstellen. So ist ein typisches Programmziel, dass bestimmte erwünschte Zustände bei den Zielgruppen des Programms ausgelöst werden sollen. Hingegen besteht die Zwecksetzung einer Evaluation zum Beispiel darin, ein bestehendes Programm zu verbessern. Diese terminologische Unterscheidung soll es erleichtern, über Evaluationszwecke und über Ziele im Bereich des Evaluationsgegenstandes klar zu kommunizieren.



■ Evaluationen sollen in ihren Zwecken, den angewandten Methoden, der Datenbasis, ihrer Auswertung und ihren Bewertungsgrundlagen nachvollziehbar und kritisierbar sein. Ihre Grundlagen und Schritte sollen dokumentiert werden. Es sollen schriftliche Berichte erstellt oder es soll in anderer geeigneter Form über die Evaluation und ihre Ergebnisse berichtet werden. Evaluationen sollen sich an fachlichen Standards orientieren, die nachfolgend beschrieben sind. Evaluatoren und Evaluatorinnen sollen sich mit ihrer Arbeit der Kritik der Fachöffentlichkeit stellen.

■ Evaluationen können unterschiedliche Leistungsschwerpunkte haben und damit unterschiedliche Zwecksetzungen anstreben. „Formative Evaluation“, deren Leistung es ist, die Gestaltung des Evaluationsgegenstandes zu begleiten, zielt vorrangig auf Verbesserungen. Sie soll den Verantwortlichen und Beteiligten helfen, den Evaluationsgegenstand und seinen Nutzen zu verbessern und Ressourcen möglichst gut einzusetzen.

■ Davon zu unterscheiden ist „summative Evaluation“, deren Leistung es ist, zu einem Evaluationsgegenstand eine zusammenfassende Bilanz zu ziehen. Häufig ist ihre Zwecksetzung, grundlegende Entscheidungen über den Evaluationsgegenstand zu ermöglichen. Neben Verbesserung und Entscheidungsvorbereitung ist der Gewinn von Erkenntnissen ohne unmittelbare Verwendungsabsicht eine weitere mögliche Zwecksetzung von Evaluationen (vgl. Standard N2). Dazu kann sowohl formative als auch summative Evaluation dienen.⁴

■ Evaluationen werden in sehr unterschiedlichen Bereichen durchgeführt. Einen Eindruck davon geben die Arbeitskreise der DeGEval mit ihren verschiedenen thematischen Schwerpunkten.⁵ Der finanzielle, zeitliche und räumliche Umfang von Evaluationen, die Zahl der am Evaluationsvorhaben beteiligten Personen – sei es als Evaluatoren bzw. Evaluatorinnen oder als Adressaten und Adressatinnen – kann sehr unterschiedlich sein. So kann z. B. die Befragung weniger Personen wertvolle und ausreichende Grundlagen für Bewertungen, Verbesserungen und Entscheidungen liefern. In einem anderen Fall kann dazu aber auch ein langfristig geplanter und wesentlich aufwendigerer Prozess der Organisations- oder Programmentwicklung notwendig sein. Beide Beispiele können den Standards für Evaluation entsprechen.

Bei allen Unterschieden weisen alle Evaluationen auf:⁶

- Auftraggeber⁷
- Finanziers⁸
- Nutzer und Nutzerinnen, Adressatinnen und Adressaten, Beteiligte und Betroffene⁹

⁴ Vgl. zum Begriffspaar "formativ" – "summativ" und zu weiteren Themen der Evaluations-Fachsprache Scriven (1991).

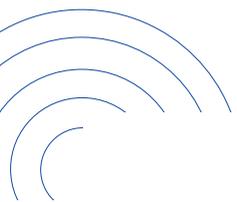
⁵ Zum Zeitpunkt der Schlussfassung dieser Standards (im Sommer 2001) waren dies: Evaluation an Hochschulen, Evaluation an Schulen, Evaluation im Umweltbereich, Evaluation in der beruflichen und betrieblichen Bildung, Evaluation von Entwicklungspolitik, Evaluation von Forschungs-, Technologie- und Innovationspolitik, Medienevaluation, Evaluation von sozialen Dienstleistungen, Evaluation im Gesundheitswesen, Verwaltung und Evaluation, Evaluation von Strukturpolitik. Diese Aufzählung ist nicht abschließend sondern illustrativ zu verstehen; vgl. auch <http://www.degeval.de/arbeitskreise.htm>.

⁶ Einen Teil dieser Aspekte behandeln Wottawa/Thierau (1998) im Kapitel "Gestaltungsaspekte von Evaluationsstudien", S. 55-66.

⁷ Die Evaluation beauftragen kann insbesondere im Fall der Selbstevaluation auch das Evaluationsteam selbst.

14 ⁸ Diese können, müssen aber nicht mit den Auftraggebern identisch sein.

⁹ Zuerst sind "Nutzer und Nutzerinnen" der Evaluation genannt. Diejenigen, welche die Evaluation und ihre Ergebnisse wie im vorgesehen Evaluationsplan nutzen, gehören zu den "Adressatinnen und Adressaten". Nicht alle Adressatinnen und Adressaten nutzen tatsächlich. Schließlich sind "Beteiligte und Betroffene" angesprochen. "Beteiligte" sind Personen, Gruppen oder auch Organisationen, die in bezug auf den Evaluationsgegenstand eine aktive Rolle spielen, z. B. Geldgeber des Programms, Akteure, die das Programm steuern oder die im Rahmen des Evaluationsgegenstandes gegen Entgelt oder ehrenamtlich tätig sind. "Betroffene" sind insbesondere Personen mit wenig Einfluss; sie finden sich nicht



- durchführende Evaluationsteams¹⁰
- Evaluationszwecke
- Evaluationsgegenstände
- Orte der Evaluation
- Werte, die der Evaluation und ihrem Gegenstand zu Grunde liegen
- Evaluationsfragestellungen
- Evaluationspläne
- Untersuchungsmethodiken (bzgl. Erhebung, Auswertung und Interpretation)
- Bewertungen
- sowie - i.d.R. schriftlich fest gehaltene - Ergebnisse

■ Es gibt viele verschiedene Ansätze professioneller Evaluation, die in vergleichenden Studien synoptisch aufgearbeitet sind (Wulf 1972, Beywl 1988, Owen/Rogers 1999, Uhl 1999, Kromrey 2000, Uhl 2000, Stufflebeam 2001), unterscheidbar z. B. danach,

- welche erkenntnistheoretischen Fundamente die Evaluation aufweist (kritisch-rational hypothesenprüfend, pragmatisch, konstruktivistisch usw.)
- welche Partizipationsvorstellungen mit Evaluation und ihrer Nutzung verbunden sind (Evaluationsansätze mit konkurrierenden „Anwaltsteams“, parlamentarisch-repräsentative und damit auf legale Entscheider zugeschnittene, stark oder gemäßigt beteiligungsorientierte, basisdemokratische, ...)
- welche Bezugswissenschaften (dominant) heran gezogen werden (Wirtschaftswissenschaft, Soziologie, Politikwissenschaft, Erziehungswissenschaft, Ethnologie, Ingenieurwissenschaft, ...)
- welche Zwecksetzungen die Evaluation verfolgen soll (z. B. Vorbereitung von Entscheidungen über den Evaluationsgegenstand, Verbesserung von Organisationen oder Programmen, Unterstützung von Lern- und Reflexionsprozessen, Erweiterung der Erkenntnisse über einen bestimmten Gegenstandsbereich)
- welche Leistungen die Evaluation erbringen soll (präformativ – entwickelnd, formativ – gestaltend, summativ – bilanzierend)
- welches die zentralen Steuerungsfaktoren der Evaluation sind (z. B. die explizierten Zwecksetzungen der Evaluation; Werte, die von Beteiligten und Betroffenen im Bereich des Evaluationsgegenstandes getragen werden; Ziele des evaluierten Programms; Ad-hoc- oder systematisch hergeleitete Hypothesen; Interessen von Beteiligten und Betroffenen; Kosten-Nutzen-Relationen)

selten unter den Zielgruppen z. B. eines Programms; oder es sind von der Nutzung eines Programms Ausgeschlossene oder durch das Programm Benachteiligte, mitunter ohne von der Existenz des Programms zu wissen. Besonders im Rahmen partizipativer Evaluationsansätze kann angestrebt sein, Betroffene zu Beteiligten zu machen. Die Trennlinie zwischen Beteiligten und Betroffenen ist nicht scharf zu ziehen. Die genannten Personenkreise überschneiden sich oft: Mitarbeiter und Mitarbeiterinnen von Organisationen, in denen evaluiert wird, sind in der Regel beteiligt bzw. betroffen, können gleichzeitig Adressatinnen und Adressaten der Evaluation sein und diese dann auch tatsächlich nutzen. Ein Gegenbeispiel ist eine Stiftung, die Evaluationen fördert und eine Adressatin des Schlussberichtes ist, die jedoch keine direkten Interessen am Evaluationsgegenstand hat, also nicht zum Kreis der Beteiligten und Betroffenen zählt. Ähnlich verhält es sich, wenn die allgemeine oder die wissenschaftliche Öffentlichkeit Adressatin ist. Es handelt sich bei diesen Benennungen um analytische Trennungen. Sie signalisieren, ob vorrangig Personen/Organisationen in Bezug auf den Evaluationsgegenstand oder in Bezug aber auf die Evaluation angesprochen sind. Dadurch sollen die für die Evaluation Verantwortlichen angeregt werden zu klären, wer in welcher Rolle durch die verschiedenen Aktivitäten im Rahmen der Evaluation angesprochen ist.

¹⁰In vielen Fällen sind Teams aus zwei und mehr Personen für eine Evaluation verantwortlich. Es kommt auch vor, dass ein Evaluator oder eine Evaluatorin allein tätig ist. Soweit im folgenden Text vom Evaluationsteam gesprochen wird, sind auch solche mit gemeint. Vgl. Fußnote 3!



- auf welche Phase der Entwicklung des Evaluationsgegenstandes (Entwicklungs- und Routinephase) sich die Evaluation bezieht (proaktiv, klärend, interaktiv, dokumentierend, Wirksamkeit überprüfend bzw. ex ante, on going, ex post)
- wie umfangreich und komplex der Evaluationsgegenstand ist
- welche Dimension des Evaluationsgegenstandes vorrangig betrachtet wird (Kontext, Struktur, Konzept, Input, Prozess oder Wirkung)
- ob die Evaluation extern durchgeführt wird, d. h. durch ein von außerhalb der Organisation beauftragtes Evaluationsteam, oder von den Personen, die für den Evaluationsgegenstand verantwortlich sind (Selbstevaluationen), oder ob es sich um eine interne Evaluation handelt, die ein Mitglied der Organisation durchführt, das dabei nicht für den zu evaluierenden Gegenstand verantwortlich ist

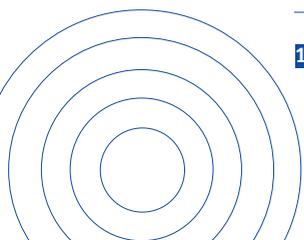
■ Die Liste der aufgeführten Unterscheidungsdimensionen ist ebenso wenig vollständig wie die in den einzelnen Dimensionen genannten Ausprägungen. Einzelne Evaluationen sind in der Regel Mischformen und Kombinationen aus den verschiedenen aufgeführten Ausprägungen. Wichtig ist, dass die konkrete Ausgestaltung der Evaluation den Zwecken der Evaluation, den Besonderheiten des Evaluationsgegenstandes, den konkreten Bedingungen und Möglichkeiten und den zur Verfügung stehenden Ressourcen (Zeit und Geld) so gut wie möglich angepasst ist und auf diese hin geplant und zugeschnitten wird.

■ Es gibt auch Selbstevaluationen, die dadurch gekennzeichnet sind, dass sie von Personen aus dem Kreis der für den Evaluationsgegenstand Verantwortlichen selbst durchgeführt und oft auch selbst beauftragt werden (Heiner 1998, v. Spiegel 1993, Allgäuer 1997, Buhren/Killus/Müller 1998). Dabei wird oft die Hilfe von externen oder internen Evaluationsteams in Anspruch genommen. Die nachfolgend vorgestellten Standards haben für Selbstevaluationen nur bedingt Geltung, insbesondere da dort häufig andere Rollen- und Interessenkonstellationen gegeben sind. Im Zusammenhang mit der Weiterentwicklung und Ausdifferenzierung der "Standards für Evaluation" wird die DeGEval die Frage von Standards, die auch oder speziell für Selbstevaluationen gelten, aufgreifen (Müller-Kohlenberg/Beywl 2001).

3. ANWENDUNG der Standards

Die DeGEval-Standards sollen der Sicherung und Entwicklung der Qualität von Evaluationen dienen. Dazu formulieren sie wichtige zu beachtende Aspekte und anzustrebende Ziele. Sie sollen als Orientierung bei der Durchführung und Bewertung von Evaluationen dienen. Entscheidend ist ihre Umsetzung. Sie kann nicht schematisch erfolgen. Mit den Qualitätsstandards ist nicht die Absicht verbunden, eine Evaluation, bei der ein bestimmter Standard nicht auf eine ganz bestimmte Weise erfüllt wurde, abzuwerten. Auch wird es Evaluationen geben, in denen einzelne Standards nicht anwendbar sind. Die Nicht-Anwendung eines Standards soll dann kurz begründet werden.

■ Die Frage bei der Bewertung der Qualität von Evaluationen soll sein, ob die anwendbaren Stan-



dards bei Planung und Durchführung der Evaluation berücksichtigt wurden und versucht wurde, ihnen im Rahmen der konkreten Bedingungen möglichst gerecht zu werden.

■ Es wird häufig nötig sein, zwischen Vor- und Nachteilen bestimmter Handlungsalternativen in der Evaluation abzuwägen. Es ist nicht immer möglich, allen Standards in gleicher Weise gerecht zu werden. Gelegentlich werden sich die Anforderungen widersprechen.¹² Es ist Aufgabe des Evaluations-teams und aller an der Evaluation Mitwirkenden, eine den jeweiligen Zwecken und Rahmenbedingungen der konkreten Evaluation angemessene Lösung zu finden.

■ Die Standards gelten für Evaluationen im Allgemeinen, nicht jedoch für die Evaluation von Personen, wie sie z. B. im Rahmen von Leistungsbemessungsprozessen oder in der Beurteilung von Mitarbeitern und Mitarbeiterinnen durchgeführt werden (Ausschluss von Personalevaluation).

4. ENTSTEHUNG der Standards

Die Deutsche Gesellschaft für Evaluation hat im Jahr 2000 eine Mitgliederbefragung mit dem Ziel durchgeführt, ein klares Meinungsbild für die Erarbeitung von Standards für Evaluationen bzw. Leitlinien für Evaluatoren und Evaluatorinnen zu erhalten. Die deutliche Mehrheit plädierte dafür, vorrangig Standards für Evaluation zu erarbeiten. Viele Mitglieder wünschten sich entweder zusätzlich dazu oder auch erstrangig Leitlinien, vergleichbar denen der American Evaluation Association (1999/1995) oder der Canadian Evaluation Society (o. J.), die das Handeln, die Haltung und die Kompetenz von Evaluatoren und Evaluatorinnen thematisieren. Schließlich fasste die DeGEval auf ihrer Mitgliederversammlung im Jahr 2000 in Berlin den Beschluss, Standards für Evaluation zu formulieren. Sie beauftragte den Vorstand, ein Arbeitsprogramm zu erstellen und eine Kommission einzurichten.

■ Der Beschluss der Mitgliederversammlung 2000 und das Arbeitsprogramm enthalten den Auftrag, die DeGEval-Standards an die Standards des Joint Committee on Standards for Educational Evaluation (JC-Standards 2000/1994) und die mit diesen verwandten Standards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards 2001) anzuschließen. Die Standards der DeGEval bauen damit auf einer inzwischen über 20-jährigen Erfahrung auf. Es wird auf eine vollständige Neuformulierung und Neugliederung von Standards bewusst verzichtet, um den internationalen Erfahrungsaustausch zu erleichtern und um die in der us-amerikanischen und der schweizerischen Evaluationsprofession vorliegenden begleitenden Materialien und Fachveröffentlichungen nutzbar zu machen (ibs. Joint Committee 2000; vgl. auch das Schaubild auf S. 2 dieser Broschüre).

■ Der Vorschlag für die „Standards für Evaluation“ der DeGEval wurde von einer Kommission erstellt. Diese bestand aus sechs Evaluatoren und einer Evaluatorin sowie zwei Vertretern von Auftraggebern. Die neun Kommissionsmitglieder stammen aus verschiedenen Anwendungsfeldern und wissenschaftlichen Disziplinen. Der Entwurf wurde mehrfach umgearbeitet und von insgesamt 13 fachkundigen Kommentatoren und Kommentatorinnen begutachtet (siehe die Listen im Anhang). Die der Mitgliederversammlung zur Abstimmung vorgelegte Version wurde schließlich auch vom Vorstand der

¹² Vgl. zu solchen Dilemmata Hager/Patry/Brezing (2000).

DeGEval zur Kenntnis genommen und der Mitgliederversammlung zur Beschlussfassung empfohlen. Vor dieser Mitgliederversammlung wurden die „Standards für Evaluation“ allen Mitgliedern der DeGEval zur Kenntnis gegeben, damit diese sich eine Meinung bilden und an der abschließenden Beschlussfassung informiert teilnehmen können. Mit dem einstimmig erfolgten Beschluss der Mitgliederversammlung am 4. Oktober 2001 wurde der Vorstand beauftragt, die Standards für Evaluation zusammen mit den Erläuterungen, welche die Kommission erarbeitet hat, zu verbreiten und in die fachliche und öffentliche Diskussion einzubringen.

5. WEITERES Vorgehen

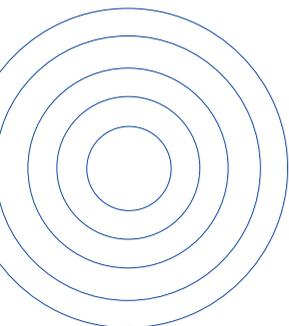
Anschließend an den Beschluss der Mitgliederversammlung sollen die Standards innerhalb der DeGEval weiter breit diskutiert sowie mit wissenschaftlichen Gesellschaften und Berufs- und Fachverbänden verschiedenster Disziplinen sowie weiteren an Evaluation interessierten Zusammenschlüssen abgestimmt werden.

■ In einigen Evaluationsfeldern existieren spezifische Regelwerke, die unabhängig von den Standards sowohl im internationalen als auch im deutschsprachigen Raum Geltung haben. Beispielfhaft genannt seien die DAC-Principles for Evaluation of Development Assistance (OECD 1998) oder die Public Management Service Best Practice Guidelines for Evaluation (PUMA Guidelines der OECD 1998a). An der us-amerikanischen Originalausgabe der Standards des Joint Committee (2000) orientiert sich das MEANS-Handbook Nr. 5 – Quality Assessment of Evaluation Reports, European Union – Regional Policy and Cohesion (1996). Für spezifische Evaluationsfelder erarbeitet worden sind die Leitlinien für Evaluation in der Drogenprävention (EMCDDA 1998), die Empfehlungen für die PR-Evaluation (DPRG 2000) oder die Empfehlungen für die Evaluation in der Jugendhilfe (Beywl u.a. 2001).

■ Die DeGEval und ihre Arbeitskreise sind aufgefordert, die Verbindungen der „Standards für Evaluation“ mit diesen und weiteren einschlägigen Regelwerken zu diskutieren und fachliche Stellungnahmen zu erarbeiten. Im Revisionsprozess der Standards sollen solche bestehende und in Entwicklung befindliche Regelwerke Berücksichtigung finden. Beim Abgleich mit politikfeldspezifischen Standards kommt den jeweiligen Arbeitskreisen der DeGEval eine besondere Rolle zu.

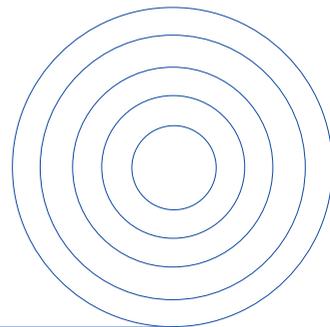
■ Es ist wünschenswert, dass eine revidierte Fassung der Standards für Evaluation ausführliche Erläuterungen enthält und auch Hinweise für die praktische Anwendung der einzelnen Standards, z. B. in Form von Leitfragen. Auch auf mögliche Fehlanwendungen soll hingewiesen werden und es sollen exemplarische Fallbeispiele dargestellt und besprochen werden, welche die Anwendung des jeweiligen Standards demonstrieren und nachvollziehbar machen. Schließlich soll ein Glossar mit der Definition zentraler Begriffe und der Erläuterung von Fachtermini eingeschlossen sein.¹³

¹³ Vorbild hierfür kann die Buchveröffentlichung des Joint Committee (2000) sein.



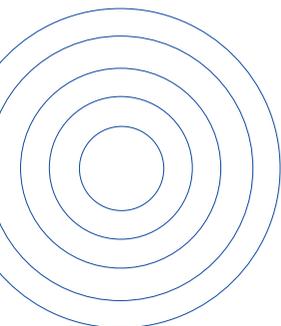
Anregungen und Stellungnahmen zu den Standards sind stets willkommen. Sie sind zu richten an die Geschäftsstelle der Deutschen Gesellschaft für Evaluation, vorzugsweise per E-Mail (standards@degeval.de). Die aktuelle Postanschrift steht unter „Kontakt“ auf den Internet-Seiten der DeGEval (<http://www.degeval.de>).

Auf diesem Weg soll in absehbarer Zeit ein breit getragenes Dokument für die Qualitätssteuerung und –einschätzung von Evaluationen vorliegen. Der Vorstand der DeGEval wird fortlaufend dafür werben, dass sich an Evaluation Interessierte auch außerhalb der DeGEval mit den Standards auseinandersetzen und Stellung nehmen. Sobald eine gewisse Zahl von Anregungen und Kommentaren vorliegt, wird der Vorstand ein Verfahren zur Revision der Standards einleiten. Dieses soll bis Ablauf des Jahres 2004 abgeschlossen sein.



ERLÄUTERUNGEN ZU DEN STANDARDS FÜR EVALUATION

■ N Nützlichkeit	21
N1 Identifizierung der Beteiligten und Betroffenen	21
N2 Klärung der Evaluationszwecke	22
N3 Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin	22
N4 Auswahl und Umfang der Informationen	23
N5 Transparenz von Werten	23
N6 Vollständigkeit und Klarheit der Berichterstattung	24
N7 Rechtzeitigkeit der Evaluation	24
N8 Nutzung und Nutzen der Evaluation	25
■ D Durchführbarkeit	26
D1 Angemessene Verfahren	26
D2 Diplomatisches Vorgehen	26
D3 Effizienz von Evaluation	27
■ F Fairness	28
F1 Formale Vereinbarungen	28
F2 Schutz individueller Rechte	28
F3 Vollständige und faire Überprüfung	29
F4 Unparteiische Durchführung und Berichterstattung	29
F5 Offenlegung der Ergebnisse	30
■ G Genauigkeit	31
G1 Beschreibung des Evaluationsgegenstandes	31
G2 Kontextanalyse	31
G3 Beschreibung von Zwecken und Vorgehen	32
G4 Angabe von Informationsquellen	32
G5 Valide und reliable Informationen	33
G6 Systematische Fehlerprüfung	34
G7 Analyse qualitativer und quantitativer Informationen	34
G8 Begründete Schlussfolgerungen	35
G9 Meta-Evaluation	35



NÜTZLICHKEIT

Die Nützlichkeitsstandards sollen sicherstellen, dass die Evaluation sich an den geklärten Evaluationszwecken sowie am Informationsbedarf der vorgesehenen Nutzer und Nutzerinnen ausrichtet.

N 1 Identifizierung der Beteiligten und Betroffenen

Die am Evaluationsgegenstand beteiligten oder von ihm betroffenen Personen bzw. Personengruppen sollen identifiziert werden, damit deren Interessen geklärt und so weit wie möglich bei der Anlage der Evaluation berücksichtigt werden können.

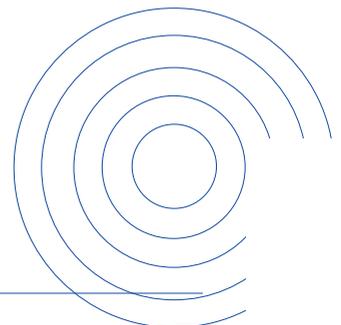
Diese Akteure, die im Umfeld einer Evaluation und ihres Gegenstandes berücksichtigt werden sollen, können durch Fragen wie die folgenden erschlossen werden:

- Wer hat über die Zukunft des Evaluationsgegenstandes zu entscheiden?
- Wer ist verantwortlich für die Konzeption und die Gestaltung des Evaluationsgegenstandes?
- Wer ist an der praktischen Umsetzung des zu untersuchenden Gegenstandes beteiligt?
- Wer soll durch den Evaluationsgegenstand direkt oder indirekt erreicht werden (Zielgruppen und deren soziales Umfeld)?

■ Die so identifizierten Personen, Gruppen und Organisationen werden als „Beteiligte und Betroffene“ bezeichnet. Es soll dabei auch festgestellt werden, wem durch den zu evaluierenden Gegenstand Nachteile erwachsen können oder wer z. B. von einem Programm ausgeschlossen wird. Soweit dies die finanziellen und zeitlichen Bedingungen zulassen, sollen die identifizierten Beteiligten und Betroffenen den Evaluationszwecken entsprechend in die Planung und Durchführung der Evaluation einbezogen werden (Heiner 1998).

■ Es ist besonders wichtig, die Informationsbedürfnisse der Adressatinnen und Adressaten der Evaluation und ihrer Ergebnisse zu ermitteln, sie bei der Klärung der Zwecksetzungen und Fragestellungen zu beteiligen und die Evaluation entsprechend auszurichten. Dabei sollen auch weitere an den Evaluationsergebnissen interessierte Gruppen berücksichtigt werden (z. B. Entscheidungsträger, welche ähnliche Projekte planen, Fachöffentlichkeit, allgemeine Öffentlichkeit).

■ Eine Evaluation, welche die Beteiligten und Betroffenen mitgestaltet und die sich nach deren Informationsbedürfnissen ausrichtet, hat beste Aussichten, von den Adressaten und Adressatinnen beachtet und tatsächlich genutzt zu werden.



N2 Klärung der Evaluationszwecke

Es soll deutlich bestimmt sein, welche Zwecke mit der Evaluation verfolgt werden, so dass die Beteiligten und Betroffenen Position dazu beziehen können und das Evaluationsteam einen klaren Arbeitsauftrag verfolgen kann.

Die Zwecke einer Evaluation sollen frühzeitig angesprochen und im Verlauf einer Evaluation ausgehandelt, schriftlich fixiert und so weit wie möglich den Beteiligten und Betroffenen bekannt gemacht werden, damit diese sich in die Evaluationsplanung einbringen können. Die Tätigkeit der beauftragenden Person bzw. Institution und des Evaluationsteams wird durch klare Zwecksetzungen erleichtert, da sie damit eine gute Grundlage haben, um die Verbreitung der Evaluationsergebnisse und deren Nutzung vorzubereiten.

■ Für die Klärung der Zwecksetzungen ist die idealtypische Unterscheidung dreier Hauptzwecke von Evaluationen hilfreich (Patton 1997, S. 79):

- Informationen bereit zu stellen, die geeignet sind, der schrittweisen Gestaltung des Evaluationsgegenstandes, z. B. bezüglich der Optimierung von Konzepten oder Prozessen, Orientierung zu geben
- Informationen bereit zu stellen, die geeignet sind, einer grundlegenden Entscheidung zum Evaluationsgegenstand Orientierung zu geben, z. B. Einführung, Weiterführung, Ausweitung oder Einstellung eines Programms betreffend
- Erkenntnisse bereit zu stellen, welche der öffentlichen und der politischen oder der wissenschaftlichen Diskussion Anregungen geben können

■ Steht bei den ersten beiden Hauptzwecken der instrumentelle Nutzen im Vordergrund, so ist mit dem Erkenntnisgewinn der konzeptionelle Nutzen angesprochen (Cronbach u.a. 1980, S. 112 ff.). Die gleichzeitige und gleichrangige Verfolgung mehrerer Hauptzwecke kann die Nutzung der Evaluation beeinträchtigen. Es bietet sich daher oft an, eine klare Priorität zu setzen und verschiedene Hauptzwecke in zeitlich getrennten Phasen oder arbeitsteilig durch unterschiedliche Evaluationsteams zu bearbeiten.

N3 Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin

Wer Evaluationen durchführt, soll persönlich glaubwürdig sowie methodisch und fachlich kompetent sein, damit bei den Evaluationsergebnissen ein Höchstmaß an Glaubwürdigkeit und Akzeptanz erreicht wird.

Die Glaubwürdigkeit der Evaluatoreninnen und Evaluatoren beeinflusst entscheidend die Durchführung

und Wirksamkeit der Evaluation. Um von den verschiedenen betroffenen Gruppen als glaubwürdig beurteilt zu werden, sind vor allem folgende Eigenschaften bedeutsam: fachliche und methodische Kompetenz, Integrität, Unabhängigkeit sowie soziale und kommunikative Kompetenzen. Diese Eigenschaften sollen bei der Auswahl der Evaluatoren und Evaluatorinnen berücksichtigt werden. Oft ist es hilfreich, die erforderlichen Kompetenzen durch die Zusammenstellung eines Evaluationsteams sicher zu stellen.

N4 Auswahl und Umfang der Informationen

Auswahl und Umfang der erfassten Informationen sollen die Behandlung der zu untersuchenden Fragestellungen zum Evaluationsgegenstand ermöglichen und gleichzeitig den Informationsbedarf des Auftraggebers und anderer Adressaten und Adressatinnen berücksichtigen.

Bei der Planung eines Evaluationsprojektes soll heraus gearbeitet werden, welche Informationen zur Beantwortung der Evaluationsfragestellungen unabdingbar sind und welche Informationen zwar interessant und wünschenswert, aber für die Beantwortung der Kernfragestellungen unerheblich sind. Es ist dabei darauf zu achten, dass die vorhandenen Ressourcen zur Datenerhebung entsprechend den Prioritäten zur Beantwortung der Kernfragestellungen und dem Informationsbedarf der wichtigsten vorgesehenen Nutzer und Nutzerinnen der Evaluation eingesetzt werden (Bundesamt für Gesundheit 1997).

N5 Transparenz von Werten

Die Perspektiven und Annahmen der Beteiligten und Betroffenen, auf denen die Evaluation und die Interpretation der Ergebnisse beruhen, sollen so beschrieben werden, dass die Grundlagen der Bewertungen klar ersichtlich sind.

In mehreren Phasen der Evaluation, bereits bei der Festlegung der Evaluationszwecke und -fragestellungen, spielen die Werte der Beteiligten und Betroffenen eine große Rolle. Deren Identifikation und Thematisierung durch die Evaluation ist dann besonders wichtig, wenn die Beteiligten und Betroffenen zur aktiven Teilnahme an der Evaluation und schließlich zur Nutzung der Evaluationsergebnisse ermuntert werden sollen (Wottawa/Thierau 1998). Die in der Schlussphase vorzunehmende Interpretation der gesammelten Informationen und Ergebnisse stellt einen der wichtigsten und kritischsten Punkte in einem Evaluationsprozess dar. In diesem Interpretationsvorgang spielen ge-



sellschaftlich vermittelte Werthaltungen (Normen) unvermeidlich eine große Rolle. Um den Interpretationsvorgang überzeugend, nachvollziehbar und beurteilbar zu machen, ist es unerlässlich, die zu Grunde liegenden Werthaltungen möglichst transparent zu machen.

N 6 Vollständigkeit und Klarheit der Berichterstattung

Evaluationsberichte sollen alle wesentlichen Informationen zur Verfügung stellen, leicht zu verstehen und nachvollziehbar sein.

Eine Bedingung für die erfolgreiche Kommunikation von Evaluationsergebnissen ist die Vollständigkeit und Klarheit des schriftlichen Berichts (oder anderer Formen der Berichterstattung). Die Sprache soll für die Adressaten und Adressatinnen verständlich sein und wichtige Begriffe sollen definiert und konsistent verwendet werden. Die Zusammenstellung der wichtigsten Ergebnisse – z. B. in einer voran gestellten Zusammenfassung, in Tabellenform oder in Graphiken – kann das Verständnis der berichteten Inhalte erheblich fördern.

■ Die Gestaltung und Art der Berichterstattung soll auf die Wahrnehmungsgewohnheiten der Adressaten und Adressatinnen einer Evaluation abgestimmt werden. Ein umfangreicher Schlussbericht stellt dabei nicht für jede adressierte Gruppe die optimale Form zur Weitergabe der Informationen dar. Zuweilen können – je nach Situation und Zielpublikum – Referate, Workshops oder ähnliche Kommunikationsformen zweckdienlicher sein.

N 7 Rechtzeitigkeit der Evaluation

Evaluationsvorhaben sollen so rechtzeitig begonnen und abgeschlossen werden, dass ihre Ergebnisse in anstehende Entscheidungsprozesse bzw. Verbesserungsprozesse einfließen können.

Der Termin, zu dem eine Evaluation vorbereitet, ggf. ausgeschrieben, beauftragt und schließlich durch erste Aktivitäten im Evaluationsfeld begonnen wird, soll ermöglichen, dass die erforderlichen Arbeiten zum Zeitpunkt der vorgesehenen Nutzung der Evaluationsergebnisse abgeschlossen sind. Der Zeitplan soll so beschaffen sein, dass notwendige Abstimmungsprozesse, Vorbereitungen für Datenerhebungen, Zeiten für die Fehlerprüfung und die Auswertung und Interpretation der Informationen realistisch eingeplant sind. Für Unvorhersehbares sollen Pufferzeiten vorgesehen werden.

■ Wichtige Zwischenergebnisse und Schlussberichte sind den vorgesehenen Nutzerinnen und Nutzern so zur Kenntnis zu bringen, dass diese sie rechtzeitig verwenden können. Ein Evaluationspro-



jekt erzielt erfahrungsgemäß größere Wirkung, wenn sein zeitlicher Ablauf auf anstehende Entscheidungsprozesse und/oder Verbesserungsprozesse abgestimmt wird. Andernfalls verliert es einen Großteil seiner Wirkungen. Dabei ist darauf zu achten, dass in vielen Fällen (z. B. Auftrag durch öffentliche Verwaltung) eine beträchtliche Vorlaufzeit eingeplant werden muss, da der Evaluationsbericht noch intern verarbeitet wird (z. B. Mitzeichnungsverfahren, Stellungnahmen), bevor Entscheidungen getroffen oder Maßnahmen eingeleitet werden können. In vielen Evaluationsprojekten ist es sinnvoll, vorläufige Resultate und Zwischenberichte bereits im Verlauf der Untersuchung rückzumelden, vor allem, wenn die Informationen für die Adressaten und Adressatinnen aktuell handlungsrelevant sind. Solche Rückmeldungsschleifen sollen bereits in der Evaluationsplanung berücksichtigt und entsprechende Ressourcen sollen dafür vorgesehen werden.

N8 Nutzung und Nutzen der Evaluation

Planung, Durchführung und Berichterstattung einer Evaluation sollen die Beteiligten und Betroffenen dazu ermuntern, die Evaluation aufmerksam zur Kenntnis zu nehmen und ihre Ergebnisse zu nutzen.

Selbst mit hoher Professionalität erzeugte Ergebnisse werden häufig nicht genutzt oder ganz anders, als dies in den Vereinbarungen zur Evaluation geplant war. Wichtige Voraussetzungen für eine intendierte Verwendung liegen in einer hohen Durchführbarkeit, Fairness und Genauigkeit von Evaluationen sowie den Aspekten, die in den vorausgegangenen sieben Nützlichkeitsstandards ausgeführt sind. Besonders die Umsetzung von Schlussfolgerungen oder Empfehlungen hängt stark davon ab, ob die Adressaten und Adressatinnen die Erwartung haben, dass das Evaluationsprojekt ihnen und ihren Vorhaben von Nutzen ist. Dabei ist es nicht selten so, dass unterschiedliche Adressaten auch unterschiedliche Nutzenerwartungen haben, die sich widersprechen oder sich gegenseitig ausschließen können. Eine wichtige Voraussetzung, um eine vereinbarte bzw. intendierte Nutzung und damit auch den Nutzen von Evaluationen zu fördern, ist die angemessene Einbeziehung der verschiedenen Adressaten und Adressatinnen in die Planung und die Durchführung der Evaluation sowie die Aufbereitung von Ergebnissen. Ebenfalls positiv wirkt es sich aus, wenn im Verlauf der Untersuchung kontinuierlich und transparent Rückmeldungen gegeben werden (Stockbauer 2000). Dies fördert, dass abgesehen vom Ergebnisnutzen bereits im Verlauf einer Evaluation ein Prozessnutzen eintritt (Patton 1998).



DURCHFÜHR- BARKEIT

Die Durchführbarkeitsstandards sollen sicherstellen, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst geplant und ausgeführt wird.

D 1 Angemessene Verfahren

Evaluationsverfahren, einschließlich der Verfahren zur Beschaffung notwendiger Informationen, sollen so gewählt werden, dass Belastungen des Evaluationsgegenstandes bzw. der Beteiligten und Betroffenen in einem angemessenen Verhältnis zum erwarteten Nutzen der Evaluation stehen.

Verfahren der Evaluation sollen einerseits den Anforderungen nach wissenschaftlicher Güte entsprechen und andererseits weder den Evaluationsgegenstand noch die Beteiligten und Betroffenen unnötig belasten oder stören. Die aus wissenschaftlicher Sicht aussagekräftigsten Methoden können oft nicht verwendet werden, da sie zu aufwändig (zeitraubend oder kostspielig) oder im entsprechenden Kontext ethisch nicht akzeptabel sind (Uhl 2000). Vor- und Nachteile sowie Aussagekraft der gewählten Verfahren sollen durch das Evaluationsteam transparent gemacht und begründet werden.

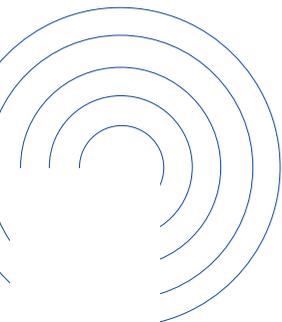
D 2 Diplomatisches Vorgehen

Evaluationen sollen so geplant und durchgeführt werden, dass möglichst hohe Akzeptanz der verschiedenen Beteiligten und Betroffenen in Bezug auf Vorgehen und Ergebnisse der Evaluation erreicht werden kann.

Eine ausgewogene Berücksichtigung der Interessen aller Beteiligten und Betroffenen fördert Akzeptanz, Zustimmung und Kooperation der verschiedenen Interessengruppen im Rahmen der Durchführung der Evaluation. Mögliche Versuche irgendeiner dieser Gruppen, die Evaluationsaktivitäten einzuschränken oder die Ergebnisse zu verzerren oder zu missbrauchen, sollen so vermieden werden. Um dieses Ziel zu erreichen, ist diplomatisches Vorgehen des Evaluationsteams gefragt.

■ Politische Tragfähigkeit und Kompromissfähigkeit sind bei Evaluationen im politischen oder organisatorischen Umfeld eine Voraussetzung für die spätere Nutzung von Evaluationsergebnissen (Wottawa/Thierau 1998, Hager/Patry/Brezing 2000, Beywl 2001, Faßmann 2001). Allerdings werden gerade im politischen Umfeld Evaluationen oft in Auftrag gegeben, um Verantwortung für schwierige Entscheidungen abzuwälzen oder bereits getroffene Entscheidungen nachträglich zu legitimieren. In einem solchen Kontext soll das Evaluationsteam besonders den Dialog zwischen den Entscheidungsträgern sowie anderen Beteiligten und Betroffenen fördern.

■ Gelingt es, eine kooperative Haltung bei den einzelnen Interessengruppen zu erzeugen, wird sich



dies positiv auf die Bereitschaft auswirken, sich am Evaluationsprozess zu beteiligen, Informationen bereit zu stellen, die Ergebnisse zu akzeptieren und ggf. zu nutzen.

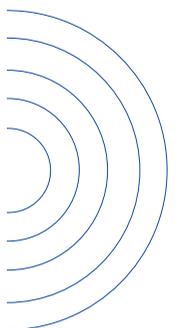
D 3 Effizienz von Evaluation

Der Aufwand für Evaluation soll in einem angemessenen Verhältnis zum Nutzen der Evaluation stehen.

Es ist sowohl zu Beginn als auch bei Abschluss einer Evaluation oft schwierig, genaue Aussagen zu Kosten und Nutzen eines Evaluationsvorhabens zu machen.¹⁴ Solche Schwierigkeiten dürfen jedoch nicht dazu führen, dass Überlegungen zu Kosten und Nutzen vollständig unterbleiben. Gerade bei der Entscheidung über die Durchführung einer Evaluation sollen Kosten und Nutzen abgeschätzt werden. In der Planung einer Evaluation soll nachvollziehbar dargestellt werden, welcher Aufwand voraussichtlich entsteht und welcher Nutzen erwartet wird.

■ Die vollständigen Kosten der Evaluation umfassen den Geldwert aller benötigten Ressourcen, z. B. Honorare für die Evaluatoren und Evaluatorinnen, den evaluationspezifischen Zeiteinsatz aller Beteiligten und Betroffenen, Reisekosten und Materialkosten. Es gibt außerdem auch sonstige Kosten, auch solche, die von Dritten getragen werden. Viele Kosten sind nicht quantifizierbar.

■ Die Quantifizierung des Nutzens ist noch schwieriger. In aller Regel kann er nur geschätzt werden. Ein Nutzen kann aus direkten und indirekten, gewollten oder ungewollten Wirkungen der Evaluation entstehen.



¹⁴Siehe z. B. die grundlegende Kritik an Kosten-Nutzen-Analysen aus ökonomischer Sicht in: Rürup/Hansmeyer (1984, S. 107ff).

FAIRNESS Die Fairnessstandards sollen sicherstellen, dass in einer Evaluation respektvoll und fair mit den betroffenen Personen und Gruppen umgegangen wird.

F1 Formale Vereinbarungen

Die Pflichten der Vertragsparteien einer Evaluation (was, wie, von wem, wann getan werden soll) sollen schriftlich festgehalten werden, damit die Parteien verpflichtet sind, alle Bedingungen dieser Vereinbarung zu erfüllen oder aber diese neu auszuhandeln.

Die formalen Vereinbarungen für eine Evaluation sollen auf der Grundlage von gegenseitigem Respekt und Vertrauen getroffen werden und zumindest in folgenden Bereichen festgelegt sein: Finanzen, Zeit, Methodik sowie mitwirkende Personen. Insbesondere die Rechte und Pflichten der Verantwortlichen und Mitwirkenden sind möglichst präzise festzuhalten. Eine formelle schriftliche Vereinbarung vermindert die Wahrscheinlichkeit, dass Missverständnisse unter den Vertragsparteien auftreten bzw. erleichtert ggf. deren Bereinigung.

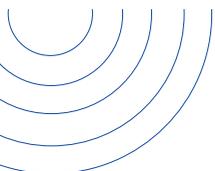
■ Dabei ist zu bedenken, dass rigide Festlegungen bezüglich der Detailfragestellungen, der Methodik und des Vorgehens leicht zu Hemmschuhen werden können, die den Erkenntnisgewinn nachhaltig behindern. Werden im Verlauf der Zeit Änderungen nötig, besteht zwar die Möglichkeit, die Abweichungen zu begründen und Bedingungen neu auszuhandeln, aber das ist leichter möglich, wenn das Evaluationsteam alle an der Vereinbarung Beteiligten von Anfang an über die potentielle Notwendigkeit von Anpassungen aufgeklärt hat.

F2 Schutz individueller Rechte

Evaluationen sollen so geplant und durchgeführt werden, dass Sicherheit, Würde und Rechte der in eine Evaluation einbezogenen Personen geschützt sind.

Evaluatorinnen und Evaluatoren sollen darauf achten, die Würde und Selbstachtung der Personen, mit denen sie im Verlauf der Evaluation in Kontakt treten, nicht in irgendeiner Form zu verletzen.¹⁵ Wo eine Evaluation auch Schwächen offenlegt, kann es zu massiven Beeinträchtigungen des Interessenschutzes kommen. In diesen Fällen sind die Interessen sorgfältig abzuwägen. Besonders gilt dies, wenn Verstöße gegen Recht und Gesetz offensichtlich werden.

28 ¹⁵Vgl. für diesen Standard z. B. AERA/APA/NCME (1999), Deutsche Gesellschaft für Erziehungswissenschaft (1997), IHK/ESOMAR (o. J.), Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e. V. u. a. (2001) sowie als Internet-Ressource (<http://www.datenschutz-berlin.de>).



F3 Vollständige und faire Überprüfung

Evaluationen sollen die Stärken und die Schwächen des Evaluationsgegenstandes möglichst vollständig und fair überprüfen und darstellen, so dass die Stärken weiter ausgebaut und die Schwachpunkte behandelt werden können und die Rechte der in eine Evaluation einbezogenen Personen geschützt sind.

Es ist selbstverständlich, dass die Schwächen eines Evaluationsgegenstandes umfassend überprüft und dargestellt werden sollen. Nur so können sie behoben werden. Doch auch die Darstellung der Stärken ist wichtig. Beides ist für die umfassende Bewertung des Evaluationsgegenstandes und für seine Verbesserung notwendig.

■ Stärken und Schwächen eines Evaluationsgegenstandes hängen oft eng zusammen. Daher ist zu beachten, dass unter Umständen durch die Korrektur von Schwächen bestehende Stärken gefährdet werden können. Andererseits ist es eventuell möglich, bestehende Schwächen durch vorhandene Stärken auszugleichen. Häufig ist nicht nur der beabsichtigte Abbau von Schwächen, sondern auch der Ausbau von Stärken Anlass einer Evaluation. Darüber hinaus steigt in der Regel die Bereitschaft zur Mitarbeit und die Akzeptanz der Evaluation durch die Beteiligten und Betroffenen, wenn auch die Stärken erfasst werden.

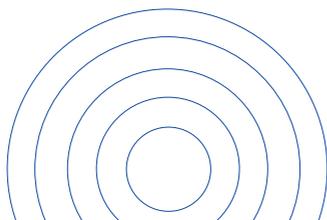
■ Die Überprüfung und Darstellung der Stärken und Schwächen soll so vollständig wie möglich sein, damit alle wichtigen Aspekte zum Vorschein kommen. Sie soll gegenüber den Beteiligten und Betroffenen fair sein. Auftraggeber oder andere Beteiligte sollen nicht versuchen, einseitig Einfluss auf die Evaluation und den Bericht zu nehmen. Es ist Aufgabe der Evaluatoren und Evaluatorinnen, eine vollständige, ausgewogene und faire Evaluation sicherzustellen.

F4 Unparteiische Durchführung und Berichterstattung

Die Evaluation soll unterschiedliche Sichtweisen von Beteiligten und Betroffenen auf Gegenstand und Ergebnisse der Evaluation in Rechnung stellen. Berichte sollen ebenso wie der gesamte Evaluationsprozess die unparteiische Position des Evaluationsteams erkennen lassen. Bewertungen sollen fair und möglichst frei von persönlichen Gefühlen getroffen werden.

Das Umfeld einer Evaluation ist durch vielfältige Sichtweisen geprägt. Die Beteiligten und Betroffenen haben oft divergierende Ansichten über den Evaluationsgegenstand. Es besteht die Gefahr, dass Evaluationen von einer bestimmten Gruppe vereinnahmt oder instrumentalisiert werden.

■ Die Evaluation soll darauf verzichten, lediglich eine spezifische Sichtweise zu übernehmen. Viel-



mehr soll sie eine faire Behandlung aller relevanten Interessen anstreben. Deshalb ist dafür zu sorgen, dass das Evaluationsteam eine möglichst unabhängige Position einnimmt. So ist ein zu enges Verhältnis zu Auftraggebern aber auch beispielsweise zu den für den Evaluationsgegenstand verantwortlichen Personen zu vermeiden. Das Verhältnis der Evaluatorinnen und Evaluatoren gegenüber den Auftraggebern der Evaluation und anderen relevanten Gruppen ist zu Beginn des Evaluationsprozesses zu klären. Diese Rollenklärung umfasst auch Festlegungen über die Veröffentlichung von Evaluationsberichten (Müller-Kohlenberg/Münstermann 2000).

F5 Offenlegung der Ergebnisse

Die Evaluationsergebnisse sollen allen Beteiligten und Betroffenen soweit wie möglich zugänglich gemacht werden.

Grundsätzlich sollen alle Personen, die an einer Evaluation beteiligt oder von dieser betroffen sind, Zugang zu den Ergebnissen und zu dem Evaluationsbericht haben. Alle an einer Evaluation mitwirkenden Akteure (z. B. Evaluationsteam, Auftraggeber, Adressaten und Adressatinnen, Mitarbeiter und Mitarbeiterinnen der Organisation, in der evaluiert wird) sind dafür verantwortlich, dass dies gewährleistet wird. Da der Kreis der Beteiligten und Betroffenen oft sehr groß ist, kann dies in vielen Fällen bedeuten, dass der Bericht öffentlich zugänglich gemacht werden muss.

■ Die Offenlegung der Evaluationsergebnisse soll den Nutzen der Evaluation steigern. Es kann jedoch Fälle geben, in denen eine vollständige Offenlegung den Nutzen einer Evaluation mindert. Dies kann z. B. der Fall sein, wenn konkurrierende Firmen die Evaluationsergebnisse zum Schaden einer evaluierten Firma nutzen können, wenn die Bereitschaft der Beteiligten und Betroffenen zur Mitarbeit an der Evaluation und der Umsetzung ihrer Ergebnisse durch ihr Wissen um die spätere Veröffentlichung (von Schwächen) leidet, oder wenn die Gefahr besteht, dass Informationsgeber durch eine Veröffentlichung kompromittiert würden. In solchen Fällen ist es die Aufgabe der Akteure, gemeinsam eine Lösung zu finden und ggf. eine Beschränkung der Offenlegung zu vereinbaren. Beschränkungen bei der Offenlegung sollen begründet werden.

■ Die Offenlegung der Evaluationsergebnisse, ihre Art und ihr Umfang, soll zu Beginn einer Evaluation vereinbart und vertraglich festgelegt werden. Falls dies nicht möglich ist, soll zu Beginn vereinbart werden, durch wen, wie, nach welchen Kriterien und wann (z. B. erst zum Ende eines Evaluationsprojektes) über die Offenlegung der Ergebnisse entschieden wird.

GENAUIGKEIT Die Genauigkeitsstandards sollen sicherstellen, dass eine Evaluation gültige Informationen und Ergebnisse zu dem jeweiligen Evaluationsgegenstand und den Evaluationsfragestellungen hervor bringt und vermittelt.

G 1 Beschreibung des Evaluationsgegenstandes

Der Evaluationsgegenstand soll klar und genau beschrieben und dokumentiert werden, so dass er eindeutig identifiziert und so genau wie möglich zugänglich gemacht werden kann.

Der Evaluationsgegenstand, sei es nun eine Maßnahme, ein Programm oder eine Organisation, soll genau beschrieben werden. Es ist zu berücksichtigen, dass sich der Evaluationsgegenstand in unterschiedlichen zeitlichen und räumlichen Kontexten verschieden ausgestalten kann. Die Beschreibung des Evaluationsgegenstandes soll klar machen, was untersucht wird. Dies ermöglicht den Adressaten und Adressatinnen auch den Vergleich mit anderen Evaluationsgegenständen. Eine genaue Beschreibung des Evaluationsgegenstandes hilft, Zusammenhänge zwischen diesem und seinen Wirkungen festzustellen und unter Umständen bisher unbeachtete Nebenwirkungen aufzufinden. Dabei ist darauf zu achten, inwiefern sich zwischen der ursprünglich vorgesehenen Form des Evaluationsgegenstandes und der tatsächlichen Implementierung Diskrepanzen zeigen.

G 2 Kontextanalyse

Der Kontext des Evaluationsgegenstandes soll ausreichend detailliert untersucht und analysiert werden.

Den Kontext des Evaluationsgegenstandes bilden die Rahmenbedingungen, welche diesen Evaluationsgegenstand umgeben. Dies sind z. B. das soziale und politische Klima, die Charakteristika und Interessen der maßgeblich Beteiligten und Betroffenen oder die wirtschaftlichen Rahmenbedingungen. Analysen zum Umfeld des Evaluationsgegenstandes sind bedeutsam zur Gewinnung von Erkenntnissen über mögliche Bedingungsfaktoren, die sich auf den Evaluationsgegenstand und auf die Evaluationsergebnisse auswirken. Eine fundierte Kontextanalyse erlaubt es auch, die Resultate einer Evaluation inbezug auf ihre Übertragbarkeit auf andere Kontexte einschätzen zu können. Es soll vermieden werden, den Kontext zu eng zu definieren. Es ist aber auch davor zu warnen, die Kontextanalyse zu ausführlich anzulegen, da dies auf Kosten der Bearbeitung der anderen Evaluations Schritte gehen kann.

G 3 Beschreibung von Zwecken und Vorgehen

Gegenstand, Zwecke, Fragestellungen und Vorgehen der Evaluation, einschließlich der angewandten Methoden, sollen genau dokumentiert und beschrieben werden, so dass sie identifiziert und eingeschätzt werden können.

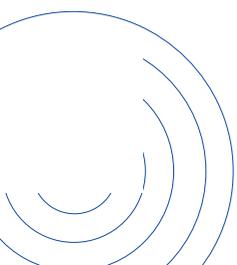
Um den Evaluationsprozess transparent zu gestalten, sind die Zwecke, die mit einer Evaluation angestrebt werden, schon im Verlauf des Evaluationsprozesses genau zu dokumentieren und im Rahmen der Berichterstattung an die Adressatinnen und Adressaten der Evaluation in nachvollziehbarer Weise zu kommunizieren. Ebenso soll mit den Fragestellungen, die bearbeitet werden, sowie den gewählten Vorgehensweisen, Methoden und den Begründungen für die dem Evaluationsplan zu Grunde liegenden Entscheidungen verfahren werden. Der Zeitplan und Abweichungen sind schriftlich festzuhalten.

■ Bei der Beschreibung der Zwecke und der Fragestellungen der Evaluation ist speziell darauf zu achten, dass auch abweichende Sichtweisen berücksichtigt werden. Die Dokumentation des Vorgehens umfasst die detaillierte Beschreibung der organisatorischen Vorkehrungen, der Erhebungen einschließlich evtl. Stichprobenverfahren, der Aufbereitung und Auswertung der Daten, des Interpretationsprozesses und schließlich der Berichterstattung. Bei der Vorgehensweise ist darauf zu achten, dass sich diese im Verlauf des Evaluationsprozesses verändern kann und dass sich damit das geplante vom tatsächlichen Vorgehen unterscheiden kann. Es soll auch nach außen deutlich dargelegt werden, welche Abweichungen entstanden sind und welche Gründe dafür verantwortlich sind.

G 4 Angabe von Informationsquellen

Die im Rahmen einer Evaluation genutzten Informationsquellen sollen hinreichend genau dokumentiert werden, damit die Verlässlichkeit und Angemessenheit der Informationen eingeschätzt werden kann.

Die Beschreibung der genutzten Informationsquellen erlaubt es den Adressaten und Adressatinnen, sich selbst ein Urteil über die Qualität der aus diesen Quellen hervor gegangenen Informationen zu bilden. Als Informationsquellen dienen in Evaluationen u. a. Individuen oder Gruppen, Dokumente, audiovisuelle Materialien, Statistiken usw.. Das Heranziehen unterschiedlicher Informationsquellen erlaubt einen Vergleich der gewonnenen Informationen. Die Glaubwürdigkeit einer Evaluation kann durch eine fehlende oder unzureichende Beschreibung der Informationsquellen in Frage gestellt werden. Neben der Beschreibung der Informationsquellen sollen die aus diesen Quellen erarbeiteten



Informationen auch eingeschätzt und bewertet werden können (Deutsche Forschungsgemeinschaft 1999).

G5 Valide und reliable Informationen

Die Verfahren zur Gewinnung von Daten sollen so gewählt oder entwickelt und dann eingesetzt werden, dass die Zuverlässigkeit der gewonnenen Daten und ihre Gültigkeit bezogen auf die Beantwortung der Evaluationsfragestellungen nach fachlichen Maßstäbe sichergestellt sind. Die fachlichen Maßstäbe sollen sich an den Gütekriterien quantitativer und qualitativer Sozialforschung orientieren.

Im Rahmen der empirischen Erfassung des Evaluationsgegenstandes müssen die eingesetzten Datenerhebungsinstrumente und die mit ihnen erhobenen Daten bestimmten Gütekriterien genügen, um den Einfluss von Fehlerquellen auf den Erfassungsvorgang zu minimieren.

■ Die zentralen Gütekriterien im Bereich quantitativer Methoden besagen, dass Messungen unabhängig von der jeweiligen Person sein sollen, die das Messinstrument anwendet (Objektivität). Die Instrumente sollen konsistente, reproduzierbare und damit zuverlässige Informationen bereitstellen und möglichst unanfällig gegenüber Störeinflüssen oder Zufallsfehlern (unsystematischen Fehlerquellen) sein (Reliabilität oder Messgenauigkeit). Es soll gewährleistet sein, dass Instrumente die Merkmale oder Verhaltensweisen, die sie zu messen vorgeben, auch tatsächlich effektiv erfassen (Validität oder Gültigkeit). Dabei bilden Objektivität und Reliabilität notwendige aber nicht hinreichende Voraussetzungen für valide Messungen (AERA/APA/NCME 1999, Dieckmann 1995, Bortz/Döring 1995).

■ In Evaluationen kann sehr häufig nicht auf standardisierte, qualitätsgesicherte quantitative Instrumente zurückgegriffen werden, deren Reliabilitäts- und Validitäts-Kennwerte bekannt sind. Oft ist ein qualitativer Zugang notwendig und bezogen auf Gegenstand und Fragestellungen auch angemessener. Um die Zuverlässigkeit und Gültigkeit von nicht-standardisierten Instrumenten, qualitativen Daten und subjektiven Interpretationen zu sichern, wurden spezifische Gütekriterien qualitativer Forschung entwickelt (z. B. „Verfahrensdokumentation“, „Argumentative Interpretationsabsicherung“, „Nähe zum Gegenstand“ und „Triangulation“ nach Mayring 1999).

■ Validität betrifft nicht nur die Qualität von Datenerhebungsinstrumenten und Daten, sondern auch die Güte und Glaubwürdigkeit der Schlüsse, die aus den Ergebnissen der Datenerhebung gezogen werden. So verstanden lässt sich Validität nur im konkreten Evaluationskontext und in Hinblick auf die spezifische Zielsetzung und Fragestellung der empirischen Erfassung einschätzen. Validierung ist das Zusammenbringen von Belegen, welche die datenbasierte Interpretation unterstützen. Es sollen deshalb mehrere – quantitative und/oder qualitative – Verfahren zur Datengewinnung

verwendet werden. Die Schlussfolgerungen aus allen verwendeten Verfahren sollen – einzeln und in Kombination – validiert werden. Die hier vorgestellten wissenschaftlichen Gütekriterien besitzen für Entscheidungen über Auswahl, Entwicklung und Einsatz von Methoden Orientierungsfunktion und müssen bezogen auf die Anforderungen der Evaluationspraxis spezifiziert und ergänzt werden (Hager/Patry/Brezing 2000).

G 6 Systematische Fehlerprüfung

Die in einer Evaluation gesammelten, aufbereiteten, analysierten und präsentierten Informationen sollen systematisch auf Fehler geprüft werden.

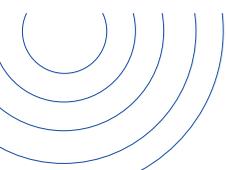
Bei der Erhebung, Aufbereitung, Auswertung und Interpretation von Informationen sowie bei der Ergebnispräsentation bestehen mannigfaltige Möglichkeiten, dass Fehler auftreten. Solche Fehler können methodische Fehler im engeren Sinn sein (Beispiele für den Bereich quantitativer Forschung: Messfehler, Fehler im Untersuchungsdesign durch mangelnde Kontrolle von Störvariablen, Verfälschungen der Untersuchungssituation durch Einflussnahme des Evaluationsteams, Fehler bei der Stichprobenziehung, Missachtung der Anwendungsvoraussetzungen statistischer Verfahren usw.) aber auch den Bereich mangelnder Sorgfalt betreffen (Verwechslung von Erhebungs-Kennnummern, Vertauschen von Mitschriften oder Tonbändern, Fehleingaben von Daten in die Datenverarbeitung, Vertauschen von Beschriftungen in Ergebnispräsentationen usw.).

■ Deshalb ist es zwingend erforderlich, den Evaluationsprozess so zu gestalten, dass potentielle Fehlerquellen frühzeitig erkannt und Fehler soweit wie möglich vermieden oder korrigiert werden. Das kann durch systematische Schulungsprogramme der Mitwirkenden sowie durch systematische Kontrollen und Genauigkeitsprüfungen (Plausibilitätstests, Parallelerfassung, kommunikative Validierung usw.) geschehen. Mögliche Fehler und deren Konsequenzen sind im Rahmen der Berichterstattung zur Evaluation in transparenter Weise zu diskutieren. Fehler führen zu fehlgeleiteten Interpretationen und können die gesamte Evaluation diskreditieren.

G 7 Analyse qualitativer und quantitativer Informationen

Qualitative und quantitative Informationen einer Evaluation sollen nach fachlichen Maßstäben angemessen und systematisch analysiert werden, damit die Fragestellungen der Evaluation effektiv beantwortet werden können.

Die in Evaluationen erhobenen quantitativen und qualitativen Daten werden im Prozess der Daten-



analyse geordnet, zusammengefasst und ausgewertet, um die Grundlage für Interpretationen und Schlussfolgerungen im Rahmen der Beantwortung der Evaluationsfragestellungen zu schaffen. Die Auswahl der adäquaten Erhebungs- und Analyseverfahren soll sich nach den Fragestellungen der Evaluation, nach dem Erkenntnisstand bezogen auf den Evaluationsgegenstand und nach Kontextvariablen des Evaluationsfeldes richten. Bei dieser Entscheidung sollen Vorlieben der auswertenden Personen keine Rolle spielen. Die Auswertung der erhobenen Informationen soll in systematischer Weise erfolgen. Dabei sollen für qualitative Verfahren (z. B. Mayring 1999, Flick 1995, Lamneck 1995, Miles/Huberman 1994) ebenso wie für quantitative Verfahren (z. B. Bortz/Döring 1995, Bortz 1999, Dieckmann 1995, Kromrey 1998) eigene Gütekriterien und Regeln Beachtung finden. Die heran gezogenen Regeln und deren methodische Grundlagen sollen – z. B. durch Literaturhinweise – offen gelegt werden. Die Auswahl sowie der Einsatz der Verfahren soll transparent und nachvollziehbar erfolgen, so dass Auswahlentscheidungen ebenso wie Ergebnisse kritisierbar bleiben. Verwendete Kennwerte und Formeln sollen allgemein verständlich erklärt werden. Auf die Aussagekraft der Methoden, wie auch auf ihre Begrenzungen, ist explizit hinzuweisen.

G 8 Begründete Schlussfolgerungen

Die in einer Evaluation gezogenen Folgerungen sollen ausdrücklich begründet werden, damit die Adressatinnen und Adressaten diese einschätzen können.

Die Schlussfolgerungen einer Evaluation sollen explizit begründet und zusammen mit den zugrunde liegenden Annahmen und den eingesetzten Evaluationsverfahren transparent dargestellt werden. Dabei ist auch deutlich auf den Geltungsbereich der Schlussfolgerungen hinzuweisen. Wo erforderlich, soll die Berichterstattung eine Diskussion alternativer Interpretationen beinhalten, wobei begründet werden soll, warum diese verworfen wurden. Die Einhaltung dieses Standards erlaubt es den Nutzerinnen und Nutzern der Evaluationsergebnisse, die Aussagekraft der Folgerungen einzuschätzen. Zudem wird die Überzeugungskraft der Folgerungen gestärkt.

G 9 Meta-Evaluation

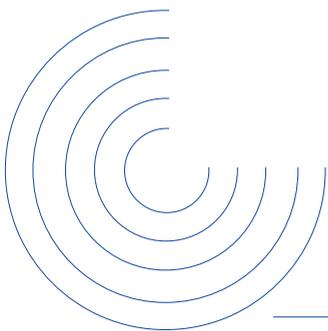
Um Meta-Evaluationen zu ermöglichen, sollen Evaluationen in geeigneter Form dokumentiert und archiviert werden.

Misslungene Evaluationen können Anlass zu Fehlentscheidungen geben. Evaluationen können aber auch – evtl. zu Unrecht – heftiger Kritik ausgesetzt sein. Um derartige Situationen zu vermeiden,

soll auch die Evaluation selbst evaluiert werden. Solche Meta-Evaluationen dienen dem Zweck, die Qualität von Evaluationsprozessen und -ergebnissen zu beurteilen und zu verbessern. Im Rahmen einer Meta-Evaluation können die hier vorliegenden Standards eingesetzt werden (Widmer 1996). Eine systematische Meta-Evaluation kann im Rahmen des Evaluationsteams (intern) oder durch Außenstehende (extern) erfolgen. Während eine umfassende und vertiefte Meta-Evaluation nur in bestimmten Fällen zweckmäßig ist, soll eine knappe Selbstevaluation der Evaluation durch die für die Evaluation verantwortlichen Personen in aller Regel vorgenommen werden. Die regelmäßige Durchführung von Meta-Evaluationen wird die Glaubwürdigkeit einzelner Evaluationen, aber auch die des Berufsstandes der Evaluatoren und Evaluatorinnen erhöhen.

■ Um Meta-Evaluationen, Evaluationssynthesen oder Meta-Analysen zu ermöglichen, sollen der Evaluationsgegenstand, die wesentlichen Zwecke, Schritte, Methoden, Daten und Ergebnisse eines Evaluationsprojektes vom Evaluationsteam dokumentiert und archiviert werden. Dies dient dem wissenschaftlichen Fortschritt und der Wissensakkumulation im jeweiligen Gegenstandsbereich sowie im Bereich der Modelle und Methoden der Evaluationsforschung.

■ Darüber hinaus ist die wissenschaftliche Publikation über Evaluationsprojekte und ihre Verfahren, Probleme und Ergebnisse grundsätzlich wünschenswert. Sie fördert den Fortschritt der Evaluation und die Verbreitung von Wissen im jeweiligen Gegenstandsbereich, ist ein Instrument der Qualitätsentwicklung und fördert darüber hinaus die Verbreitung und Akzeptanz von Evaluationen.

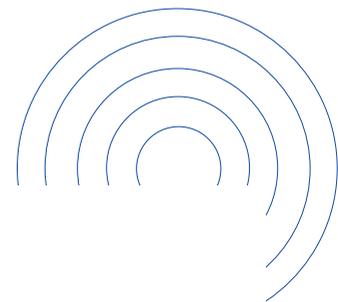


FUNKTIONALES INHALTSVERZEICHNIS

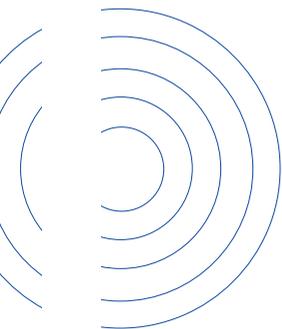
Die Standards sind in diesem Kapitel nach zehn in einer Evaluation zu bearbeitenden Hauptaufgaben gruppiert. Die Aufgaben 1 bis 6 sind chronologisch – im Ablauf einer Evaluation – zu bearbeiten und bezeichnen somit auch Phasen von Evaluationen. Die Aufgaben 7 bis 10 sind Querschnittsaufgaben des Evaluationsmanagements.

1. Entscheidung über die Durchführung einer Evaluation	38
2. Definition des Evaluationsproblems	38
3. Planung der Evaluation	38
4. Informationsgewinnung	39
5. Informationsauswertung	39
6. Berichterstattung zur Evaluation	39
7. Budgetierung der Evaluation	40
8. Evaluationsvertrag	40
9. Steuerung der Evaluation	41
10. Personelle Ausstattung der Evaluation	41

Diese Art der Darstellung soll den Lesern und Leserinnen verdeutlichen, dass die verschiedenen Standards im ganzen Evaluationsprozess angewendet werden können, angefangen bei der Planung über die Durchführung bis hin zur Berichterstattung. Die funktionale Übersicht veranschaulicht zudem, dass jeder Evaluationsaufgabe Standards zugeordnet werden können, deren Anwendung für die Aufgabenerfüllung in der Regel von besonderer Bedeutung ist. Es wäre jedoch ein Fehler, lediglich die hier bezeichneten Standards für eine bestimmte Aufgabe heranzuziehen, ohne dass geprüft worden wäre, ob auch andere relevant sein könnten. Grundsätzlich sollen alle Standards in Betracht gezogen werden, wenn es darum geht, ihre jeweilige Bedeutung für eine spezifische Evaluationsaufgabe einzuschätzen.



1. Entscheidung über die Durchführung einer Evaluation – wichtigste Standards:		
N1	Identifizierung der Beteiligten und Betroffenen	21
N2	Klärung der Evaluationszwecke	22
N7	Rechtzeitigkeit der Evaluation	24
N8	Nutzung und Nutzen der Evaluation	25
D2	Diplomatisches Vorgehen	26
D3	Effizienz von Evaluation	27
F1	Formale Vereinbarungen	28
G1	Beschreibung des Evaluationsgegenstandes	31
G2	Kontextanalyse	31
2. Definition des Evaluationsproblems – wichtigste Standards:		
N1	Identifizierung der Beteiligten und Betroffenen	21
N2	Klärung der Evaluationszwecke	22
G1	Beschreibung des Evaluationsgegenstandes	31
G2	Kontextanalyse	31
G3	Beschreibung von Zwecken und Vorgehen	32
3. Planung der Evaluation – wichtigste Standards:		
N1	Identifizierung der Beteiligten und Betroffenen	21
N2	Klärung der Evaluationszwecke	22
N3	Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin	22
N4	Auswahl und Umfang der Informationen	23
N5	Transparenz von Werten	23
N7	Rechtzeitigkeit der Evaluation	24
N8	Nutzung und Nutzen der Evaluation	25
D1	Angemessene Verfahren	26
D2	Diplomatisches Vorgehen	26
D3	Effizienz von Evaluation	27
F1	Formale Vereinbarungen	28
F3	Vollständige und faire Überprüfung	29
F4	Unparteiische Durchführung und Berichterstattung	29
G1	Beschreibung des Evaluationsgegenstandes	31
G2	Kontextanalyse	31
G3	Beschreibung von Zwecken und Vorgehen	32
G4	Angabe von Informationsquellen	32



G5	Valide und reliable Informationen	33
G7	Analyse qualitativer und quantitativer Informationen	34
G8	Begründete Schlussfolgerungen	35
G9	Meta-Evaluation	35

4. Informationsgewinnung – wichtigste Standards:

N3	Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin	22
N4	Auswahl und Umfang der Informationen	23
N5	Transparenz von Werten	23
D1	Angemessene Verfahren	26
D2	Diplomatisches Vorgehen	26
F1	Formale Vereinbarungen	28
F2	Schutz individueller Rechte	28
F3	Vollständige und faire Überprüfung	29
G1	Beschreibung des Evaluationsgegenstandes	31
G2	Kontextanalyse	31
G3	Beschreibung von Zwecken und Vorgehen	32
G4	Angabe von Informationsquellen	32
G5	Valide und reliable Informationen	33
G6	Systematische Fehlerprüfung	34
G9	Meta-Evaluation	35

5. Informationsauswertung – wichtigste Standards:

N5	Transparenz von Werten	23
F4	Unparteiische Durchführung und Berichterstattung	29
G1	Beschreibung des Evaluationsgegenstandes	31
G2	Kontextanalyse	31
G6	Systematische Fehlerprüfung	34
G7	Analyse qualitativer und quantitativer Informationen	34
G8	Begründete Schlussfolgerungen	35
G9	Meta-Evaluation	35

6. Berichterstattung zur Evaluation – wichtigste Standards:

N1	Identifizierung der Beteiligten und Betroffenen	21
N4	Auswahl und Umfang der Informationen	23
N5	Transparenz von Werten	23
N6	Vollständigkeit und Klarheit der Berichterstattung	24

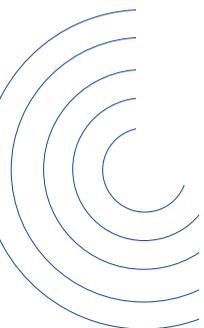
N7	Rechtzeitigkeit der Evaluation	24
N8	Nutzung und Nutzen der Evaluation	25
F2	Schutz individueller Rechte	28
F3	Vollständige und faire Überprüfung	29
F4	Unparteiische Durchführung und Berichterstattung	29
F5	Offenlegung der Ergebnisse	30
G1	Beschreibung des Evaluationsgegenstandes	31
G2	Kontextanalyse	31
G3	Beschreibung von Zwecken und Vorgehen	32
G4	Angabe von Informationsquellen	32
G8	Begründete Schlussfolgerungen	35
G9	Meta-Evaluation	35

7. Budgetierung der Evaluation – wichtigste Standards:

N2	Klärung der Evaluationszwecke	22
N4	Auswahl und Umfang der Informationen	23
D3	Effizienz von Evaluation	27
F1	Formale Vereinbarungen	28
G1	Beschreibung des Evaluationsgegenstandes	31
G3	Beschreibung von Zwecken und Vorgehen	32

8. Evaluationsvertrag – wichtigste Standards:

N1	Identifizierung der Beteiligten und Betroffenen	21
N2	Klärung der Evaluationszwecke	22
N3	Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin	22
N4	Auswahl und Umfang der Informationen	23
N7	Rechtzeitigkeit der Evaluation	24
D2	Diplomatisches Vorgehen	26
D3	Effizienz von Evaluation	27
F1	Formale Vereinbarungen	28
F2	Schutz individueller Rechte	28
F4	Unparteiische Durchführung und Berichterstattung	29
F5	Offenlegung der Ergebnisse	30
G1	Beschreibung des Evaluationsgegenstandes	31
G3	Beschreibung von Zwecken und Vorgehen	32
G9	Meta-Evaluation	35



9. Steuerung der Evaluation – wichtigste Standards:	
N1 Identifizierung der Beteiligten und Betroffenen	21
N2 Klärung der Evaluationszwecke	22
N3 Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin	22
N7 Rechtzeitigkeit der Evaluation	24
N8 Nutzung und Nutzen der Evaluation	25
D2 Diplomatisches Vorgehen	26
D3 Effizienz von Evaluation	27
F1 Formale Vereinbarungen	28
F2 Schutz individueller Rechte	28
F4 Unparteiische Durchführung und Berichterstattung	29
G3 Beschreibung von Zwecken und Vorgehen	32
G6 Systematische Fehlerprüfung	34
G9 Meta-Evaluation	35
10. Personelle Ausstattung der Evaluation – wichtigste Standards:	
N3 Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin	22
N4 Auswahl und Umfang der Informationen	23
D2 Diplomatisches Vorgehen	26
F4 Unparteiische Durchführung und Berichterstattung	29
G1 Beschreibung des Evaluationsgegenstandes	31

TRANSFORMATIONSTABELLE

Die nachfolgende Tabelle ermöglicht einen Vergleich der drei Standard-Sets, die – beginnend mit den Standards des us-amerikanischen Joint Committee on Standards for Educational Evaluation (2000) über die Standards der Schweizerischen Evaluationsgesellschaft bis zu den DeGEval-Standards – in ihrer Entwicklung verbunden sind. Die Tabelle ermöglicht insbesondere, die ausführlichen Erläuterungen und Fallbeispiele aus den Standards des Joint Committee (2000) heranzuziehen, die als deutschsprachige Buchveröffentlichung vorliegen.

■ Die Darstellung von Standards mit denselben Namen bedeutet nicht, dass diese identisch definiert sind; es ist jeweils erforderlich, die Texte daraufhin zu prüfen.

■ Die Inhalte der Standards K7 und K8 des Joint Committee sind nicht entfallen, sondern in andere Standards der DeGEval bzw. der SEVAL aufgenommen worden.

TRANSFORMATIONSTABELLE

Deutsche Gesellschaft für Evaluation	Schweizerische Evaluationsgesellschaft	Joint Committee on Standards (USA)
N1 Identifizierung der Beteiligten und Betroffenen	N1 Ermittlung der Beteiligten und Betroffenen	N1 Ermittlung der Beteiligten und Betroffenen
N2 Klärung der Evaluationszwecke	N2 Klärung der Evaluationsziele	fehlt
N3 Glaubwürdigkeit und Kompetenz des Evaluators / der Evaluatorin	N3 Glaubwürdigkeit	N2 Glaubwürdigkeit der Evaluatorin
N4 Auswahl und Umfang der Informationen	N4 Umfang und Auswahl der Informationen	N3 Umfang und Auswahl der Informationen
N5 Transparenz von Werten	N5 Transparenz der Bewertung	N4 Feststellung von Werten
N6 Vollständigkeit und Klarheit der Berichterstattung	N6 Vollständigkeit und Klarheit des Berichts	N5 Klarheit des Berichts
N7 Rechtzeitigkeit der Evaluation	N7 Rechtzeitigkeit der Berichterstattung	N6 Rechtzeitigkeit/Verbreitung des Berichts
N8 Nutzung und Nutzen der Evaluation	N8 Wirkung der Evaluation	N7 Wirkung der Evaluation
D1 Angemessene Verfahren	D1 Praktikable Verfahren	D1 Praktische Verfahren
D2 Diplomatisches Vorgehen	D2 Politische Tragfähigkeit	D2 Politische Tragfähigkeit
D3 Effizienz von Evaluation	D3 Kostenwirksamkeit	D3 Kostenwirksamkeit
entfällt	entfällt	K1 Unterst. der Dienstleistungsorient.
F1 Formale Vereinbarungen	K1 Formale Vereinbarungen	K2 Formale Vereinbarungen
F2 Schutz individueller Rechte	K2 Schutz individueller Rechte	K3 Schutz individueller Menschenrechte
	K3 Menschlich gestaltete Interaktion	K4 Human gestaltete Interaktion
F3 Vollständige und faire Überprüfung	K4 Vollständige und faire Einschätzung	K5 Vollständige und faire Einschätzung
F5 Offenlegung der Ergebnisse	K5 Offenlegung der Ergebnisse	K6 Offenlegung der Ergebnisse
	K6 Deklaration von Interessenkonflikten	K7 Deklaration von Interessenkonflikten
	in D3 Kostenwirksamkeit	K8 Finanzielle Verantwortlichkeit
G1 Beschreibung des Evaluationsgegenstandes	G1 Dokumentation des Evaluationsgegenstandes	G1 Programmdokumentation
G2 Kontextanalyse	G2 Kontextanalyse	G2 Kontextanalyse
G3 Beschreibung von Zwecken und Vorgehen	G3 Beschreibung von Zielen und Vorgehen	G3 Beschreibung von Zielen und Vorgehen
G4 Angabe von Informationsquellen	G4 Verlässliche Informationsquellen	G4 Verlässliche Informationsquellen
G5 Valide und reliable Informationen	G5 Valide und reliable Informationen	G5 Valide Informationen
		G6 Reliable Informationen
G6 Systematische Fehlerprüfung	G6 Systemat. Informationsüberprüfung	G7 Systemat. Informationsüberprüfung
G7 Analyse qualitativer und quantitativer Informationen	G7 Analyse qualitativer und quantitativer Informationen	G8 Analyse quantitativer Informationen
		G9 Analyse qualitativer Informationen
G8 Begründete Schlussfolgerungen	G8 Begründete Schlussfolgerungen	G10 Begründete Schlussfolgerungen
F4 Unparteiische Durchführung und Berichterstattung	G9 Unparteiische Berichterstattung	G11 Unparteiische Berichterstattung
G9 Meta-Evaluation	G10 Meta-Evaluation	G12 Meta-Evaluation

LITERATUR VORBEMERKUNG

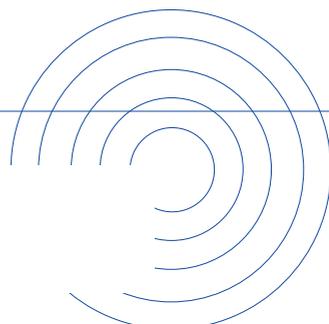
Die nachfolgende Liste umfasst ausschließlich die in der Einleitung und die in den Erläuterungen zu den Standards aufgeführten Quellen. Vorwiegend werden dort deutschsprachige Texte zitiert, wenn möglich Bücher, in Ausnahmefällen auch Artikel. Die zitierte Literatur hat entweder für Evaluationsfelder und -modelle oder Disziplinen übergreifende Bedeutung oder gibt exemplarisch Hinweise zum Umgang mit den in einem bestimmten Standard angesprochenen Themen. Dort, wo prominente – auch in Deutschland verbreitete – englischsprachige oder internationale Referenztexte vorliegen, wurde auch auf solche hingewiesen.

■ Die Verteilung der Belege auf die Standards ist sehr ungleichgewichtig. Für viele Standards konnte keine einzige einschlägige deutschsprachige Referenz identifiziert werden. Für weitere gibt es lediglich einzelne Bezugstexte. Dagegen gibt es für einige der Genauigkeitsstandards und für den Fairnessstandard F2 „Schutz individueller Rechte“ zahlreiche einschlägige deutschsprachige Quellen, von denen jeweils einige Beispiele aufgeführt sind. Zu großen Teilen sind diese nicht auf ihre Anwendbarkeit für Evaluationen geprüft. Es kann davon ausgegangen werden, dass eine Übertragung möglich ist, wenn erforderliche Anpassungen vorgenommen werden.

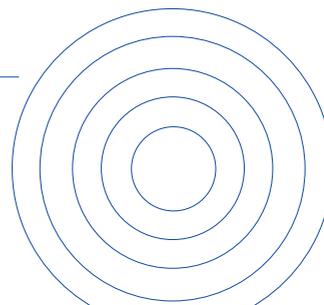
■ Im Unterschied dazu finden sich zu den 30 „Standards für die Evaluation von Programmen“ des Joint Committee (2000/1994) jeweils mindestens zwei, teilweise bis zu zehn einschlägige englischsprachige Literaturquellen, auf die hiermit ergänzend verwiesen wird. Der Zugang dazu wird durch die „Transformationstabelle“ im Abschnitt D dieses DeGEval-Papiers erleichtert.

■ Die Literaturlage im deutschsprachigen Raum ist auch Ausdruck dafür, dass sich Evaluation hier noch in der Professionalisierungsphase befindet. Im Rahmen der Verbreitung und Diskussion der „Standards für Evaluation“ der DeGEval wird auch erwartet, dass zunehmend deutschsprachige Literatur und andere Quellen beigesteuert werden.

- AEA** American Evaluation Association (1999): Leitprinzipien für Evaluatoren und Evaluatorinnen (Original: Guiding Principles for Evaluators 1995); in: Beywl, W., Schobert, B. (1999): Evaluation - Controlling - Qualitätsmanagement in der betrieblichen Weiterbildung. Kommentierte Auswahlbibliographie, 3. Auflage. Bielefeld. S. 94-101.
- AERA/APA/NCME** American Educational Research Association/American Psychological Association/National Council on Measurement in Education (1999): The Standards for Educational and Psychological Testing. Washington.
- ALLGÄUER, R.** (1997): Evaluation macht uns stark! Frankfurt am Main.
- ARBEITSKREIS** Deutscher Markt- und Sozialforschungsinstitute e.V./Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V./Berufsverband Deutscher Markt- und Sozialforscher e.V./Deutsche Gesellschaft für Online-Forschung e.V. (D.G.O.F.). (2001): Standards zur Qualitätssicherung für Online-Befragungen. Frankfurt a. M. (<http://www.adm-ev.de>).



- BEYWL, W.** (1988): Zur Weiterentwicklung der Evaluationsmethodologie. Grundlegung, Konzeption und Anwendung eines Modells der responsiven Evaluation. Frankfurt (Reprint, Köln 1999).
- BEYWL, W.** (2001): Konfliktfähigkeit der Evaluation und die "Standards für Evaluation"; in: Sozialwissenschaften und Berufspraxis Nr. 2/2001, S. 151-164.
- BEYWL, W.** u. a. (2001) (Hg.): Evaluation im Alltag. Münster.
- BORTZ, J.** (1999): Statistik für Sozialwissenschaftler. Berlin.
- BORTZ, J.** /Döring, N. (1995): Forschungsmethoden und Evaluation. Berlin.
- BUHREN, C. G.** /Killus, D./Müller, S. (1998): Wege und Methoden der Selbstevaluation. Ein praktischer Leitfaden für Schulen. Dortmund.
- BUNDESAMT** für Gesundheit (BAG) (1997) (Hg.): Leitfaden für die Planung von Projekt- und Programmevaluationen. Bern (Bundesamt für Gesundheit, Fachbereich Evaluation).
- BUSSMANN, W.** /Klöti, U./Knoepfel P. (1997): Einführung in die Politikevaluation. Zürich.
- CANADIAN** Evaluation Society (o. J.): CES Guidelines for Ethical Conduct. Ottawa (<http://www.evaluationcanada.ca>).
- CRONBACH, L.J.** u.a. (1980): Toward Reform of Program Evaluation. San Francisco.
- DEUTSCHE** Forschungsgemeinschaft/ Kaase, M. (1999) (Hg.): Qualitätskriterien der Umfrageforschung – Denkschrift. Berlin.
- DEUTSCHE** Gesellschaft für Erziehungswissenschaft (1997): Standards erziehungswissenschaftlicher Forschung; in: Friebertshäuser, B./Prenzel A. (Hg.): Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft. München. S. 858-863.
- DEUTSCHE** Norm DIN 33430. Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen. Entwurf Oktober 2000. Berlin (www.mybeuth.de).
- DIEKMANN, A.** (1995): Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek.
- DPRG/GPRA** Deutsche Public-Relations-Gesellschaft/Gesellschaft Public Relations Agenturen (2000): PR-Evaluation. Messen – Analysieren – Bewerten. Bonn.
- EMCDDA** European Monitoring Centre for Drugs and Drug Addiction (1998): Guidelines for the Evaluation of Drug Prevention: A Manual for Programme-Planners and Evaluators. Lisbon.
(Eine deutschsprachige Version ist in Vorbereitung.)
- EUROPEAN** Union– Regional Policy and Cohesion (1996): MEANS-Handbook Nr. 5. Quality Assessment of Evaluation Reports: A Framework. Bruxelles.
- FABMANN, H.** (2001): Probleme der Evaluation von Programmen im Bereich der Rehabilitation; in: Sozialwissenschaften und Berufspraxis Nr. 2/2001, S.133-149.
- FLICK, U.** (1995): Qualitative Forschung. Reinbek.
- GROHMANN, R.** (1997): Das Problem der Evaluation in der Sozialpädagogik. Bezugspunkte zur Weiterentwicklung der evaluationstheoretischen Reflexion. Frankfurt.
- GUBA, E.G.** /Lincoln, Y.S. (1981): Effective Evaluation. San Francisco.



- HAGER, W.** /Patry J.L./Brezing, H. (2000): Handbuch Evaluation psychologischer Interventionsmaßnahmen. Standards und Kriterien. Bern.
- HEINER, M.** (1988) (Hg.): Selbstevaluation in der Sozialen Arbeit. Freiburg.
- HEINER, M.** (1998) (Hg.): Experimentierende Evaluation. Ansätze zur Entwicklung lernender Organisationen. Weinheim.
- HELLSTERN, M.** /Wollmann H. (1984) (Hg.): Handbuch der Evaluierungsforschung. Opladen.
- IHK/ESOMAR** Internationale Handelskammer/Europäische Gesellschaft für Meinungs- und Marketing-Forschung (o. J.): Internationaler Kodex für die Praxis der Markt- und Sozialforscher. Amsterdam (<http://www.planung-analyse.de>).
- JOINT** Committee on Standards for Educational Evaluation (2000): Handbuch der Evaluationsstandards, 2. Auflage. Opladen. (Original: The Program Evaluation Standards, 1994. 2. Aufl. Thousand Oaks)
- KROMREY, H.** (1998): Empirische Sozialforschung. Opladen.
- KROMREY, H.** (2000): Die Bewertung von Humandienstleitungen. Fallstricke bei der Implementations- und Wirkungsforschung sowie methodische Alternativen; in: Müller-Kohlenberg/Münstermann (2000), S. 19-55.
- KROMREY, H.** (2001): Evaluation – ein vielschichtiges Konzept; in: Sozialwissenschaften und Berufspraxis Nr. 2/2001, S. 105-131.
- LAMNEK, S.** (1995): Qualitative Sozialforschung, Bd. 1 u. 2.. 3. Auflage. München.
- MAYRING, P.** (1999): Einführung in die qualitative Sozialforschung. 4. Auflage. Weinheim.
- MILES, M.B.** /Huberman, A.M. (1994): Qualitative data analysis: An expanded source book. Thousand Oaks.
- MÜLLER-KOHLBERG, H.** /Münstermann, K. (2000) (Hg.): Qualität von Humandienstleistungen. Evaluation und Qualitätsmanagement in Sozialer Arbeit und Gesundheitswesen. Opladen.
- MÜLLER-KOHLBERG, H.** /Beywl, W. (2001): Der Arbeitskreis Soziale Dienstleistungen. Entwurf eines Arbeitsplans, in: Deutsche Gesellschaft für Evaluation e.V. (Hg.): Evaluation – Reformmotor oder Reformbremse. Köln 2001.
- OECD** (1998): Review of the DAC Principles for Evaluation of Development Assistance. Paris.
- OECD** (1998a): Public Management Service Best Practice Guidelines for Evaluation. Paris.
- OWEN, J. M.** /Rogers, P. J. (1999): Program evaluation: Forms and approaches, Thousand Oaks.
- PATTON, M.Q.** (1997): Utilization-Focused Evaluation. 3. Auflage. Thousand Oaks.
- PATTON, M.Q.** (1998): Die Entdeckung des Prozessnutzens. Erwünschtes und unerwünschtes Lernen durch Evaluation; in: Heiner (1998), S. 55-66.
- RÜRUP, B.** /Hansmeyer, K.-H. (1984): Staatswirtschaftliche Planungsinstrumente. 3.Auflage. Düsseldorf.
- SCRIVEN, M.** (1991): Evaluation Thesaurus. 4. Auflage. Newbury Park.
- SPIEGEL, H.V.** (1993) (Hg.): Aus Erfahrung lernen. Qualifizierung durch Selbstevaluation. Münster.
- STOCKBAUER, U.** (2000): Was macht Evaluationen nützlich? Überblick zum Forschungsstand – Ergebnisse von Fallstudien; in: Müller-Kohlenberg/Münstermann (2000), S. 121-128.
- STOCKMANN, R.** (2000) (Hg.): Evaluationsforschung – Grundlagen und ausgewählte Forschungsfelder. Opladen.

- STUFFLEBEAM, D.L.** (2001): Evaluation Models, New Directions for Evaluation, No. 89, Spring. San Francisco.
- UHL, A.** (1999): Evaluation. in: Stimmer, F. (Hg.): Suchtlexikon. München.
- UHL, A.** (2000): The Limits of Evaluation; in: Neaman, R./Nilson, M./Solberg, U.(Hg.): Evaluation - A Key Tool for Improving Drug Prevention. EMCDDA Scientific Monograph Series, No 5. Lisbon.
- WEDEL, H.V.** (1998): Erfolgskontrolle finanzwirksamer Maßnahmen in der öffentlichen Verwaltung (Schriftenreihe der Bundesbeauftragten für die Wirtschaftlichkeit in der Verwaltung, Bd. 2). Stuttgart.
- WIDMER, T.** (1996): Meta-Evaluation. Kriterien zur Bewertung von Evaluationen. Bern.
- WIDMER, T.** /Landert, C./Bachmann, N. (2000): Evaluations-Standards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards). Bern/Genf.
- WOTTAWA, H.** /Thierau, H. (1998): Lehrbuch Evaluation, 2. Auflage. Bern.
- WULF, C.** (1972) (Hg.): Evaluation. Beschreibung und Bewertung von Unterricht, Curricula und Schulversuchen. München.

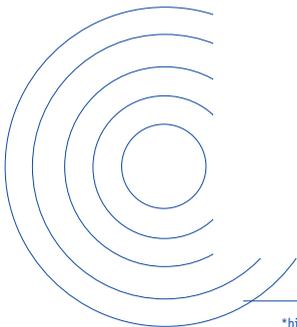
MITARBEITER UND MITARBEITERINNEN AN DER STANDARD-ENTWICKLUNG

Die Mitglieder der Kommission "Standards für Evaluation"

Dr. Wolfgang Beywl	Dieter Brauns	Dr. Hansjörg Drewello
Dr. Andreas Hellmann	Thomas Kuby	Sabine Müller
Dr. Alfred Uhl	Gerald Wagner	Hein Winnubst*

Kommentatoren und Kommentatorinnen

Dr. Alois Basler	Karin Haubrich	Dr. Christian Lüders
Alexandra Caspari	Thomas Hochgesang	Dirk Scheffler
Dr. Karin Fischer-Bluhm	Ursula Hütte	Gerlinde Struhkamp
Werner Fuchs	Marc Jelitto	Prof. Dr. Reinhard Stockmann
	Dr. Kathleen Toepel	



SUMMARY OF EVALUATION STANDARDS

Presented by the German Evaluation Society (DeGEval-Standards)

The following Evaluation Standards were ratified by the general assembly of the German Evaluation Society (Deutsche Gesellschaft für Evaluation – DeGEval) on October 4th, 2001. They are the result of a two-year discussion and preparation process which included a membership survey, an appointed Standards Committee, and a review process.

■ The twenty-five DeGEval-Standards are organized in four groups. This structure as well as many Standards, including titles and descriptive statements, were stimulated by the „Program Evaluation Standards“ of the US-American „Joint Committee on Standards for Educational Evaluation“ and adapted to the requirements of evaluation in Germany and Austria. The DeGEval-Standards were also inspired by the Swiss adaptation of the Joint Committee Standards which provides a generalization of these standards from educational to more diverse settings.

■ In its German original, this short version of the DeGEval-Standards is accompanied by a 30-page document which includes a clarification of the aims and the scope of the Standards, definitions of evaluation and other key concepts, an overview of different approaches to evaluation, comments on the application of the Standards, and a description of the development of the document itself as well as the review process.

■ By the end of 2004, the DeGEval-Standards will have been subjected to a second review process which will include research societies as well as professional associations.

For more information, please see <http://www.degeval.de> or contact Wolfgang Beywl (wolfgang.beywl@univention.org), former Chair of the Standards Committee and board member of the German Evaluation Society.

Evaluations should feature four basic attributes:

Utility – Feasibility – Propriety – Accuracy.

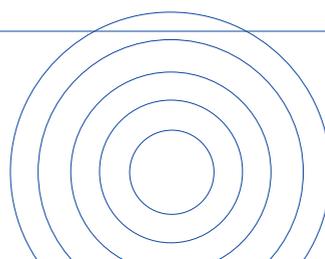
UTILITY **U** The Utility Standards are intended to ensure that an evaluation is guided by both the clarified purposes of the evaluation and the information needs of its intended users.

U1 Stakeholder Identification

Persons or groups involved in or affected by the evaluand should be identified, so that their interests can be clarified and taken into consideration when designing the evaluation.

U2 Clarification of the Purposes of the Evaluation

The purposes of the evaluation should be stated clearly, so that the stakeholders can provide relevant comments on these purposes, and so that the evaluation team knows exactly what it is expected to do.



U3 Evaluator Credibility and Competence

The persons conducting an evaluation should be trustworthy as well as methodologically and professionally competent, so that the evaluation findings achieve maximum credibility and acceptance.

U4 Information Scope and Selection

The scope and selection of the collected information should make it possible to answer relevant questions about the evaluand and, at the same time, consider the information needs of the client and other stakeholders.

U5 Transparency of Values

The perspectives and assumptions of the stakeholders that serve as a basis for the evaluation and the interpretation of the evaluation findings should be described in a way that clarifies their underlying values.

U6 Report Comprehensiveness and Clarity

Evaluation reports should provide all relevant information and be easily comprehensible.

U7 Evaluation Timeliness

The evaluation should be initiated and completed in a timely fashion, so that its findings can inform pending decision and improvement processes.

U8 Evaluation Utilization and Use

The evaluation should be planned, conducted, and reported in ways that encourage attentive follow-through by stakeholders and utilization of the evaluation findings.

FEASIBILITY **F** The Feasibility Standards are intended to ensure that an evaluation is planned and conducted in a realistic, thoughtful, diplomatic, and cost-effective manner.

F1 Appropriate Procedures

Evaluation procedures, including information collection procedures, should be chosen so that the burden placed on the evaluand or the stakeholders is appropriate in comparison to the expected benefits of the evaluation.

F2 Diplomatic Conduct

The evaluation should be planned and conducted so that it achieves maximal acceptance by the different stakeholders with regard to evaluation process and findings.

F3 Evaluation Efficiency

The relation between cost and benefit of the evaluation should be appropriate.

PROPRIETY **P** The Propriety Standards are intended to ensure that in the course of the evaluation all stakeholders are treated with respect and fairness.

P1 Formal Agreement

Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated to adhere to all conditions of the agreement or to renegotiate it.

P2 Protection of Individual Rights

The evaluation should be designed and conducted in a way that protects the welfare, dignity, and rights of all stakeholders.

P3 Complete and Fair Investigation

The evaluation should undertake a complete and fair examination and description of strengths and weaknesses of the evaluand, so that strengths can be built upon and problem areas addressed.

P4 Unbiased Conduct and Reporting

The evaluation should take into account the different views of the stakeholders concerning the evaluand and the evaluation findings. Similar to the entire evaluation process, the evaluation report should evidence the impartial position of the evaluation team. Value judgments should be made as unemotionally as possible.

P5 Disclosure of Findings

To the extent possible, all stakeholders should have access to the evaluation findings.

ACCURACY **A** The Accuracy Standards are intended to ensure that an evaluation produces and discloses valid and useful information and findings pertaining to the evaluation questions.

A1 Description of the Evaluand

The evaluand should be described and documented clearly and accurately, so that it can be unequivocally identified.

A2 Context Analysis

The context of the evaluand should be examined and analyzed in enough detail.

A3 Described Purposes and Procedures

Object, purposes, questions, and procedures of an evaluation, including the applied methods, should be accurately documented and described, so that they can be identified and assessed.

A4 Disclosure of Information Sources

The information sources used in the course of the evaluation should be documented in appropriate detail, so that the reliability and adequacy of the information can be assessed.

A5 Valid and Reliable Information

The data collection procedures should be chosen or developed and then applied in a way that ensures the reliability and validity of the data with regard to answering the evaluation questions.

A6 Systematic Data Review

The data collected, analyzed, and presented in the course of the evaluation should be systematically examined for possible errors.

A7 Analysis of Qualitative and Quantitative Information

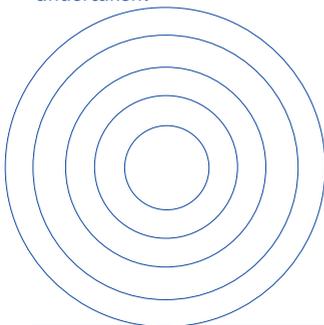
Qualitative and quantitative information should be analyzed in an appropriate, systematic way, so that the evaluation questions can be effectively answered.

A8 Justified Conclusions

The conclusions reached in the evaluation should be explicitly justified, so that the audiences can assess them.

A9 Meta-Evaluation

The evaluation should be documented and archived appropriately, so that a Meta-Evaluation can be undertaken.



Das Handbuch der Evaluationsstandards

Joint Committee on Standards for Educational Evaluation/ James R. Sanders (Hrsg.)

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.

2., durchgesehene Auflage 2000

310 Seiten. Kart.

34,90 EUR

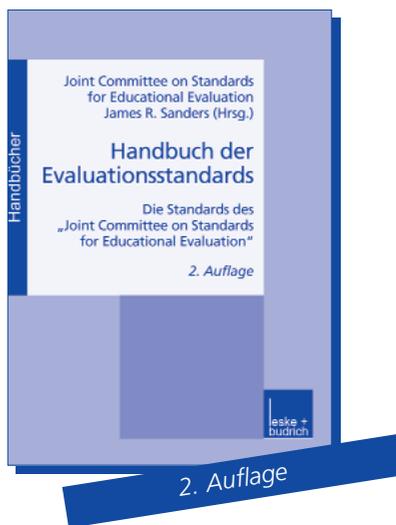
ISBN 3-8100-2766-9

Die amerikanischen ‚Program Evaluation Standards‘ sind ein Referenzwerk der Evaluationspraxis. Der Hauptteil enthält die 30 Standards, geordnet nach den Bewertungsdimensionen Nützlichkeit, Korrektheit, Anwendbarkeit sowie Genauigkeit. Jeder Standard wird detailliert beschrieben und erklärt. Außerdem illustrieren Praxisbeispiele aus Schule, Hochschule, betrieblicher Bildung und Sozialer Arbeit ihre Anwendung.

Stimmen zum Handbuch:

„Das Handbuch gibt einen reichhaltigen Überblick über die in Wissenschaft, Politik und Verwaltung im In- und Ausland gängigen Qualitätsanforderungen an Evaluationsstudien im weitesten Sinne. Das Buch verdient es, von all jenen gelesen zu werden, die sich als Auftragnehmer oder Auftraggeber im öffentlichen Sektor mit derartigen Studien zu befassen haben.“

Prof. Dr. Peter Knoepfel,
Direktor des Institut des hautes études en administration publique (IDHEAP),
Lausanne



„Die Übersetzung dieses Standardwerkes zur Evaluation erscheint gerade zur rechten Zeit. Es bietet eine klare Orientierung für eine systematische, empirisch fundierte Qualitätsentwicklung auch in der Sozialen Arbeit. (...) Ich werde das Buch in meinen Evaluationskursen und -seminaren innerhalb und außerhalb der Universität empfehlen und als Text einsetzen.“

Dr. Maja Heiner,
Professorin für Sozialpädagogik an der
Universität Tübingen

„Das Handbuch ist nicht nur eine entscheidende Informationsquelle, um vorhandene Qualitätssysteme auf den Prüfstand zu stellen, sondern liefert auch Instrumente und die in die Zukunft weisende Arbeitsphilosophie für die Entwicklung neuer Evaluationskonzepte.“

Prof. Gottfried Kleinschmidt,
Landesinstitut für Erziehung und Unterricht
Stuttgart

■ **Leske + Budrich**

**Arbeitsmarkt- und
Beschäftigungspolitik**

Agrarpolitik

Arbeitsschutz

Strukturpolitik

Betriebliches Management

Kinder- und Jugendhilfe

Ausstellungen und Museen

**Kriminal- und
Justizpolitik**

Humanitäre Aktionen

Public Relations

Sozialpolitik

**Stiftungen und Non-
Profit-Organisationen**

Unterricht

Suchtprävention

Verkehr

Verbraucherschutz