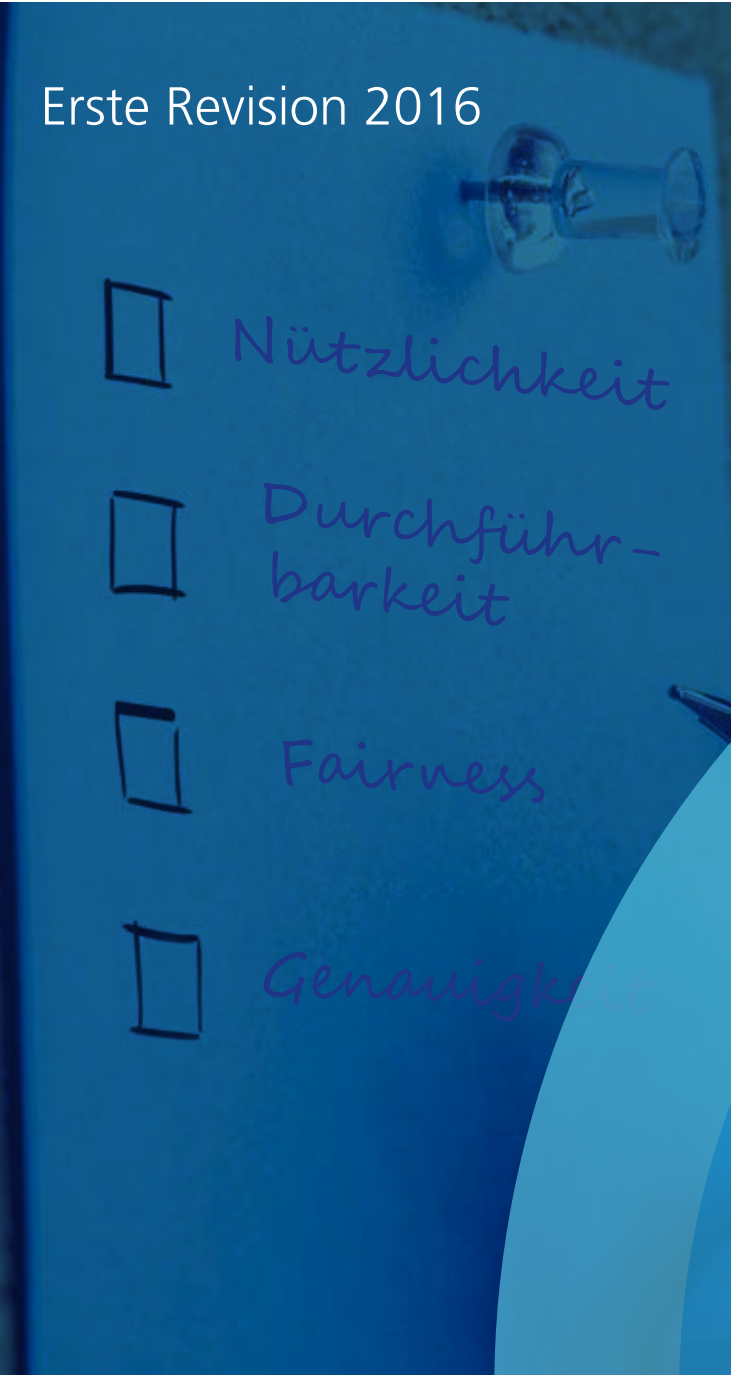


DeGEval – Gesellschaft für Evaluation Standards für Evaluation

Erste Revision 2016

- 
- Nützlichkeit
 - Durchführ-
barkeit
 - Fairness
 - Genauigkeit

Online-Fassung

Hinweis: Leere Seiten der Druck-Fassung fehlen in dieser Online-Fassung.
Die Seitenzählung ist jedoch identisch mit der gedruckten Ausgabe.

Die Deutsche Bibliothek – CIP Einheitsaufnahme
DeGEval – Gesellschaft für Evaluation e.V.
Standards für Evaluation
Erste Revision 2016
Mainz, 2017, ISBN 978-3-941569-06-5

1. Auflage

Herausgeber: DeGEval – Gesellschaft für Evaluation e.V.
Mitglieder der Revisionsgruppe:
Wolfgang Böttcher (Leitung bis September 2015)
Alexandra Caspari
Jan Hense (Leitung ab September 2015)
Michael Kalman
Wolfgang Meyer

Vertrieb: DeGEval – Gesellschaft für Evaluation e.V.
E-Mail: info@degeval.de

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 978-3-941569-06-5

Satz: Tanja Labs
Titelbild: © Gajus – shutterstock.com
Druck: Druckerei Zeidler, Mainz-Kastel

Printed in Germany

Verabschiedet durch die Mitgliederversammlung
der DeGEval e.V. am 21. September 2016

DeGEval – Gesellschaft für Evaluation
Standards für Evaluation

Erste Revision 2016

INHALTSVERZEICHNIS

Vorwort.....	5
1 – Einleitung.....	7
2 – Entstehung und Entwicklung der Standards	11
3 – Die Standards für Evaluation.....	17
4 – Konzeptionelle und begriffliche Grundlagen	23
Zielsetzung der Standards.....	24
Definition und Begriffsklärungen.....	25
Anwendung und Reichweite der Standards.....	28
5 – Erläuterungen zu den Standards für Evaluation	31
Nützlichkeit.....	34
Durchführbarkeit.....	39
Fairness	41
Genauigkeit	44
6 – Funktionales Inhaltsverzeichnis.....	51
7 – Glossar der Standards für Evaluation	63
Literatur	73

VORWORT

Seit Erscheinen ihrer ersten Fassung im Jahr 2002 definieren die „Standards für Evaluation“ der DeGEval Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit als die vier übergreifenden Merkmale, die gute Evaluationen kennzeichnen. Die „Standards“ stellen damit nicht nur einen zentralen Bezugspunkt für das professionelle Selbstverständnis der DeGEval und ihrer Mitglieder dar. Vielmehr sollen die Standards weit über die Gesellschaft hinaus Wirkungen entfalten. Sie sollen Evaluierenden ebenso wie Auftraggebenden und allen weiteren an Evaluationen Beteiligten Orientierung geben, wie gute Evaluationen zu gestalten sind. Zugleich sollen sie die Qualität von Evaluationen beurteilbar machen, als Orientierung für die Aus- und Weiterbildung dienen und stellen somit eine wichtige Grundlage für die Professionalisierung der Evaluation dar.

Nachdem schon 2002 mit erstem Erscheinen eine Revision nach wenigen Jahren angedacht war, wurde diese zunächst nach einer Mitgliederbefragung zurückgestellt, da kein grundlegender Überarbeitungsbedarf erkennbar war und man die Ressourcen eher in eine weitere Dissemination und Implementierung der Standards investieren wollte. So dauerte es schließlich eine Dekade, bis aufgrund verschiedener interner und externer Entwicklungen 2012 der Revisionsprozess in Angriff genommen wurde, dessen Ergebnis Sie nun in den Händen halten.

Die vorliegende erste Revision wurde nach Berücksichtigung unterschiedlicher Revisionsperspektiven und umfassender Einbeziehungen von Fachexpertise und der gesamten Mitgliedschaft erarbeitet und von der DeGEval-Mitgliederversammlung 2016 verabschiedet. Die hohe Qualität der ursprünglichen Fassung der Standards wurde dadurch deutlich, dass die Bearbeitung inkrementell erfolgen konnte und keine grundlegenden Änderungen an Inhalten und Struktur erforderlich waren. Investiert wurde aber v.a. in viele Aktualisierungen im Kleinen, in denen u.a. fachliche Weiterentwicklungen der vergangenen Dekade aufgegriffen wurden, sowie in die Verständlichkeit des Dokuments. In diesem Rahmen wurde auch ein Glossar der Standards für Evaluation erstellt, das Sie am Ende dieser Publikation finden und dessen Einträge im Text einheitlich durch einen Verweispfeil (→) als Glossarbegriffe gekennzeichnet werden.

Gedankt sei an dieser Stelle den Autorinnen und Autoren der Originalfassung, den im Revisionsprozess beteiligten Expertinnen und Experten sowie allen Kommentatorinnen und Kommentatoren aus der Mitgliedschaft. Ein besonderer Dank gilt der Arbeitsgruppe der DeGEval („Revisionsgruppe“), die den gesamten Überarbeitungsprozess getragen hat. Ihr gehörten an: Wolfgang Böttcher (Leitung bis September 2015), Alexandra Caspari, Jan Hense (Leitung ab September 2015), Michael Kalman und Wolfgang Meyer. In leichter Abwandlung der vier Kernstandards lässt sich sagen, dass ihr Beitrag wesentlich geholfen hat, die Nützlichkeit, Anwendbarkeit, Ausgewogenheit und Präzision der DeGEval-Standards ein gutes Stück voranzubringen.

Prof. Dr. Philipp Mayring

Vorstandsvorsitzender der Gesellschaft für Evaluation (DeGEval e.V.)

EINLEITUNG



Evaluation hat nicht nur in Deutschland und Österreich in den letzten Jahren eine bemerkenswerte Karriere vorzuweisen. Sie gewinnt in vielen Staaten und in den meisten Bereichen gesellschaftspolitischer Gestaltung an Bedeutung. Mit mehr als einem Dutzend Arbeitskreisen, die in unterschiedlichen Handlungsfeldern aktiv sind, belegt die DeGEval die Anwendungsbreite der Evaluation sowie die Notwendigkeit, Evaluation immer auch gegenstandsbezogen zu denken. Trotz aller Vielfältigkeit hat Evaluation als generisches Konzept immer einen gemeinsamen Wesenskern, weshalb allgemeine Gütestandards sinnvoll und wichtig sind.

Die vorliegende Publikation ist eine Überarbeitung der seit 2002 gültigen Ursprungsfassung der DeGEval-Standards. Mit den „Standards für Evaluation“ wollen die in der **DeGEval – Gesellschaft für Evaluation** organisierten Personen und Organisationen einen Orientierungsrahmen zur Bestimmung einer „guten“ Evaluation geben. Die Standards beanspruchen, die Idee einer professionellen Evaluation zu konkretisieren.

Das vorliegende Dokument enthält neben dieser Einleitung folgende Bestandteile:

- Kapitel 2 geht auf die Entstehung, den Revisionsprozess und die geplante zukünftige Entwicklung der Standards ein.
- Kapitel 3 enthält die Standards im eigentlichen Sinne, also je eine Kurzbezeichnung und den Wortlaut aller 25 Einzelstandards, gruppiert nach den vier Standardgruppen Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit.
- Kapitel 4 fasst in knapper Form die wesentlichen konzeptionellen und begrifflichen Grundlagen für das Verständnis der Standards zusammen.
- Kapitel 5 enthält eine begleitende Kommentierung der 25 Einzelstandards, die das Verständnis vertiefen soll, indem jeweils erstens eine Begründung für den Einzelstandard gegeben wird und zweitens einige Umsetzungshinweise, die in der Praxis relevant werden können.
- Kapitel 6 beinhaltet ein **funktionales Inhaltsverzeichnis**, das aufzeigt, welche Standards in welchen Phasen eines idealtypischen Evaluationsprozesses besonders relevant sind.
- Kapitel 7 definiert in einem **Glossar** zentrale Begriffe, wie sie in den Standards verstanden werden, und setzt sie untereinander in Beziehung. Im Text sind alle Glossareinträge mit einem →Pfeil markiert).

Zusätzlich liegen eine Reihe von Stellungnahmen zu den Standards vor, die diese aus Sicht verschiedener Querschnittsthemen kommentieren und den jeweiligen Stellenwert für das Querschnittsthema erläutern. Im Einzelnen handelt es sich um die vier Perspektiven **Ethik, Gender, Methoden sowie Aus- und Weiterbildung in der Evaluation**. Diese Stellungnahmen wurden auf Bitte der die Revision verantwortenden Arbeitsgruppe durch Fachpersonen für die jeweilige Thematik erstellt und werden gleichzeitig mit der Druckfassung dieser revidierten Standards veröffentlicht.

ENTSTEHUNG UND ENTWICKLUNG DER STANDARDS



Die **DeGEval – Gesellschaft für Evaluation** hatte im Jahr 2000 eine Mitgliederbefragung mit dem Ziel durchgeführt, ein Meinungsbild zur Erarbeitung von Standards für Evaluationen bzw. Leitlinien für Evaluatorinnen und Evaluatoren zu erhalten. Die deutliche Mehrheit plädierte dafür, vorrangig Standards für Evaluation zu erarbeiten. Viele Mitglieder wünschten sich entweder zusätzlich 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 Evaluatorinnen und Evaluatoren beschreiben. 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 und das Arbeitsprogramm enthielten den Auftrag, die DeGEval-Standards an die Standards des Joint Committee on Standards for Educational Evaluation (JCSEE-Standards 2000/1994) und die mit diesen verwandten Standards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards 2001) anzuschließen. Damit hatte man sich dafür entschieden, an eine 20-jährige fachliche Entwicklungshistorie anzuknüpfen. Einer der tragenden Gründe war die Überzeugung, hierdurch den internationalen Erfahrungsaustausch zu erleichtern und die bereits vorliegenden Materialien und Fachveröffentlichungen nutzbar machen zu können (vgl. Joint Committee 2000).

Die Erstfassung der „Standards für Evaluation“ der DeGEval wurde von einer Kommission aus Evaluierenden und Auftraggebenden erstellt. Sie trat mit dem einstimmigen Beschluss der Mitgliederversammlung am 4. Oktober 2001 in Kraft. Die Standards waren seitdem wichtige Referenz für mit Evaluation befassten Personen und Organisationen und stellen bis heute das wichtigste inhaltliche Produkt der DeGEval dar.

Selbst einer selbstkritischen, evaluativen Haltung verpflichtet, war eine Revision der Standards mit ihrer ersten Verabschiedung von Anfang an geplant. Eine umfassende Mitgliederbefragung 2004 ergab jedoch, dass die Mitgliedschaft keinen grundlegenden Bedarf für eine Revision der Originalfassung sah, sondern eher deren weitere Verbreitung und Implementierung vorantreiben wollte. Im Jahr 2012, eine Dekade nach ihrem ersten Inkrafttreten, ergab sich jedoch aus verschiedenen internen und externen Entwicklungen heraus der Impuls, die Standards einer Überprüfung und Überarbeitung zu unterziehen.

- Ein Jahr zuvor war nach einem langen Diskussionsprozess die dritte Auflage der „Program Evaluation Standards des Joint Committee on Standards for Educational Evaluation“ erschienen (Yarbrough, Shulha, Hopson & Caruthers, 2011). Nachdem die DeGEval-Standards wesentlich auf der zweiten Auflage basierten, stellte sich die Frage, inwiefern Veränderungen der dritten Auflage wie etwa die Einführung einer fünften Kriteriendimension „Meta-Evaluation“ nachvollzogen werden sollten.
- Andere nationale Evaluationsgesellschaften planten zeitgleich die Aufnahme eines Revisionsprozesses ihrer Standards. Maßgeblich war v.a. unsere schweizerische Schwestergesellschaft SEVAL, deren Standards ebenfalls ein wichtiger Einfluss auf die DeGEval-Standards gewesen waren.
- In den Jahren seit Erscheinen der Originalfassung gab es in dem immer noch jungen Feld der Evaluation verschiedene fachliche Entwicklungen theoretischer und methodischer Art, deren Einfluss auf die Standards zu überprüfen war.
- Insbesondere durch eine in jüngerer Zeit deutlich verstärkte Forschung über Evaluation stand inzwischen eine bessere Evidenzbasis zur Begründung fachlicher Standards zur Verfügung.
- In einer Revision wurde die Chance gesehen, neue Impulse für eine weitere Verbreitung und verstärkte Implementation der Standards zu geben.
- Schließlich ist es ein Gebot einer eigenen evaluativen Grundhaltung, auch Bewährtes immer wieder einer kritischen Prüfung zu unterziehen und auf mögliche Verbesserungen hin zu diskutieren.

Aufgrund dessen wurden 2012 durch den Vorstand Beteiligte der Entwicklung der Originalfassung sowie interessierte Mitglieder eingeladen, die Frage einer möglichen Revision zu diskutieren. Die Diskussion erbrachte den Konsens, dass nur moderater Änderungsbedarf gesehen wurde, der sich u.a. auf folgende Punkte bezog: Ergänzung durch Operationalisierung, Verbesserung der Verständlichkeit, stärkere Berücksichtigung der Multiperspektivität einer Evaluation, Beseitigung bzw. Moderation von Widersprüchen sowie deutlichere Abgrenzung zu verwandten Begriffen und Konzepten.

Mit der Revision, so eine grundsätzliche Überlegung, könnte eine Brücke zur Anwendung geschlagen werden: „Von der Orientierung zur Handlungsanleitung“. Der Mitgliederversammlung 2012 wurde ein Vorschlag für die Durchführung des Revisionsprozesses vorgelegt. Sie beauftragte den Vorsitzenden einstimmig, eine Arbeitsgruppe für die Durchführung dieses Prozesses und die Entwicklung eines Revisionsvorschlags zu bilden.

In der Folge begann ein intensiver Diskussionsprozess. Die Mitgliedschaft wurde anlässlich der Mitgliederversammlungen über den Verlauf informiert. Im Kontext der Jahrestagung 2013 fand eine Session statt, bei der die notwendigen weiteren Prozesse und Inhalte diskutiert wurden. Im Winter 2013 fand in Bern (CH) ein von EvalPartners geförderter Workshop mit der Schweizerischen Evaluationsgesellschaft SEVAL statt, die parallel an der Revision ihrer Standards arbeitete. Hier hat sich im gemeinsamen Austausch und nach Beiträgen eingeladenen Experten der Ansatz gefestigt, die Standards lediglich inkrementell, also „konservativ“ zu revidieren.

Zur Jahrestagung 2015 legte die Arbeitsgruppe den Mitgliedern der DeGEval einen Entwurf der Revision vor. Die Sprecherinnen und Sprecher der Arbeitskreise wurden gebeten, eingehende Kommentierungen und Vorschläge in eine endgültige Konzeption einzuspeisen und mit der Arbeitsgruppe und dem Vorstand der DeGEval abzustimmen. Auf der Jahrestagung fand zudem eine Session statt, in der es insbesondere um die Sammlung von Ansätzen gehen sollte, wie die Praxisrelevanz, Verbreitung und Anwendung der Standards gefördert werden kann.

Die daraus resultierende erste Entwurfsfassung der revidierten Standards wurde im November 2015 der gesamten Mitgliedschaft zur Kommentierung vorgelegt. Die wesentlichen Neuerungen – und Konstanten – der revidierten Fassung sind:

- Beibehaltung der bewährten Struktur mit den vier Attributen guter Evaluationen (Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit). Die in der dritten Auflage der JCSEE-Standards eingeführte neue Standardgruppe „Meta-Evaluation“ wurde als strukturell wenig überzeugend beurteilt und nicht übernommen. Auch wurde bewusst keine Neugliederung nach typischen Prozessphasen der Evaluation vorgenommen, da eine solche nicht einheitlich für alle Evaluationsfälle anwendbar wäre und die Ableitung von Kriterien für Meta-Evaluationen auf Basis der Standards erschwert.
- Moderate Anpassungen im Wortlaut einzelner Einzelstandards. Insbesondere wurde auf eine stärkere Passung zwischen der Kurzbezeichnung eines Einzelstandards und dessen Wortlaut geachtet.
- Berücksichtigung neuerer theoretischer, methodischer und praxisrelevanter Entwicklungen. In den konzeptionellen Grundlagen (Kapitel 4) sowie an geeigneten Stellen der Erläuterungen wurden u.a. folgende inhaltliche Perspektiven gestärkt bzw. ergänzt: Grundlagen und Logik des Bewertens (→Kriterien, →Zielwerte), Datenschutz, Programmtheorien und Wirkmodelle sowie Implementationstreue.

- Überarbeitung aller Erläuterungen (Kapitel 5). Hier wurde eine klare Trennung von Begründung und Umsetzungshinweisen vorgenommen bzw. dort ergänzt, wo sie fehlten. Es wird nun also jeweils eine explizite inhaltliche Begründung für die Notwendigkeit eines Standards gegeben. Zusätzlich werden in aller Kürze zentrale Hinweise zur Umsetzung des Standards in der Evaluationspraxis gegeben.
- Glossar der Standards für Evaluation. Die wesentlichen Begriffe, die zum Verständnis der Standards erforderlich sind, werden in einem Glossar im Anhang definiert und im Text einheitlich mit einem Verweispeil (→) versehen. Bei der Erstellung des Glossars wurde v.a. auf interne Konsistenz geachtet. Daneben sollte in möglichst großem Maße die heterogene Begriffsverwendung in möglichst vielen Anwendungsfeldern der Evaluation Berücksichtigung finden.

Zu der solchermaßen revidierten Entwurfsfassung gingen bis Ende der Kommentierungsfrist im Februar 2016 insgesamt 185 Rückmeldungen zu verschiedenen Textstellen, zum Glossar und zu übergreifenden Aspekten ein. Diese wurden von der Arbeitsgruppe umfassend diskutiert, geprüft und inhaltlich kommentiert, um die Entscheidung über die jeweilige Textstelle zu begründen. Es ergab sich eine Vielzahl von textlichen Änderungen und Ergänzungen, die nicht nur zu einer partizipativen Absicherung der vorliegenden Standards, sondern auch zur Qualität des fertigen Ergebnisses substantiell beigetragen haben.

Die auf dieser Basis überarbeitete Entwurfsfassung wurde am 21. September 2016 der Mitgliederversammlung der DeGEval e.V. zur Abstimmung vorgelegt und mit großer Mehrheit verabschiedet.

DIE STANDARDS FÜR EVALUATION

3

**Evaluationen sollen vier grundlegende Eigenschaften aufweisen:
Nützlichkeit, Durchführbarkeit, Fairness, Genauigkeit**

NÜTZLICHKEIT

N1 – Identifizierung der Beteiligten und Betroffenen

Die am Evaluationsgegenstand oder an der Evaluation Beteiligten sowie die von Evaluationsgegenstand oder Evaluation Betroffenen sollen vorab identifiziert werden, damit deren Interessen und Informationsbedürfnisse 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 und die Evaluierenden einen klaren Arbeitsauftrag verfolgen können.

N3 – Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin

Wer Evaluationen durchführt, soll fachlich und methodisch kompetent sein, damit für die Evaluation und ihre Ergebnisse ein Höchstmaß an Glaubwürdigkeit und Akzeptanz erreicht wird.

N4 – Auswahl und Umfang der Informationen

Auswahl und Umfang der erfassten Informationen sollen die adäquate Beantwortung der zu untersuchenden Fragestellungen zum Evaluationsgegenstand ermöglichen und die Informationsbedürfnisse der Auftraggebenden und weiterer Beteiligter und Betroffener berücksichtigen.

N5 – Transparenz von Werthaltungen

Werthaltungen der Beteiligten und Betroffenen, die sich in deren Perspektiven und Annahmen manifestieren und einen Einfluss haben auf die Evaluation und Interpretation ihrer Ergebnisse, sollten transparent dokumentiert werden, um Evaluationsergebnisse besser einordnen zu können.

N6 – Vollständigkeit und Klarheit der Berichterstattung

Evaluationsberichte sollen alle wesentlichen Informationen zur Verfügung stellen und für ihre Adressatinnen und Adressaten verständlich und nachvollziehbar sein.

N7 – Rechtzeitigkeit der Evaluation

Evaluationsvorhaben sollen so rechtzeitig begonnen und abgeschlossen werden, dass ihre Ergebnisse in anstehende Entscheidungs-, Verbesserungs- oder sonstige Nutzungsprozesse einfließen können.

N8 – Nutzung und Nutzen der Evaluation

Planung, Durchführung und Berichterstattung einer Evaluation sollen die Beteiligten und Betroffenen dazu ermuntern, die Evaluation mitzutragen und ihre Ergebnisse zu nutzen.

DURCHFÜHRBARKEIT

D1 – Angemessene Verfahren

Evaluationsverfahren, einschließlich der Verfahren zur Beschaffung notwendiger Informationen, sollen so gewählt werden, dass einerseits die Evaluation professionell und den Erfordernissen entsprechend umgesetzt wird und andererseits der Aufwand für die Beteiligten und Betroffenen in einem adäquaten Verhältnis zum intendierten Nutzen der Evaluation gehalten wird.

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

F1 – Formale Vereinbarungen

Die Rechte und Pflichten der an einer Evaluation beteiligten Parteien (was, wie, von wem, wann getan werden soll und darf) sollen schriftlich festgehalten werden.

F2 – Schutz individueller Rechte

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

F3 – Umfassende und faire Prüfung

Evaluationen sollen die Stärken und die Schwächen des Evaluationsgegenstandes möglichst fair und umfassend prüfen und darstellen.

F4 – Unparteiische Durchführung und Berichterstattung

Die Evaluation soll unterschiedliche Sichtweisen von Beteiligten und Betroffenen auf Gegenstand und Ergebnisse der Evaluation beachten. Der gesamte Evaluationsprozess sowie die Evaluationsberichte sollen die unparteiische Position der Evaluierenden erkennen lassen.

F5 – Offenlegung von Ergebnissen und Berichten

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

GENAUIGKEIT

G1 – Beschreibung des Evaluationsgegenstandes

Sowohl das Konzept des Evaluationsgegenstands als auch seine Umsetzung sollen genau und umfassend beschrieben und dokumentiert werden.

G2 – Kontextanalyse

Der Kontext des Evaluationsgegenstandes soll ausreichend umfassend und detailliert analysiert sowie bei der Interpretation von Ergebnissen berücksichtigt werden.

G3 – Beschreibung von Zwecken und Vorgehen

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

G5 – Valide und reliable Informationen

Erhebungsverfahren und Datenquellen sollen so gewählt 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 der empirischen Forschung orientieren.

G6 – Systematische Fehlerprüfung

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

G7 – Angemessene 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 beantwortet werden können.

G8 – Begründete Bewertungen und Schlussfolgerungen

Die in einer Evaluation getroffenen wertenden Aussagen sollen auf expliziten Kriterien und Zielwerten basieren. Schlussfolgerungen sollen ausdrücklich und auf Grundlage der erhobenen und analysierten Daten begründet werden, damit sie nachvollzogen und beurteilt werden können.

G9 – Meta-Evaluation

Meta-Evaluationen evaluieren Evaluationen. Um dies zu ermöglichen, sollen Evaluationen in geeigneter Form dokumentiert, archiviert und soweit wie möglich zugänglich gemacht werden.

KONZEPTIONELLE UND BEGRIFFLICHE GRUNDLAGEN

4

Die hier in der ersten Revisionsfassung vorgelegten „Standards für Evaluation“ (kurz: „Standards“) formulieren übergreifende Qualitätsansprüche, an denen sich professionelle Evaluation in der Praxis orientieren und messen lassen sollte.

Gemäß den Standards zeichnen sich gute →Evaluations durch **Nützlichkeit, Durchführbarkeit, Fairness** und **Genauigkeit** aus, wobei diese vier Hauptkriterien durch insgesamt 25 Einzelstandards präzisiert werden. Für die Einhaltung der Standards müssen alle an einer Evaluation →beteiligten Personen bzw. Gruppen an jeweils passender Stelle ihren Beitrag leisten, wobei besonders →Evaluierende und →Auftraggebende in der Pflicht sind.

Dieses Kapitel erläutert im Folgenden genauer die Zielsetzungen der Standards, nimmt grundlegende Definitionen und begriffliche Klärungen vor und gibt übergreifende Hinweise zur Anwendung und Reichweite der Standards.

ZIELSETZUNG DER STANDARDS

Die Gesellschaft für Evaluation (DeGEval) verfolgt mit den Standards für Evaluation das übergreifende Ziel der →Qualitätssicherung und -entwicklung von Evaluationen. Die Standards richten sich damit an alle →Beteiligte und Betroffene einer Evaluation, die an verschiedenen Stellen Einfluss auf Qualität von Evaluationen haben. Insbesondere richten sie sich an →Evaluierende und →Auftraggebende. Um zum Ziel der Qualitätssicherung und -entwicklung beizutragen, sollen die Standards verschiedene Funktionen erfüllen.

- Die Standards sollen als **fachlicher Bezugspunkt und Dialoginstrument** für einen Austausch über die Qualität professioneller Evaluation innerhalb der Fachgemeinschaft dienen.
- Die Standards sollen nach Außen **Transparenz über Evaluation** als professionelle Praxis schaffen und – nicht zuletzt in möglichen Konfliktfällen – die Kommunikation von →Evaluierenden mit anderen →Beteiligten und Betroffenen einer Evaluation unterstützen.
- Die Standards sollen **Orientierung bei der Initiierung, Planung und der Durchführung von Evaluationen** geben. Dies bezieht sich auf die Ausgestaltung der konkreten Arbeitsschritte einer Evaluation von der Entscheidung über eine Evaluation bis hin zur Berichterstattung und Nutzungsförderung.
- Die Standards sollen in der **Aus- und Weiterbildung in Evaluation** zum Einsatz kommen. Sie sind als Grundlage für die Festlegung von Kompetenzziele und Inhalten entsprechender Angebote gleichermaßen relevant.

- Die Standards sollen schließlich als **Kriterien für die →Meta-Evaluation**, also die Evaluation von Evaluationen, dienen, indem sie definieren, welche Merkmale gute Evaluationen aufweisen sollten.
- Die Evaluationsstandards bilden schließlich einen Rahmen, anhand dessen die **Entwicklungsstufe der Evaluationspraxis** in einem bestimmten (z.B. regionalen, nationalen oder politikfeldspezifischen) Einsatzbereich von Evaluation beurteilt werden kann.

Insgesamt stellen die Standards damit ein wesentliches Element und eine zentrale Voraussetzung für die Professionalisierung der Evaluation dar. Sie verstehen sich aber nicht als Schritt zu möglichen Akkreditierungen oder Zertifizierungen im Evaluationsfeld. Auch sind die Standards nicht für die Bewertung von Personen (Personalevaluation) gedacht, wie sie z.B. im Rahmen von individuellen Leistungsbemessungen oder in der Personalrekrutierung durchgeführt werden.

DEFINITION UND BEGRIFFSKLÄRUNGEN

Evaluation ist die systematische Untersuchung der →Güte oder des →Nutzens eines →Evaluationsgegenstands. Evaluation definiert sich demnach

- durch ein nachvollziehbares **systematisches Vorgehen** auf Basis von empirisch gewonnenen Daten, womit sie im Gegensatz zu alltäglichen Bewertungsvorgängen steht,
- durch eine transparente, kriteriengeleitete **Bewertung**, die vor dem Hintergrund eines bestimmten Verwendungskontexts (Untersuchung des Nutzens) oder übergreifend vorgenommen wird (Untersuchung der Güte), womit sie im Gegensatz zu reinen Forschungsstudien steht, sowie
- durch ihre Anwendbarkeit auf **unterschiedliche Gegenstände**. Zu den wichtigsten Evaluationsgegenständen gehören Projekte, Maßnahmen und andere Interventionen (→Programme), Organisationen, Produkte sowie Evaluationen selbst (→Meta-Evaluation).

Als Evaluation wird sowohl der **Prozess der Evaluation** selbst als auch sein **Produkt** im engeren Sinne, z.B. als Kurzform für „Evaluationsbericht“, bezeichnet, wobei sich die Standards explizit auf den gesamten Prozess der Evaluation beziehen nicht nur auf seine Produkte.

Leistungen und Produkte, die als „Evaluation“ gekennzeichnet werden, sollten sich selbstverständlich den in den vorliegenden Standards formulierten Qualitätsansprüchen verpflichtet fühlen. Dennoch impliziert der Begriff „Evaluation“ nach dem oben formulierten Verständnis **kein Qualitätsurteil** in dem Sinne, dass nur „gute“ Evaluationen, die bestimmte Qualitätskriterien erfüllen, so genannt

werden könnten. Es gibt gute und weniger gute Evaluationen. Die Standards dienen dabei zur Unterscheidung und ggf. Orientierung im Sinne einer Qualitätsverbesserung.

Zentral für einen transparenten Bewertungsprozess in Evaluationen ist die Festlegung oder Vereinbarung von →**Kriterien und** →**Zielwerten**, an denen sich ein Werturteil orientiert:

- Kriterien sind jene Merkmale eines Evaluationsgegenstands, in denen sich seine Güte oder sein Nutzen zeigen. Da Auswahl und Priorisierung von Kriterien grundlegende Werte und Werthaltungen verschiedener Akteurinnen und Akteure spiegeln (z.B. „Worin zeigt sich der Erfolg einer Maßnahme?“), ist eine transparente Dokumentation und ggf. Begründung essentielle Voraussetzung für jede systematische Bewertung.
- Zielwerte (Vergleichsstandards, Schwellenwerte, Maßstäbe) geben vor, welche Ausprägung ein Kriterium aufweisen sollte, um ein bestimmtes Werturteil (z.B. „gut“, „ausreichend“, „verbesserungswürdig“) zu rechtfertigen.

Evaluation erfolgt immer mit einer **Verwendungsabsicht**, auf die sie in Planung und Durchführung ausgerichtet werden sollte. Zur Abgrenzung gegenüber der Ebene des evaluierten →Gegenstands wird hier von →Evaluationszwecken gesprochen und der Zielbegriff der Gegenstandsebene vorbehalten (→Ziele des Evaluationsgegenstands vs. Zwecke der Evaluation). Häufige Evaluationszwecke sind:

- die gezielte Verbesserung des evaluierten Gegenstands,
- das Ermöglichen von Entscheidungen über den Gegenstand,
- die Erweiterung von Erkenntnissen über den Gegenstand,
- die Förderung von Lernen und Dialog über den Gegenstand sowie
- die Rechenschaftslegung über den Gegenstand.

Je nach Evaluationszweck kann eine Evaluation →formativ oder →summativ ausgerichtet sein. Ausgehend von den Evaluationszwecken orientieren sich Evaluationen an konkreten **Evaluationsfragestellungen**, die den Untersuchungsfokus genauer abgrenzen.

Evaluationen kommen in unterschiedlichen **Bereichen und Politikfeldern** zur Anwendung. Zu diesen gehören u.a. Bildung, Entwicklungszusammenarbeit, Forschung, Technologie und Innovation, Gesundheitswesen, Kultur, soziale Arbeit, Stadt- und Regionalentwicklung, Umwelt, Verwaltung und Wirtschaft. Einen genaueren Eindruck geben die Arbeitskreise der DeGEval (www.degeval.de) mit ihren verschiedenen thematischen Schwerpunkten. Die Standards haben den Anspruch, politikfeldübergreifend zu gelten und zur Anwendung zu kommen.

Der finanzielle, zeitliche, personelle und räumliche **Umfang von Evaluationen** kann sehr unterschiedlich sein. Die Standards gelten gleichermaßen für größere wie für kleinere Vorhaben. Besonders zu beachten ist bei kleinen wie bei großen Evaluationsprojekten die Verhältnismäßigkeit der intendierten Evaluationszwecke, der investierten Mittel und des gewählten Vorgehens.

Unterschiedliche **Personen, Personengruppen und Institutionen** (→Beteiligte und Betroffene) sind in der Regel an einer Evaluation beteiligt und von ihr betroffen:

- **Ebene der Evaluation:** →Auftraggebende, →Evaluierende, →Nutzerinnen und Nutzer, →Adressatinnen und Adressaten
- **Ebene des Evaluationsgegenstands:** →Zielgruppe, →Programmverantwortliche (Programmdesign, Programmumsetzung, Programmmanagement), →Datengebende

Die genannten Akteurinnen und Akteure bzw. Gruppen sind in den meisten Evaluationen vertreten. Je nach →Kontext können auch mehrere Rollen zusammen fallen (z.B. fungieren oft →Zielgruppen und Programmumsetzende als Datengebende) und sich die beiden genannten Ebenen überschneiden (wenn etwa in einer →formativen Evaluation →Programmverantwortliche gleichzeitig →Nutzerinnen und Nutzer sind).

Evaluationen können in verschiedenen **Konstellationen** umgesetzt werden (Ort der Evaluation). Bei →internen Evaluationen arbeiten Evaluierende in derselben Organisation, die auch für den Evaluationsgegenstand verantwortlich ist, während diese bei →externen Evaluationen von Außerhalb kommen. In →Selbstevaluationen, bei denen Personen, die einen Gegenstand entwickeln und umsetzen, diesen eigenständig evaluieren, fallen die Rollen Auftraggebende, Evaluierende und Programmdesign bzw. -umsetzung zusammen. Bei einer →Fremdevaluation (auch bei einer internen Fremdevaluation) sind diese Rollen getrennt.

In einer **zeitlichen Dimension** können Evaluationen danach unterschieden werden, aus welcher zeitlichen Perspektive sie einen Gegenstand untersuchen bzw. welche Phase im Lebenszyklus eines →Programms sie fokussieren. →Ex-ante Evaluationen untersuchen Plan oder →Konzept eines Gegenstands vor seiner Realisierung etwa im Hinblick auf Bedarf, Umsetzbarkeit und Erfolgsaussichten. Zwischenevaluationen finden während der Durchführungsphase eines Gegenstandes statt und können sowohl bilanzierend (→summativ) als auch gestaltend konstruktiv (→formativ) sein. Schlussevaluationen erfolgen nach Beendigung eines Gegenstandes und überprüfen vornehmlich summativ die Zielerreichung. →Ex-post Evaluationen betrachten summativ die langfristigen →Wirkungen und die Nachhaltigkeit eines Gegenstandes mehrere Jahre nach Beendigung. →Prozessevaluationen fokussieren die Umsetzung eines Gegenstands. Ergebnisevaluationen sind vor allem an den →Outcomes interessiert.

In der Evaluation gibt es vielfältige **Evaluationsansätze**, die bestimmte Aspekte des Evaluationsgeschehens betonen und in Bezug auf die Planung oder Umsetzung der Evaluation bestimmte Setzungen oder Strategien bevorzugen. Sie unterscheiden sich z.B. darin,

- auf welchem erkenntnistheoretischen Fundament Evaluationen stehen,
- welche Partizipationsvorstellungen mit ihnen verbunden sind,
- inwiefern sie die Nutzung von Evaluationsergebnissen optimieren wollen,
- welche Bezugswissenschaften (dominant) heran gezogen werden,
- inwiefern die Wirklogik (→Programmtheorie, →logisches Modell) zu Grunde gelegt wird,
- welche Zwecksetzungen priorisiert werden,
- welche Bewertungskriterien und -maßstäbe vorgegeben werden,
- wie das Verhältnis von Erkenntnisinteresse und Veränderungsabsicht gewichtet ist und
- welche Dimensionen bzw. Aspekte des Evaluationsgegenstandes vorrangig betrachtet werden.

Die Standards gelten grundsätzlich unabhängig vom konkret angewendeten Evaluationsansatz.

Professionelle Evaluation setzt **Evaluationskompetenzen** voraus. Als zentrale Anforderungsprofile hat die DeGEval (2004) die fünf Kompetenzfelder Theorie und Geschichte der Evaluation, Methodenzuständigkeiten, Organisations- und Feldkenntnisse, Sozial- und Selbst-Kompetenzen sowie Praxis der Evaluation definiert.

ANWENDUNG UND REICHWEITE DER STANDARDS

Die Standards sollen der Sicherung und Entwicklung der Qualität von Evaluationen dienen. Ihre Anwendung obliegt **allen an einer Evaluation** →**Beteiligten**. Insbesondere sind →Auftraggebende und →Evaluierende in der Pflicht, die notwendigen Voraussetzungen guter Evaluation zu schaffen und eine hohe Umsetzungsqualität sicher zu stellen.

Die Standards sind als **Maximalstandards** formuliert in dem Bewusstsein, dass nicht alle Standards in allen Konstellationen vollumfänglich zu realisieren sind. Dies liegt auch an einem natürlichen Konkurrenzverhältnis, in dem sich einzelne Standards gegenüber stehen (z.B. Durchführbarkeit vs.

Genauigkeit). Es ist Aufgabe der Evaluierenden und aller an der Evaluation Beteiligten, eine den jeweiligen Zwecken und Rahmenbedingungen angemessene Lösung zu finden.

Die teilweise oder vollständige **Nichterfüllung von Einzelstandards** sollte immer offen und nachvollziehbar etwa im Rahmen der Berichterstattung dokumentiert und begründet werden. Damit wird nicht zuletzt eine Beurteilung der Evaluationsqualität möglich. Diese fragt primär, ob die im jeweiligen Kontext anwendbaren Standards identifiziert und bei Planung und Durchführung der Evaluation berücksichtigt wurden und inwiefern versucht wurde, ihnen im Rahmen der konkreten Bedingungen möglichst weitgehend gerecht zu werden.

Zur konkreten Anwendung der 25 Einzelstandards wurden in der revidierten Fassung zusätzlich zu einer Begründung jeweils explizite **Umsetzungshinweise** je Standard formuliert (vgl. Kapitel 5). Ausführlichere Umsetzungshinweise finden sich in verschiedenen verwandten bzw. Vorläuferpublikationen, auf denen die DeGEval-Standards beruhen. Da die dritte Auflage der JCSEE-Standards (Yarbrough, Shulha, Hopson & Caruthers, 2011) deutlichere Revisionen enthält, die in den vorliegenden DeGEval-Standards nicht nachvollzogen wurden, empfiehlt sich dazu aufgrund der stärkeren strukturellen Nähe nach wie vor insbesondere die zweite Auflage der JCSEE-Standards (Joint Committee on Standards for Educational Evaluation, 1994).

Anregungen und Stellungnahmen zu den Standards sind stets willkommen.

Sie sind zu richten an die Geschäftsstelle der DeGEval – Gesellschaft für Evaluation, vorzugsweise per E-Mail (info@degeval.de). Die aktuelle Postanschrift finden Sie unter „Kontakt“ auf den Internetseiten der DeGEval (www.degeval.de).

ERLÄUTERUNGEN
ZU DEN STANDARDS
FÜR EVALUATION

5

Die DeGEval – Gesellschaft für Evaluation definiert mit den „Standards für Evaluation“, was professionelle Evaluationen ausmacht. Den Kern der Standards bilden die vier Standardgruppen mit 25 Einzelstandards, die in den folgenden Erläuterungen jeweils um eine Begründung und einige Hinweise zur Umsetzung ergänzt wurden.

NÜTZLICHKEIT	34
N1 – Identifizierung der Beteiligten und Betroffenen	34
N2 – Klärung der Evaluationszwecke	35
N3 – Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin	36
N4 – Auswahl und Umfang der Informationen.....	36
N5 – Transparenz von Werthaltungen	37
N6 – Vollständigkeit und Klarheit der Berichterstattung	37
N7 – Rechtzeitigkeit der Evaluation.....	38
N8 – Nutzung und Nutzen der Evaluation.....	38
DURCHFÜHRBARKEIT	39
D1 – Angemessene Verfahren	39
D2 – Diplomatisches Vorgehen	40
D3 – Effizienz von Evaluation.....	40
FAIRNESS	41
F1 – Formale Vereinbarungen.....	41
F2 – Schutz individueller Rechte	41
F3 – Umfassende und faire Prüfung.....	42
F4 – Unparteiische Durchführung und Berichterstattung.....	42
F5 – Offenlegung von Ergebnissen und Berichten	43

GENAUIGKEIT	44
G1 – Beschreibung des Evaluationsgegenstandes	44
G2 – Kontextanalyse.....	44
G3 – Beschreibung von Zwecken und Vorgehen	45
G4 – Angabe von Informationsquellen	45
G5 – Valide und reliable Informationen	46
G6 – Systematische Fehlerprüfung.....	46
G7 – Angemessene Analyse qualitativer und quantitativer Informationen	47
G8 – Begründete Bewertungen und Schlussfolgerungen.....	48
G9 – Meta-Evaluation.....	48

NÜTZLICHKEIT

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

N1 – Identifizierung der Beteiligten und Betroffenen

Die am Evaluationsgegenstand oder an der Evaluation Beteiligten sowie die von Evaluationsgegenstand oder Evaluation Betroffenen sollen vorab identifiziert werden, damit deren Interessen und Informationsbedürfnisse geklärt und so weit wie möglich bei der Anlage der Evaluation berücksichtigt werden können.

Begründung: Eine Evaluation, welche die →Beteiligten und Betroffenen einbezieht, die sich nach deren Informationsbedürfnissen ausrichtet und deren Fragestellungen aufgreift, hat bessere Aussichten, von diesen beachtet und tatsächlich genutzt zu werden. Daher ist es besonders wichtig, die Beteiligten und Betroffenen bei der Klärung der →Evaluationszwecke und Fragestellungen zu beteiligen und bei der Evaluation entsprechend zu berücksichtigen.

Umsetzungshinweise: Die Akteurinnen und Akteure, die im Umfeld einer Evaluation und ihres →Gegenstandes berücksichtigt werden sollen, können vor Beginn und während des Evaluationsprozesses durch Fragen wie die folgenden erschlossen werden:

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

Die so identifizierten Personen, Gruppen und Organisationen werden als →Beteiligte und Betroffene oder Stakeholder bezeichnet. Es soll dabei auch reflektiert werden, wem durch den zu evaluierenden Gegenstand Nachteile erwachsen können oder wer z.B. von einer Maßnahme ausgeschlossen wird.

Dabei sollen gegebenenfalls auch weitere an den Evaluationsergebnissen interessierte Gruppen berücksichtigt werden wie z.B. Entscheidungsträgerinnen und Entscheidungsträger, welche ähnliche Projekte planen, die Fachöffentlichkeit oder die allgemeine Öffentlichkeit.

Die identifizierten Beteiligten und Betroffenen oder sie repräsentierende Einzelpersonen sollten idealerweise dazu befragt werden, welche Bedeutung sie der Evaluation beimessen, welche Informationen ihnen nützlich wären und inwiefern sie die Ergebnisse verwenden möchten.

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 und die Evaluierenden einen klaren Arbeitsauftrag verfolgen können.

Begründung: Die Verantwortung und Rolle der →Auftraggebenden und die Arbeit der →Evaluierenden sind auf die klare Festlegung von →Evaluationszwecken angewiesen, da sie eine notwendige Grundlage für eine konsensfähige und erwartungstreue Planung und Durchführung der Evaluation darstellen.

Umsetzungshinweise: Die Zwecke einer Evaluation sollen frühzeitig angesprochen und soweit möglich im Vorlauf einer Evaluation ausgehandelt und schriftlich fixiert werden. Nützlich ist dabei auch die Frage nach der konkreten vorgesehenen Nutzung von Ergebnissen.

Für die Klärung der Zwecksetzungen ist die idealtypische Unterscheidung dreier Hauptzwecke von Evaluationen hilfreich:

- Informationen bereit zu stellen, die geeignet sind, der schrittweisen Gestaltung des →Evaluationsgegenstandes, z.B. bezüglich der Optimierung von →Konzepten oder Prozessen, Orientierung zu geben („→formative Evaluation“).
- 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 („→summative Evaluation“).
- 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 konzeptuelle Nutzen angesprochen. Die gleichzeitige und gleichrangige Verfolgung mehrerer Hauptzwecke kann die Umsetzbarkeit 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 Evaluierende zu bearbeiten.

N3 – Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin

Wer Evaluationen durchführt, soll fachlich und methodisch kompetent sein, damit für die Evaluation und ihre Ergebnisse ein Höchstmaß an Glaubwürdigkeit und Akzeptanz erreicht wird.

Begründung: Die Evaluationskompetenz der →Evaluatorinnen und Evaluatoren beeinflusst entscheidend die Qualität und Glaubwürdigkeit der Evaluation und erhöht damit die Wahrscheinlichkeit ihrer Nutzung.

Umsetzungshinweise: Um von den verschiedenen →Beteiligten und Betroffenen als glaubwürdig beurteilt zu werden, sind vor allem folgende Eigenschaften bedeutsam: fachliche und methodische Kompetenzen, Integrität, Unabhängigkeit sowie soziale und kommunikative Kompetenzen. Ausführlich wurden die für eine professionelle Evaluation erforderlichen Evaluationskompetenzen in Form von Anforderungsprofilen durch die DeGEval (2004) beschrieben. Diese Kompetenzen sollen bei der Auswahl der Evaluatorinnen und Evaluatoren berücksichtigt werden. Dies kann durch die Zusammenstellung eines →Evaluationsteams, in dem sich Kompetenzen der Evaluatorinnen und Evaluatoren ergänzen, sichergestellt werden.

N4 – Auswahl und Umfang der Informationen

Auswahl und Umfang der erfassten Informationen sollen die adäquate Beantwortung der zu untersuchenden Fragestellungen zum Evaluationsgegenstand ermöglichen und die Informationsbedürfnisse der Auftraggebenden und weiterer Beteiligter und Betroffener berücksichtigen.

Begründung: Um nützliche und sachdienliche Antworten auf Evaluationsfragestellungen und ggf. weitergehende Informationsbedürfnisse geben zu können, muss die Evaluation sich auf Informationen stützen, die inhaltlich relevant und im Umfang ausreichend detailliert sind. Dazu sind nach Möglichkeit auch die verschiedenen Gruppen von Beteiligten und Betroffenen und ihre Sichtweisen zu berücksichtigen. Dabei ist das Prinzip der Datensparsamkeit (vgl. D1) zu berücksichtigen, indem keine darüber hinaus gehenden Informationen erfasst werden.

Umsetzungshinweise: Bei der Planung einer Evaluation soll herausgearbeitet 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 entsprechend der Prioritäten zur Beantwortung der Kernfragestellungen und des Informationsbedarfs der wichtigsten →Beteiligten und Betroffenen der Evaluation eingesetzt werden. Gleichzeitig müssen ausreichend Informationen und ggf. Perspektiven einbezogen werden, um begründete Antworten auf die Evaluationsfragestellungen geben zu können.

In bestimmten Evaluationskontexten, z.B. bei noch wenig ausgereiften Evaluationsgegenständen oder in längerfristigen Evaluationsprojekten, können sich Fragestellungen oder Informationsbedarfe im Verlauf der Evaluation ändern oder neu ergeben. Es sollte daher frühzeitig geklärt werden, inwiefern mit solchen Änderungen zu rechnen sein wird und wie offen die Evaluation diesen gegenüber bleiben soll.

N5 – Transparenz von Werthaltungen

Werthaltungen der Beteiligten und Betroffenen, die sich in deren Perspektiven und Annahmen manifestieren und einen Einfluss haben auf die Evaluation und Interpretation ihrer Ergebnisse, sollten transparent dokumentiert werden, um Evaluationsergebnisse besser einordnen zu können.

Begründung: In mehreren Phasen der Evaluation, häufig bereits bei der Festlegung der →Evaluationszwecke und -fragestellungen, spielen die Werthaltungen 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.

Umsetzungshinweise: 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 gesellschaftlich vermittelte Werthaltungen und 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. Bei unzureichender Wertetransparenz besteht die Gefahr, dass Bewertungen des →Evaluationsgegenstandes von →beteiligten Akteuren nicht akzeptiert werden.

N6 – Vollständigkeit und Klarheit der Berichterstattung

Evaluationsberichte sollen alle wesentlichen Informationen zur Verfügung stellen und für ihre Adressatinnen und Adressaten verständlich und nachvollziehbar sein.

Begründung: Die Vollständigkeit und Klarheit relevanter Formen der Berichterstattung ist eine Mindestvoraussetzung für die Nachvollziehbarkeit, Glaubwürdigkeit und somit Nützlichkeit von Evaluationsergebnissen. Zusätzlich ist sie eine wichtige Voraussetzung der →Meta-Evaluation (vgl. G9).

Umsetzungshinweise: Wichtige Elemente der Berichterstattung sind die Dokumentation von Anlass und →Zwecken der Evaluation, des →Gegenstandes und der Fragestellungen, der Projektschritte und des methodischen Herangehens, der verwendeten Informationsquellen, gewonnenen Daten und ihrer Interpretation, der Ergebnisse und Schlussfolgerungen sowie deren Begründung. Auch methodisch bedingte Limitierungen bezüglich der Tragfähigkeit, des Erkenntnisniveaus und der Reichweite der Ergebnisse sind Bestandteile einer vollständigen Berichterstattung.

Berichterstattung kann verschiedene Formen umfassen. Neben schriftlichen Berichten, gehören dazu unterschiedliche Arten der Kommunikation wie Referate, Präsentationen, Workshops, digitale Medien etc. Dabei sollte jeweils auf eine präzise und für die →Adressatinnen und Adressaten verständliche Darstellung geachtet werden.

Es kann sinnvoll sein, den spezifischen Berichts- und Kommunikationsmix bereits zu Beginn der Evaluation zu vereinbaren.

N7 – Rechtzeitigkeit der Evaluation

Evaluationsvorhaben sollen so rechtzeitig begonnen und abgeschlossen werden, dass ihre Ergebnisse in anstehende Entscheidungs-, Verbesserungs- oder sonstige Nutzungsprozesse einfließen können.

Begründung: Eine Evaluation kann ihre →Zwecke nur erfüllen, wenn ihr zeitlicher Ablauf auf anstehende Entscheidungsprozesse und/oder Verbesserungsprozesse abgestimmt wird. Andernfalls droht sie, einen Großteil ihrer möglichen →Wirkungen zu verlieren. Dies erfordert, dass die Evaluation durch die →Auftraggebenden rechtzeitig vorbereitet, ggf. ausgeschrieben und beauftragt wird.

Umsetzungshinweise: 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.

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 Evaluationen ist es sinnvoll, vorläufige Resultate und Zwischenberichte bereits im Verlauf der Untersuchung zurückzumelden, vor allem, wenn die Informationen für die intendierten Nutzerinnen und Nutzer 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 mitzutragen und ihre Ergebnisse zu nutzen.

Begründung: Evaluationen und ihre Ergebnisse werden häufig nicht oder nicht adäquat genutzt. Daher sollten die →Evaluierenden im Rahmen des Möglichen alle Anstrengungen unternehmen, um eine Nutzung der Ergebnisse wahrscheinlich werden zu lassen. Wichtige Voraussetzungen für eine

Nutzung liegen in einer hohen Durchführbarkeit, Fairness und Genauigkeit von Evaluationen sowie den Aspekten, die in den vorausgegangenen sieben Nützlichkeitsstandards ausgeführt sind.

Umsetzungshinweise: Besonders die Umsetzung von Schlussfolgerungen oder Empfehlungen hängt stark davon ab, ob die →Adressatinnen und Adressaten 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 gar gegenseitig ausschließen können. Eine wichtige Voraussetzung, um eine Nutzung und damit auch den Nutzen von Evaluationen zu fördern, ist die angemessene Einbeziehung der verschiedenen Adressatinnen und Adressaten 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. Dies fördert, dass abgesehen vom Ergebnisnutzen bereits im Verlauf einer Evaluation ein Prozessnutzen eintreten kann.

Die Nützlichkeit von Evaluationen kann über eine unmittelbare Nutzung hinaus erheblich gesteigert werden, wenn die Ergebnisse auf ähnliche Gegenstände verallgemeinerbar oder übertragbar sind. Soweit möglich sollte also geprüft und dokumentiert werden, inwiefern und unter welchen Bedingungen Ergebnisse generalisierbar sind.

DURCHFÜHRBARKEIT

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 einerseits die Evaluation professionell und den Erfordernissen entsprechend umgesetzt wird und andererseits der Aufwand für die Beteiligten und Betroffenen in einem adäquaten Verhältnis zum intendierten Nutzen der Evaluation gehalten wird.

Begründung: Eine Evaluation bringt in der Regel einen zusätzlichen Aufwand für am →Evaluationsgegenstand →beteiligte Personen und/oder Organisationen mit sich, etwa wenn diese Datenerhebungen unterstützen oder selbst als →Datengebende agieren. Um die Mitarbeit der →Beteiligten und Betroffenen zu sichern, die Akzeptanz der Evaluation nicht zu gefährden und die für die Evaluation verfügbaren →Ressourcen zu schonen, sollten Belastungen soweit möglich minimiert werden.

Umsetzungshinweise: Verfahren der Evaluation müssen einerseits den Anforderungen wissenschaftlicher Güte entsprechen und sollen andererseits die beteiligten Personen und/oder Organisationen nicht unnötig belasten oder stören. Zur Umsetzung ist es erforderlich, dass der mindestens notwen-

dige und maximal mögliche methodische Aufwand – auch vor dem Hintergrund der verfügbaren →Ressourcen – abgewogen wird. Vor- und Nachteile sowie Angemessenheit und Aussagekraft der gewählten Verfahren sollen durch die →Evaluierenden transparent gemacht und begründet werden.

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.

Begründung: Politische Tragfähigkeit und ein sensibles Vorgehen sind bei Evaluationen im politischen oder organisatorischen Umfeld eine Voraussetzung sowohl für die Durchführung der Evaluation als auch für die spätere Nutzung von Evaluationsergebnissen. Gelingt es, eine kooperative Haltung bei den einzelnen Interessengruppen zu erzeugen, wird sich dies in der Regel positiv auf die Bereitschaft auswirken, sich am Evaluationsprozess zu beteiligen, Informationen bereitzustellen, die Ergebnisse zu akzeptieren und zu nutzen.

Umsetzungshinweise: Für ein diplomatisches Vorgehen sollten die Interessen aller →Beteiligten und Betroffenen ausgewogen berücksichtigt werden, um deren Akzeptanz, Zustimmung und Kooperation zu sichern. Mögliche Versuche einer dieser Gruppen, die Evaluationsaktivitäten einzuschränken oder die Ergebnisse zu verzerren oder zu missbrauchen, sollen so vermieden werden.

Ein diplomatisches Vorgehen findet z.B. seine Grenzen, wenn versucht wird, Evaluationen in Auftrag zu geben, um Verantwortung für schwierige Entscheidungen abzuwälzen oder bereits getroffene Entscheidungen nachträglich zu legitimieren.

D3 – Effizienz von Evaluation

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

Begründung: Um die erforderliche Akzeptanz für Evaluation nicht zu gefährden, sollte immer das Aufwand-Nutzen-Verhältnis reflektiert werden. Idealerweise sollte also der →Nutzen einer Evaluation Kosten und Aufwand übertreffen oder wenigstens ausgleichen.

Umsetzungshinweise: Es ist sowohl zu Beginn als auch bei Abschluss einer Evaluation oft schwierig, genaue Aussagen zu Aufwand und Nutzen eines Evaluationsvorhabens zu machen. Solche Schwierigkeiten dürfen jedoch nicht dazu führen, dass einschlägige Überlegungen unterbleiben. Gerade bei der Entscheidung über die Durchführung einer Evaluation sollen Aufwand und Nutzen abgeschätzt werden. In der Planung einer Evaluation soll nachvollziehbar dargestellt werden, welcher Aufwand voraussichtlich entsteht und welcher Nutzen erwartet wird. Ein Nutzen kann aus direkten und indirekten, intendierten und nicht-intendierten →Wirkungen der Evaluation entstehen.

FAIRNESS

Die Fairnessstandards sollen sicherstellen, dass in einer Evaluation respektvoll und fair mit allen →beteiligten und betroffenen Personen und Gruppen umgegangen wird.

F1 – Formale Vereinbarungen

Die Rechte und Pflichten der an einer Evaluation beteiligten Parteien (was, wie, von wem, wann getan werden soll und darf) sollen schriftlich festgehalten werden.

Begründung: Eine formelle schriftliche Vereinbarung verpflichtet die Parteien, alle Bedingungen dieser Vereinbarung zu erfüllen oder aber diese neu auszuhandeln. Sie vermindert die Wahrscheinlichkeit, dass Missverständnisse unter den Vertragsparteien auftreten bzw. erleichtert ggf. deren Bereinigung. Gleichzeitig ist sie eine zuverlässige Referenz, um Rechte und Pflichten ggf. neu auszuhandeln.

Umsetzungshinweise: Die formalen Vereinbarungen für eine Evaluation sollen auf der Grundlage von gegenseitigem Respekt und Vertrauen getroffen werden und zumindest in folgenden Bereichen möglichst präzise festgelegt sein: Finanzen, Zeit, Methodik, Veröffentlichungsrechte sowie mitwirkende Personen.

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 die →Evaluierenden alle an der Vereinbarung →Beteiligten von Anfang an über die potentielle Notwendigkeit von Anpassungen aufgeklärt haben.

F2 – Schutz individueller Rechte

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

Begründung: Evaluation kann elementare Personenrechte berühren. Vor allem, wenn eine Evaluation auch Schwächen offenlegt, kann es zu massiven Beeinträchtigungen des Interessenschutzes kommen. Der Schutz individueller Rechte steht über dem Interesse an Information.

Umsetzungshinweise: →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. Im besonderen Maße ist hier der Schutz individueller Rechte zu berücksichtigen. Elementar ist dabei auch die Kenntnis und Einhaltung der gesetzlichen, behördlichen und/oder organisationsbezogenen Bestimmungen zum Datenschutz. Dies impliziert insgesamt den Schutz von Daten vor dem unbefugten Zugriff Dritter.

F3 – Umfassende und faire Prüfung

Evaluationen sollen die Stärken und die Schwächen des Evaluationsgegenstandes möglichst fair und umfassend prüfen und darstellen.

Begründung: Fairness impliziert gleichermaßen die Berücksichtigung von Stärken wie Schwächen, selbst wenn diese ungleich verteilt sein sollten. Eine einseitige Fokussierung von Schwächen birgt das Risiko, dass eventuell vorhandene Stärken übersehen werden, obwohl sie bei anschließenden Veränderungen oder Entscheidungen mitberücksichtigt werden sollten. Fairness impliziert auch, die Kontextabhängigkeit von Stärken und Schwächen offen zu legen und die unterschiedlichen Sichtweisen von →Beteiligten und Betroffenen zu berücksichtigen und zu dokumentieren.

Umsetzungshinweise: Die Kenntnis von Stärken und Schwächen eines Gegenstands sind für seine umfassende Bewertung, für seine Verbesserung oder für Entscheidungen über seine Zukunft notwendig und 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.

Bei der Identifizierung von Stärken und Schwächen eines →Evaluationsgegenstandes sollten immer auch nicht intendierte →Wirkungen berücksichtigt werden.

→Auftraggebende oder andere →Beteiligte sollen nicht versuchen, einseitig Einfluss auf die Evaluation und den Bericht zu nehmen. Es ist Aufgabe der →Evaluierenden, eine umfassende, ausgewogene und faire Evaluation sicherzustellen. Hierzu ist auch erforderlich, differierende Sichtweisen, Interpretationen und Schlussfolgerungen zu berücksichtigen und zu dokumentieren.

F4 – Unparteiische Durchführung und Berichterstattung

Die Evaluation soll unterschiedliche Sichtweisen von Beteiligten und Betroffenen auf Gegenstand und Ergebnisse der Evaluation beachten. Der gesamte Evaluationsprozess sowie die Evaluationsberichte sollen die unparteiische Position der Evaluierenden erkennen lassen.

Begründung: Das Umfeld einer Evaluation ist durch vielfältige Sichtweisen geprägt. Die →Beteiligten und Betroffenen haben oft divergierende Ansichten über den →Evaluationsgegenstand und ihre eigene Funktion oder Rolle. Es besteht die Gefahr, dass Evaluationen von einer bestimmten Gruppe vereinnahmt oder instrumentalisiert werden oder dass in Folge eigener Werthaltungen der →Evaluierenden eine Voreingenommenheit besteht oder erwächst.

Umsetzungshinweise: Unparteilichkeit schließt nicht aus, dass eine begründete Position durch die Evaluation eingenommen wird, solange diese aus einem sichtbar unparteiisch gestalteten Evaluationsprozess resultiert und die Begründung transparent und nachvollziehbar dargelegt wird. Die Evaluation soll aber darauf verzichten, lediglich eine spezifische Sichtweise zu übernehmen. Vielmehr soll sie eine faire Behandlung aller relevanten Interessen anstreben. Deshalb ist dafür zu sorgen, dass die →Evaluierenden eine möglichst unabhängige Position einnehmen. So ist ein zu enges Verhältnis zu →Auftraggebenden aber auch beispielsweise zu den für den Evaluationsgegenstand verantwortlichen Personen zu vermeiden. Das Verhältnis der →Evaluatorinnen und Evaluatoren gegenüber den Auftraggebenden der Evaluation und anderen relevanten Gruppen ist dauerhaft zu reflektieren, insbesondere in Evaluationskonstellationen mit geringer Distanz, wie →internen oder →Selbstevaluationen.

F5 – Offenlegung von Ergebnissen und Berichten

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

Begründung: Es ist ein Gebot der Fairness, dass die Evaluationsergebnisse und -berichte möglichst allen →Beteiligten und Betroffenen mindestens aber den unmittelbar an der Evaluation Beteiligten zur Kenntnis gegeben werden, insbesondere denjenigen, die sich aktiv an der Durchführung der Evaluation beteiligt haben.

Umsetzungshinweise: Da der Kreis der Beteiligten und Betroffenen oft sehr groß ist, kann Offenlegung bedeuten, dass Ergebnisse oder Berichte öffentlich zugänglich gemacht werden müssen, auch wenn in manchen Fällen eine vollständige Offenlegung nicht in Betracht kommt. In solchen Fällen ist es die Aufgabe der Akteurinnen und Akteure, gemeinsam eine Lösung zu finden und ggf. eine Beschränkung der Offenlegung zu vereinbaren und zu begründen.

Die Offenlegung der Evaluationsergebnisse, ihre Art und ihr Umfang, soll wenn möglich zu Beginn einer Evaluation vereinbart und dokumentiert 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 und nachvollziehbare Informationen und Ergebnisse zu dem jeweiligen →Evaluationsgegenstand und den Evaluationsfragestellungen hervorbringt und vermittelt. In diesem Zusammenhang kommt wissenschaftlichen Gütekriterien eine besondere Bedeutung zu.

G1 – Beschreibung des Evaluationsgegenstandes

Sowohl das Konzept des Evaluationsgegenstands als auch seine Umsetzung sollen genau und umfassend beschrieben und dokumentiert werden.

Begründung: Eine genaue und vollständige Beschreibung des →Evaluationsgegenstandes hilft nachzuvollziehen, worauf genau sich die Evaluation und ihre Ergebnisse beziehen. So ist ihr Geltungsbereich korrekt einzuordnen. Dies ermöglicht auch den Vergleich mit ähnlichen Evaluationsgegenständen. Die Beschreibung ist relevant, um bei einer Ausweitung oder dem Transfer z.B. einer Maßnahme deren wesentliche Komponenten und Bedingungen identifizieren zu können.

Umsetzungshinweise: Die Beschreibung des Evaluationsgegenstandes soll klar machen, was untersucht wird. Zu einer umfassenden Beschreibung gehören die Erfassung →beteiligter Akteurinnen und Akteure, →Ziele, →Aktivitäten, Strukturen und →Ressourcen sowie eines →Wirkmodells des Gegenstands.

Wenn sich der Evaluationsgegenstand in unterschiedlichen zeitlichen und räumlichen →Kontexten verschieden ausgestaltet, sollte auch dies dokumentiert werden. Auch wenn sich zwischen der ursprünglich vorgesehenen Form des Evaluationsgegenstandes (→Konzept, Plan, Antrag, Projektskizze, Manual, Verordnung etc.) und der tatsächlichen Implementierung Diskrepanzen zeigen (Problem der Implementierungstreue), sollte dies dokumentiert werden. Nichtbeachtung von Variabilität und/oder von Abweichungen vom →Konzept der Maßnahme kann zu fehlerhaften Interpretationen über den Evaluationsgegenstand führen.

G2 – Kontextanalyse

Der Kontext des Evaluationsgegenstandes soll ausreichend umfassend und detailliert analysiert sowie bei der Interpretation von Ergebnissen berücksichtigt werden.

Begründung: 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 in Bezug auf ihre Übertragbarkeit auf andere →Kontexte einschätzen zu können.

Umsetzungshinweise: Den Kontext des Evaluationsgegenstandes bilden die Rahmenbedingungen, welche diesen Evaluationsgegenstand umgeben. Dies sind v.a. politische, ökonomische, soziale, technologische, gesetzliche und ökologische Faktoren. Die Kontextanalyse sollte sich auf solche Faktoren beschränken, die Umsetzung oder →Wirkung des Evaluationsgegenstands maßgeblich beeinflussen können.

G3 – Beschreibung von Zwecken und Vorgehen

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

Begründung: Die genaue und vollständige Dokumentation von →Zwecken und Vorgehen sowie dem Gegenstand der Evaluation (vgl. G1) dient vor allem dazu, die Evaluation und ihre Ergebnisse nachvollziehen und deren Qualität beurteilen zu können. Zusätzlich ist sie relevant für →Meta-Evaluationen und für eine Diskussion von Fragen der Generalisierbarkeit der Evaluationsergebnisse, etwa im Rahmen von Replikationsstudien und zusammenfassenden Studien wie →Meta-Analysen oder →systematischen Reviews.

Umsetzungshinweise: Um den Evaluationsprozess transparent zu gestalten, sind die Zwecke, die mit einer Evaluation angestrebt werden, genau zu dokumentieren und in nachvollziehbarer Weise zu kommunizieren. Gleiches gilt für die Fragestellungen, Zeitplan, Vorgehensweisen und Methoden. Abweichungen sind ggf. schriftlich festzuhalten.

Bei der Beschreibung der Zwecke und der Fragestellungen der Evaluation ist speziell darauf zu achten, dass ggf. abweichende Sichtweisen berücksichtigt werden. 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 deutlich dargelegt werden, falls Abweichungen entstanden sind und welche Gründe dafür verantwortlich sind.

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 können.

Begründung: Die Beschreibung der genutzten Informationsquellen ermöglicht es, sich ein Urteil über die Qualität der aus diesen Quellen hervorgegangenen Informationen zu bilden. Die Glaubwürdigkeit einer Evaluation wird durch eine fehlende oder unzureichende Beschreibung der Informationsquellen in Frage gestellt.

Umsetzungshinweise: Wenn möglich sollten mehrere unterschiedliche Informationsquellen herangezogen werden, um einen Vergleich und damit eine Abwägung der gewonnenen Informationen zu ermöglichen (vgl. G5 und N4). Neben der Beschreibung der Informationsquellen sollen die aus diesen Quellen erarbeiteten Informationen auch eingeschätzt und bewertet werden können.

G5 – Valide und reliable Informationen

Erhebungsverfahren und Datenquellen sollen so gewählt 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 der empirischen Forschung orientieren.

Begründung: Die Einhaltung von Gütekriterien bei der Erfassung des →Evaluationsgegenstandes mittels qualitativer und quantitativer Methoden der empirischen Sozialforschung oder sonstiger Erhebungs- oder Messmethoden reduziert den Einfluss von Fehlerquellen und trägt so zur Glaubwürdigkeit und Akzeptanz der Evaluationsergebnisse bei.

Umsetzungshinweise: Die eingesetzten, qualitativen wie quantitativen Instrumente der Datenerhebung sowie sonstige Erhebungs- und Messverfahren sollen konsistente, reproduzierbare bzw. intersubjektiv nachvollziehbare, zuverlässige und valide Informationen bereitstellen. Falls möglich und angemessen, sollen mehrere Verfahren zur Datengewinnung verwendet werden. Selbstverständlich finden auch im Falle kombinierter Methodennutzung (Mixed Methods oder Triangulation) die Gütekriterien Anwendung.

Es gilt aber immer, die Wahl der empirischen Zugänge zum Evaluationsgegenstand im Hinblick auf die spezifischen →Zwecksetzungen und Fragestellungen der Evaluation abzustimmen. Die wissenschaftlichen Gütekriterien besitzen für Entscheidungen über Auswahl, Entwicklung und Einsatz von Methoden Orientierungsfunktion und müssen bezogen auf die Anforderungen der konkreten Evaluation spezifiziert und ggf. ergänzt werden.

G6 – Systematische Fehlerprüfung

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

Begründung: Bei der Erhebung, Aufbereitung, Auswertung und Interpretation von Informationen sowie bei der Ergebnispräsentation können verschiedene Fehler auftreten. Diese führen zu unangemessenen oder verzerrten Interpretationen und können die gesamte Evaluation diskreditieren.

Umsetzungshinweise: Der Evaluationsprozess soll so gestaltet werden, dass potentielle Fehlerquellen frühzeitig erkannt und Fehler soweit wie möglich vermieden oder korrigiert werden. Im Prozess der Evaluation sollen daher geeignete Reflexions- und Feedbackphasen vorgesehen sein. Mögliche Fehlerquellen und deren Konsequenzen sind im Rahmen der Berichterstattung zur Evaluation in transparenter Weise zu diskutieren. Es ist für Evaluationen so unabdingbar wie für den →Evaluationsgegenstand im Rahmen der Evaluation, die Suche nach möglichen Fehlern als eine positive Aktivität zu betrachten.

Es kann sich um methodische Fehler im engeren Sinn handeln (z.B. unangemessenes Untersuchungsdesign, mangelnde Kontrolle von Störvariablen, Verfälschungen der Untersuchungssituation, unzulängliche Stichprobenziehung, Messfehler, Fehlinterpretation eines Beobachtungsschemas, Ignorieren von Interviewregeln, Missachtung der Anwendungsvoraussetzungen statistischer Verfahren). Fehler können aber auch durch mangelnde Sorgfalt entstehen (z.B. Verwechslung von Erhebungs-Kennnummern, Vertauschen von Mitschriften oder Tonaufnahmen, Fehleingaben von Daten).

G7 – Angemessene 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 beantwortet werden können.

Begründung: Sowohl qualitative als auch quantitative Verfahren können fruchtbare Daten und Informationen erbringen. Bedingung dafür ist, dass diese jeweils systematisch auf Basis angemessener Regeln ausgewertet werden. Je nach Evaluationsfragestellungen kann aber erst in der kombinierten Reflexion der Auswertungsergebnisse die nötige Tiefenschärfe erzeugt werden.

Umsetzungshinweise: Die Auswahl der adäquaten Auswertungsverfahren soll sich nach den Fragestellungen der Evaluation, nach dem Erkenntnisstand bezogen auf den →Evaluationsgegenstand und nach Kontextvariablen des Evaluationsfeldes richten. Voraussetzung dafür ist, dass bei den →Evaluierenden die erforderlichen Kompetenzen in der notwendigen Breite vorhanden sind.

Die herangezogenen Regeln und deren methodische Grundlagen sollen – z.B. durch Literaturhinweise – offen gelegt werden. Die Auswahl sowie der Einsatz der Verfahren sollen transparent und nachvollziehbar erfolgen, so dass Auswahlentscheidungen ebenso wie Ergebnisse kritisierbar bleiben. Auf die Aussagekraft der Methoden, wie auch auf ihre Begrenzungen, ist explizit hinzuweisen.

G8 – Begründete Bewertungen und Schlussfolgerungen

Die in einer Evaluation getroffenen wertenden Aussagen sollen auf expliziten Kriterien und Zielwerten basieren. Schlussfolgerungen sollen ausdrücklich und auf Grundlage der erhobenen und analysierten Daten begründet werden, damit sie nachvollzogen und beurteilt werden können.

Begründung: Die Einhaltung dieses Standards erlaubt es, die Aussagekraft von Bewertungen und Folgerungen einzuschätzen. Zudem wird deren Überzeugungskraft gestärkt.

Umsetzungshinweise: Jede Art von Bewertung basiert auf →Kriterien und →Zielwerten. Kriterien sind dabei jene Merkmale eines Gegenstands, in denen sich seine →Güte äußert (z.B. Kompetenzzuwachs, Reduktion von Straftaten, Akzeptanz einer Maßnahme oder Kosteneinsparungen). Zielwerte geben eine Ausprägung vor, die ein Kriterium aufweisen muss, um eine bestimmte Wertzuweisung wie „gut“ oder „verbesserungswürdig“ zu ermöglichen. Zielwerte benötigen immer einen Vergleichsmaßstab. Bei bewertenden Aussagen in einer Evaluation sollte immer deutlich sein, auf welchen Kriterien und Zielwerten sie beruhen, damit die Zuweisung nachvollziehbar ist.

Ebenso sollten die Schlussfolgerungen einer Evaluation explizit begründet und zusammen mit den zugrunde liegenden Annahmen und den eingesetzten Evaluationsverfahren nachvollziehbar dargestellt werden. Dabei ist auch deutlich auf den Geltungsbereich der Schlussfolgerungen hinzuweisen, wobei die Verlässlichkeit und Aussagekraft der Daten sowie Einflüsse des →Kontexts berücksichtigt werden sollten. Keinesfalls dürfen sich Bewertungen eher auf subjektive Einschätzungen der →Evaluierenden als auf die ermittelten Daten stützen. Wo erforderlich, soll die Berichterstattung eine Diskussion alternativer Interpretationen beinhalten.

G9 – Meta-Evaluation

Meta-Evaluationen evaluieren Evaluationen. Um dies zu ermöglichen, sollen Evaluationen in geeigneter Form dokumentiert, archiviert und soweit wie möglich zugänglich gemacht werden.

Begründung: Professionelle Evaluation zeichnet sich durch ein hohes Qualitätsbewusstsein aus, daher soll auch die Evaluation selbst evaluiert werden können. →Meta-Evaluationen sind wichtig, um die Glaubwürdigkeit einzelner Evaluationen, aber auch die des Tätigkeitsfelds der →Evaluatorinnen und Evaluatoren insgesamt zu erhöhen. Außerdem ermöglichen nur Evaluationen, die den Standards genügen, eine Aggregation und Systematisierung von Evaluationen, um generalisierbares Wissen zu erzeugen.

Umsetzungshinweise: Meta-Evaluationen dienen dem Zweck, die Qualität von Evaluationsprozessen und -ergebnissen zu beurteilen. Im Rahmen einer Meta-Evaluation können die hier vorliegenden Standards eingesetzt werden. Eine systematische Meta-Evaluation kann begleitend oder retrospektiv nach Ende des Evaluationsprozesses durchgeführt werden. Sie kann innerhalb eines Evaluations-

teams (→intern) oder durch Außenstehende (→extern) erfolgen, wobei eine externe Meta-Evaluation die Zugänglichkeit der untersuchten Evaluationen voraussetzt. Eine →Selbstevaluation der Evaluation sollte in aller Regel durch die →Evaluierenden vorgenommen werden.

Eine adäquate Dokumentation und Berichterstattung ermöglichen nicht nur Meta-Evaluationen, sondern auch →Evaluationssynthesen und →Meta-Analysen und dienen damit auch dem wissenschaftlichen Fortschritt und der Wissensakkumulation im jeweiligen Gegenstandsbereich.

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. Sie ist ein Instrument der Qualitätsentwicklung und fördert darüber hinaus die Verbreitung und Akzeptanz von Evaluationen.

FUNKTIONALES INHALTSVERZEICHNIS

6

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	53
2. Definition des Evaluationsproblems	53
3. Planung der Evaluation	54
4. Informationsgewinnung	55
5. Informationsauswertung	56
6. Berichterstattung zur Evaluation	57
7. Budgetierung der Evaluation	58
8. Evaluationsvertrag	59
9. Steuerung der Evaluation	60
10. Personelle Ausstattung der Evaluation	61

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	34
N2 – Klärung der Evaluationszwecke	35
N7 – Rechtzeitigkeit der Evaluation.....	38
N8 – Nutzung und Nutzen der Evaluation.....	38
D2 – Diplomatisches Vorgehen	40
D3 – Effizienz von Evaluation.....	40
F1 – Formale Vereinbarungen.....	41
G1 – Beschreibung des Evaluationsgegenstandes	44
G2 – Kontextanalyse.....	44

2. DEFINITION DES EVALUATIONSPROBLEMS

Wichtigste Standards:

N1 – Identifizierung der Beteiligten und Betroffenen	34
N2 – Klärung der Evaluationszwecke	35
G1 – Beschreibung des Evaluationsgegenstandes	44
G2 – Kontextanalyse.....	44
G3 – Beschreibung von Zwecken und Vorgehen	45

3. PLANUNG DER EVALUATION

Wichtigste Standards:

N1 – Identifizierung der Beteiligten und Betroffenen	34
N2 – Klärung der Evaluationszwecke	35
N3 – Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin	36
N4 – Auswahl und Umfang der Informationen	36
N5 – Transparenz von Werthaltungen	37
N7 – Rechtzeitigkeit der Evaluation	38
N8 – Nutzung und Nutzen der Evaluation	38
D1 – Angemessene Verfahren	39
D2 – Diplomatisches Vorgehen	40
D3 – Effizienz von Evaluation	40
F1 – Formale Vereinbarungen	41
F3 – Umfassende und faire Prüfung	42
F4 – Unparteiische Durchführung und Berichterstattung	42
G1 – Beschreibung des Evaluationsgegenstandes	44
G2 – Kontextanalyse	44
G3 – Beschreibung von Zwecken und Vorgehen	45
G4 – Angabe von Informationsquellen	45
G5 – Valide und reliable Informationen	46
G7 – Angemessene Analyse qualitativer und quantitativer Informationen	47
G8 – Begründete Bewertungen und Schlussfolgerungen	48
G9 – Meta-Evaluation	48

4. INFORMATIONSGEWINNUNG

Wichtigste Standards:

N3 – Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin	36
N4 – Auswahl und Umfang der Informationen.....	36
N5 – Transparenz von Werthaltungen.....	37
D1 – Angemessene Verfahren	39
D2 – Diplomatisches Vorgehen	40
F1 – Formale Vereinbarungen.....	41
F2 – Schutz individueller Rechte	41
F3 – Umfassende und faire Prüfung.....	42
G1 – Beschreibung des Evaluationsgegenstandes	44
G2 – Kontextanalyse.....	44
G3 – Beschreibung von Zwecken und Vorgehen	45
G4 – Angabe von Informationsquellen	45
G5 – Valide und reliable Informationen	46
G6 – Systematische Fehlerprüfung.....	46
G9 – Meta-Evaluation.....	48

5. INFORMATIONSAUSWERTUNG

Wichtigste Standards:

N5 – Transparenz von Werthaltungen	37
F4 – Unparteiische Durchführung und Berichterstattung.....	42
G1 – Beschreibung des Evaluationsgegenstandes	44
G2 – Kontextanalyse.....	44
G6 – Systematische Fehlerprüfung.....	46
G7 – Angemessene Analyse qualitativer und quantitativer Informationen	47
G8 – Begründete Bewertungen und Schlussfolgerungen.....	48
G9 – Meta-Evaluation.....	48

6. BERICHTERSTATTUNG ZUR EVALUATION

Wichtigste Standards:

N1 – Identifizierung der Beteiligten und Betroffenen	34
N4 – Auswahl und Umfang der Informationen.....	36
N5 – Transparenz von Werthaltungen.....	37
N6 – Vollständigkeit und Klarheit der Berichterstattung	37
N7 – Rechtzeitigkeit der Evaluation.....	38
N8 – Nutzung und Nutzen der Evaluation.....	38
F2 – Schutz individueller Rechte	41
F3 – Umfassende und faire Prüfung.....	42
F4 – Unparteiische Durchführung und Berichterstattung.....	42
F5 – Offenlegung von Ergebnissen und Berichten	43
G1 – Beschreibung des Evaluationsgegenstandes	44
G2 – Kontextanalyse.....	44
G3 – Beschreibung von Zwecken und Vorgehen	45
G4 – Angabe von Informationsquellen	45
G8 – Begründete Bewertungen und Schlussfolgerungen.....	48
G9 – Meta-Evaluation.....	48

7. BUDGETIERUNG DER EVALUATION

Wichtigste Standards:

N2 – Klärung der Evaluationszwecke	35
N4 – Auswahl und Umfang der Informationen.....	36
D3 – Effizienz von Evaluation.....	40
F1 – Formale Vereinbarungen.....	41
G1 – Beschreibung des Evaluationsgegenstandes	44
G3 – Beschreibung von Zwecken und Vorgehen	45

8. EVALUATIONSVERTRAG

Wichtigste Standards:

N1 – Identifizierung der Beteiligten und Betroffenen	34
N2 – Klärung der Evaluationszwecke	35
N3 – Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin	36
N4 – Auswahl und Umfang der Informationen.....	36
N7 – Rechtzeitigkeit der Evaluation.....	38
D2 – Diplomatisches Vorgehen	40
D3 – Effizienz von Evaluation.....	40
F1 – Formale Vereinbarungen.....	41
F2 – Schutz individueller Rechte	41
F4 – Unparteiische Durchführung und Berichterstattung.....	42
F5 – Offenlegung von Ergebnissen und Berichten	43
G1 – Beschreibung des Evaluationsgegenstandes	44
G3 – Beschreibung von Zwecken und Vorgehen	45
G9 – Meta-Evaluation.....	48

9. STEUERUNG DER EVALUATION

Wichtigste Standards:

N1 – Identifizierung der Beteiligten und Betroffenen	34
N2 – Klärung der Evaluationszwecke	35
N3 – Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin	36
N7 – Rechtzeitigkeit der Evaluation	38
N8 – Nutzung und Nutzen der Evaluation	38
D2 – Diplomatisches Vorgehen	40
D3 – Effizienz von Evaluation	40
F1 – Formale Vereinbarungen	41
F2 – Schutz individueller Rechte	41
F4 – Unparteiische Durchführung und Berichterstattung	42
G3 – Beschreibung von Zwecken und Vorgehen	45
G6 – Systematische Fehlerprüfung	46
G9 – Meta-Evaluation	48

10. PERSONELLE AUSSTATTUNG DER EVALUATION

Wichtigste Standards:

N3 – Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin	36
N4 – Auswahl und Umfang der Informationen.....	36
D2 – Diplomatisches Vorgehen	40
F4 – Unparteiische Durchführung und Berichterstattung.....	42
G1 – Beschreibung des Evaluationsgegenstandes	44

GLOSSAR DER STANDARDS FÜR EVALUATION



Das vorliegende Glossar hat zwei Funktionen: Es soll erstens in den Standards für Evaluation häufig verwendete Fachbegriffe inhaltlich präzisieren und dadurch missverständlichen Interpretationen vorbeugen. Zweitens soll es die Konsistenz der Begriffsverwendung innerhalb der Standards und ihrer Erläuterungen sicherstellen.

Die im Glossar erfolgenden Festlegungen orientieren sich so weit wie möglich an in der Evaluationsliteratur etablierten Begriffsverständnissen und versuchen, wo diese nicht eindeutig sind, tragfähige Kompromisse einzugehen. Dies schließt nicht aus, dass auch abweichende Begriffsverständnisse anzutreffen sind, insbesondere in unterschiedlichen Politikfeldern oder der Evaluation benachbarten Disziplinen.

Für die Erstellung des Glossars wurden die zum gegenwärtigen Zeitpunkt verfügbaren Standardlehrbücher zur Evaluation systematisch berücksichtigt. Vorausgesetzt, und daher hier nicht enthalten, sind Begriffe der empirischen Sozialforschung.

Begriff	Definition	Alternative Begriffe
Adressatinnen und Adressaten	Institutionen und Personen, die Evaluationsergebnisse zur Kenntnis nehmen sollen. Dies umschließt die intendierten →Nutzerinnen und Nutzer der Evaluation, umfasst zusätzlich aber weitere mögliche Rezipierende wie z.B. die Fachöffentlichkeit oder andere mit dem →Evaluationsgegenstand befasste Personen (z.B. →Beteiligte und Betroffene, die nicht direkt die Evaluationsergebnisse nutzen).	
Aktivitäten	Die im Zuge der →Implementation eines →Evaluationsgegenstands durchgeführten Arbeitsschritte, Tätigkeiten, Leistungen etc. Aktivitäten sind ein Standardelement von →logischen Modellen.	Prozesse
Audit	Überprüfung der Erfüllung festgelegter Anforderungen (Standards, Merkmale, Parameter, Grenzwerte,...) an ein System, Verfahren, Prozess oder Produkt mittels Soll-Ist-Vergleich.	
Auftraggebende	Institution oder Person, die eine Evaluation bei →Evaluierenden direkt oder nach einem Ausschreibungsprozess beauftragt, den Evaluationsauftrag managt und die Evaluationsergebnisse abnimmt	

Begriff	Definition	Alternative Begriffe
Beteiligte und Betroffene	Alle an einer Evaluation und/oder ihrem →Gegenstand beteiligte oder von ihr betroffene Personen, Gruppen und Institutionen. Wird das Begriffspaar differenziert, wirken Beteiligte in einer eher aktiv-gestaltenden Rolle an der Evaluation mit (z.B. →Auftraggebende), Betroffene dagegen in einer eher passiven Rolle (z.B. →Datengebende). Beteiligte und Betroffene des Evaluationsgegenstands können auf verschiedene Weise in die Evaluation eingebunden werden, sind aber auf jeden Fall zumindest indirekt von den Ergebnissen der Evaluation aufgrund ihrer Beziehung zum Evaluationsgegenstand betroffen.	Stakeholder
Bruttowirkungen	Summe aller Veränderungen, die sich im Zielbereich eines →Evaluationsgegenstands im Zeitverlauf ergeben haben. Schließt neben den →Nettowirkungen auch Wirkungen ein, die nicht kausal auf den Evaluationsgegenstand zurückgehen, sondern andere Ursachen haben.	
Controlling	Kontinuierliche Sammlung und Interpretation von relevanten Daten für das Management von Organisationen mit dem Ziel, Planungs- und Steuerungsprozesse möglichst effektiv und effizient zu gestalten. Kann dafür auf Daten und Ergebnisse von Monitoring und Evaluation zurückgreifen. Im Unterschied zu Monitoring und Evaluation bezieht sich Controlling immer auf organisationale Prozesse, ist als Daueraufgabe angelegt, richtet sich ausschließlich an das Management und enthält keine originär wertende Komponente.	
Datengebende	Personen, die für eine Evaluation Informationen bereitstellen oder im Rahmen von empirischen Erhebungen bei der Evaluation befragt, beobachtet oder in sonstiger Weise untersucht werden.	
Effekt	→Wirkung	
Effektivität	Wirksamkeit eines →Gegenstands im Sinne des Verhältnisses von Zielerreichung (genauer: der →Nettowirkungen) zu intendierten →Zielen.	

Begriff	Definition	Alternative Begriffe
Effizienz	Verhältnis der →(Netto-)Wirkungen eines →Programms zu den investierten →Ressourcen. Je größer die (Netto-)Wirkungen und je geringer der Ressourceneinsatz, desto höher ist die Effizienz des Programms.	
Evaluation	Die systematische Untersuchung von →Nutzen und/oder →Güte eines Gegenstands (→Evaluationsgegenstand) auf Basis von empirisch gewonnenen Daten. Impliziert eine Bewertung anhand offengelegter →Kriterien für einen bestimmten →Zweck.	
Evaluationsgegenstand	→Gegenstand einer Evaluation	
Evaluations-synthese	Überblicksartige Zusammenfassung der Befunde aus mehreren Evaluationsstudien zu einem gemeinsamen →Gegenstand.	
Evaluationsteam	→Evaluierende	
Evaluationszweck	→Zweck einer Evaluation	
Evaluierende	Einzelpersonen oder Teams, die eine →Evaluation konzipieren und durchführen.	Evaluator, Evaluatorin
Ex-ante Evaluation	Evaluation, die vor der Durchführung eines →Programms auf Basis von Konzepten, Plänen oder Anträgen vorgenommen wird und Aspekte wie Bedarf, Realisierbarkeit oder Erfolgsaussichten beurteilt.	
Ex-post Evaluation	Evaluation, die ein →Programm retrospektiv nach dessen Beendigung beurteilt, wobei auch Daten einfließen können, die vor oder während der Programmdurchführung erhoben wurden. Der Fokus liegt häufig auf Nachhaltigkeit sowie langfristigen Wirkungen.	
Externe Evaluation	Evaluation, bei der die →Evaluierenden nicht der Organisation angehören, die für den →Evaluationsgegenstand verantwortlich ist. Ist immer eine →Fremdevaluation, aber nicht dazu synonym, da Fremdevaluationen auch intern realisiert werden können (z.B. durch andere Organisationseinheiten im Rahmen einer „in-house-evaluation“).	

Begriff	Definition	Alternative Begriffe
Formative Evaluation	Evaluation, die dem →Zweck der Verbesserung und Steuerung des →Evaluationsgegenstands dienen soll. Richtet sich primär an →Programmverantwortliche und wird in der Regel begleitend zur Maßnahme und oft zyklisch durchgeführt.	
Fremdevaluation	Evaluation, bei der die →Evaluierenden nicht an Entwicklung, Umsetzung oder Management des →Evaluationsgegenstands beteiligt sind. Die Rollen →Evaluierende und →Programmverantwortung sind hier also im Gegensatz zur →Selbstevaluation getrennt. Kann als →interne oder →externe Evaluation realisiert werden.	
Gegenstand (einer Evaluation)	Das, was im Rahmen der →Evaluation untersucht und bewertet wird und worauf sich mögliche Konsequenzen in Folge der Evaluation beziehen (→Zweck einer Evaluation). →Programme sind in vielen Anwendungsfeldern der häufigste Typ von Evaluationsgegenstand.	
Güte (eines Evaluationsgegenstands)	„Innerer“ oder „intrinsischer“ Wert eines Evaluationsgegenstands, unabhängig von jedem Anwendungskontext (→Nutzen).	
Impact	Einwirkungen eines Evaluationsgegenstands außerhalb seiner intendierten Reichweite, also auf Objekte (Systeme, Institutionen, Organisationen, Personen, Produkte etc.), die durch den evaluierten →Gegenstand mitverursacht sind.	
Implementation	Umsetzung des →Konzepts eines →Evaluationsgegenstands in der Praxis. Ergeben sich dabei Abweichungen vom Konzept, liegt unvollständige Implementationstreue vor.	Implementierung
Indikator	Empirisch erfassbares Merkmal, das ein nicht direkt beobachtbares und für den →Evaluationsgegenstand bedeutsames Konstrukt zugänglich für eine qualitative oder quantitative Erhebung macht.	
Input	→Ressourcen	
Interne Evaluation	Evaluation, bei der die →Evaluierenden derselben Organisation angehören, die auch für den →Evaluationsgegenstand verantwortlich ist. Kann als →Selbstevaluation oder →Fremdevaluation („in-house-evaluation“) realisiert werden.	

Begriff	Definition	Alternative Begriffe
Kontext	Faktoren im Umfeld eines →Evaluationsgegenstands, die seine →Implementation und seine →Wirkungen beeinflussen können (z.B. politische, rechtliche, organisationale, soziale und kulturelle Aspekte oder auch tagesaktuelle Ereignisse).	
Konzept	Plan oder Modell eines →Evaluationsgegenstands, wie er idealerweise umgesetzt werden sollte (→Implementation). Je nach →Evaluationsgegenstand kann das Konzept etwa in Form eines Antrags, eines Manuals, eines Leitbilds, einer Spezifikation, eines Gesetzes oder einer Verordnung vorliegen.	Konzeption
Kriterium	Merkmal eines →Evaluationsgegenstands, anhand dessen Ausprägung durch Vergleich mit einem →Zielwert seine →Güte oder sein →Nutzen festgestellt wird, wobei einer Bewertung in der Regel mehrere Kriterien zugrunde liegen. (s. a. →Zielwert)	
Logisches Modell	Ablauflogik eines →Evaluationsgegenstands mit Elementen wie →Ressourcen, →Aktivitäten, →Outcomes etc. (→Wirkmodell)	
Meta-Analyse	Statistische Zusammenfassung der Befunde mehrerer (Evaluations-)Studien zum gleichen →Gegenstand in quantitativer Form (Effektstärke). Akkumuliert die Ergebnisse z.B. zu einer bestimmten Maßnahme über Einzelevaluationen hinweg. (→Evaluationssynthese, →systematisches Review)	
Meta-Evaluation	Evaluation von Evaluationen, die also Evaluationen als →Gegenstand hat, und →Güte oder →Nutzen dieser Evaluationen bestimmt. Mögliche →Kriterien der Meta-Evaluation sind die vorliegenden Standards der Evaluation. Nicht zu verwechseln mit →Meta-Analyse.	
Monitoring	Routinemäßige, regelmäßige und kriteriengeleitete Sammlung (Protokollierung) von Vergleichsdaten mit dem Zweck, rechtzeitig Steuerungsbedarfe zu erkennen. Ist im Unterschied zur Evaluation nicht bewertend und immer längsschnittlich angelegt.	

Begriff	Definition	Alternative Begriffe
Nettowirkungen	Summe aller →Wirkungen, die direkt oder indirekt durch den →Evaluationsgegenstand entstehen. Entspricht der Differenz zum (kontrafaktischen) Zustand, der ohne diese gegeben wäre.	
Nutzen (eines Evaluationsgegenstands)	Nützlichkeit, Tauglichkeit oder Eignung eines →Evaluationsgegenstands in einem konkreten Verwendungskontext (→Güte).	
Nutzerinnen und Nutzer	Institutionen und Personen, die Evaluationsergebnisse in Konsequenzen umsetzen sollen (intendierte Nutzerinnen und Nutzer) bzw. umsetzen.	
Outcome	Auswirkungen des evaluierten →Gegenstands auf Ebene der →Zielgruppe(n) des →Gegenstands.	
Outputs	Die zählbaren Dienstleistungen und Produkte eines zu evaluierenden →Programms, über die →Wirkungen erreicht werden sollen.	
Programm	Zielgerichtete Vorhaben jeder Art (Projekt, Intervention, Modellversuch, Maßnahme, Maßnahmenbündel etc.), die auf Basis eines →Konzepts planvoll →Aktivitäten durchführt, die zu bestimmten →Wirkungen führen sollen.	
Programmtheorie	→Wirkmodell mit differenzierteren Wirkannahmen, die erklären, wie die →Aktivitäten bzw. →Outputs eines →Programms zu den intendierten (und ggf. auch nicht-intendierten) →Outcomes und →Impacts beitragen. Enthält im Gegensatz zur einfachen linearen Ablauflogik eines →logischen Modells spezifischere Wirkbeziehungen (Wirkketten, Wirkungspfade) zwischen konkreten programminternen und -externen Faktoren.	
Programmverantwortliche	Institutionen und Personen, die Konzept, Umsetzung und/oder Steuerung des Evaluationsgegenstands verantworten. Je nach Kontext fallen diese drei Rollen zusammen oder können nach Programmkonzeption, Programmumsetzung und Programmmanagement differenziert werden.	
Prozessevaluation	Evaluation, die sich auf die Umsetzung (→Implementati-on) eines →Evaluationsgegenstands und nicht auf seine →Wirkungen konzentriert und daher in der Regel begleitend zu diesem durchgeführt wird.	

Begriff	Definition	Alternative Begriffe
Qualitätsmanagement	Alle Aktivitäten und Maßnahmen einer Organisation, die im Rahmen einer Qualitätspolitik Ziele und Verantwortlichkeiten festlegen, um Kundenanforderungen zufrieden zu stellen.	
Qualitätssicherung	Allgemein alle Maßnahmen zur Sicherstellung festgelegter Qualitätsanforderungen. Im Kontext →Qualitätsmanagement verstanden als alle geplanten und systematischen Tätigkeiten zur Erfüllung von Qualitätsanforderungen.	
Ressourcen	Finanzielle, personelle, materielle, administrative, organisationale und andere Mittel, die in ein →Programm investiert werden, um dessen →Ziele zu erreichen.	Inputs
Selbstevaluation	Evaluation, bei der Personen, die einen Gegenstand entwickeln und/oder umsetzen, diesen selbstverantwortlich und eigenständig evaluieren. Die Rollen →Evaluierende, →Programmverantwortliche und teils auch →Auftraggebende fallen hier zusammen. Ist immer eine →interne Evaluation, aber nicht synonym, da auch →Fremdevaluationen intern realisiert werden können („in-house-evaluation“).	
Stakeholder	→Beteiligte und Betroffene	
Summative Evaluation	Bilanzierende Evaluation, die dem →Zweck der Entscheidungsfindung über den →Evaluationsgegenstand (z.B. Fortführung, Ausweitung, Einstellung) dienen soll. Impliziert eine retrospektive Gesamtbewertung und richtet sich primär an Entscheidungspersonen.	
Systematisches Review	Zusammenfassung der Befunde aus mehreren Evaluationsstudien zu einem gemeinsamen →Gegenstand auf Basis eines dokumentierten und replizierbaren Verfahrens. Überbegriff zu →Evaluationssynthesen und →Meta-Analysen.	
Wirkmodell	Ablauflogik und Wirkungsannahmen eines →Programms, meist in visueller Form dargestellt. Umschließt als Überbegriff →logische Modelle und →Programmtheorien.	

Begriff	Definition	Alternative Begriffe
Wirkung	Im Allgemeinen eine auf Ursachen zurückzuführende Veränderung. In Evaluationen interessieren zumeist die direkt oder indirekt durch den →Evaluationsgegenstand ausgelösten Veränderungen (→Nettowirkungen), die von anderen Ursachen analytisch getrennt betrachtet werden sollen. Häufig werden dabei kurz-, mittel- und langfristige Einwirkungen auf →Zielgruppen (→Outcomes) oder auf andere Personen, Gruppierungen, Institutionen, Systeme etc. unterschieden (→Impacts).	Effekt
Ziel (des Evaluationsgegenstands)	Intendierte →Wirkung des →Gegenstands; z.B. das, was mit einem Programm erreicht werden soll.	
Zielgruppe (des Evaluationsgegenstands)	Individuen, Gruppen, Institutionen oder sonstige Entitäten, bei denen der →Evaluationsgegenstand intendierte →Wirkungen erzielen soll.	
Zielwert	Vergleichsstandard, Schwellenwert oder Maßstab, der vorgibt, welche Ausprägung ein →Kriterium aufweisen sollte, um ein bestimmtes Werturteil (z.B. „gut“, „hinreichend“, „verbesserungswürdig“) zu rechtfertigen.	
Zweck (der Evaluation)	Was mit der Evaluation und vor allem ihren Ergebnissen erreicht werden soll (z.B. Verbesserung des →Evaluationsgegenstands, Entscheidungen über den Evaluationsgegenstand, Rechenschaftslegung).	

LITERATUR

DeGEval – Gesellschaft für Evaluation. (2004). Empfehlungen für die Aus- und Weiterbildung in der Evaluation. Anforderungsprofile an Evaluatorinnen und Evaluatoren. Mainz: DeGEval – Gesellschaft für Evaluation.

DeGEval – Gesellschaft für Evaluation. (2002). Standards für Evaluation. Mainz: DeGEval – Gesellschaft für Evaluation.

Joint Committee on Standards for Educational Evaluation. (1994). The program evaluation standards. How to assess evaluations of educational programs (2. Aufl.). Thousand Oaks, CA: Sage.

Joint Committee on Standards for Educational Evaluation (Hrsg.) (2000): Handbuch der Evaluationsstandards. Die Standards des Joint Committee on Standards for Educational Evaluation. 2. Auflage. Opladen.

Joint Committee on Standards for Educational Evaluation & Sanders, J. R. (2006). Handbuch der Evaluationsstandards. Die Standards des "Joint Committee on Standards for Educational Evaluation" (3., erw. und aktual. Aufl.). Übers. und für die dt. Ausg. erw. von Wolfgang Beywl und Thomas Widmer. Wiesbaden: VS Verlag für Sozialwissenschaften.

Yarbrough, D. B. (2011). The program evaluation standards. A guide for evaluators and evaluation users (3rd ed.). Thousand Oaks, CA: Sage.



DeGEval – Gesellschaft für Evaluation e.V.
c/o ZQ Johannes Gutenberg-Universität Mainz
Colonel-Kleinmann-Weg 2
55099 Mainz

www.degeval.de