



Evaluationsstandards – Leitprinzipien von Evaluationen

Vanessa van den Bogaert

Zusammenfassung

Woran sind gute Evaluationen zu erkennen? Wie lassen sich praxistaugliche Evaluationsvorhaben auf hohem Niveau realisieren? Zur Beantwortung dieser Fragen werden im gleichnamigen Beitrag die Evaluationsstandards der *DeGEval* (Deutsche Gesellschaft für Evaluation) eingeführt. Dazu werden die 25 Einzelstandards, geordnet nach den vier Qualitätsdimensionen *Nützlichkeit*, *Durchführbarkeit*, *Fairness* und *Genauigkeit* detailliert beschrieben und erklärt.

1 Einleitung

Das Themen- und Forschungsfeld der Evaluation hat sich in den letzten Jahrzehnten zu einem interdisziplinären Arbeitsgebiet entwickelt, mit speziellen Methoden und Standards, Fachgesellschaften und Kongressen, Handbüchern und Zeitschriften (Westermann 2016). Dem Bedarf der Community nach möglichst allgemeingültigen Evaluationsstandards wurde in den USA bereits in den 1970er-Jahren mit dem sog. *Joint Committee on Standards for Educational Evaluation* nachgegangen (Sanders 2013). Zeitlich versetzt wurden auch in den D-A-CH-Ländern – von den jeweils dort ansässigen nationalen Evaluationsgesellschaften

V. van den Bogaert (✉)

Institut für Erziehungswissenschaft, Ruhr-Universität Bochum, Bochum, Deutschland

E-Mail: vanessa.vandenbogaert@ruhr-uni-bochum.de

© Der/die Autor(en) 2023

P. Niemann et al. (Hrsg.), *Evaluationsmethoden der*

Wissenschaftskommunikation, https://doi.org/10.1007/978-3-658-39582-7_4

(*DeGEval* – Deutsche Gesellschaft für Evaluation; *fteval* – Österreichischen Plattform für Forschungs- und Technologiepolitikevaluation; *SEVAL* – Schweizerische Evaluationsgesellschaft) – Evaluationsstandards erarbeitet und verabschiedet. In diesem Beitrag liegt der Fokus daher auf den Evaluationsstandards, welche ganz allgemein als wünschbare Eigenschaften von Evaluationen gelesen werden können (Beywl 2019). Die Standards bieten Unterstützung bei der Klärung von Fragen nach dem Zweck, der Planung, Durchführung, Auswertung sowie Kommunikation während der Informationssammlung und der Berichtslegung von Evaluationen. Die Standards der drei o. g. Fachgesellschaften unterscheiden sich zwar hinsichtlich ihres inhaltlichen Aufbaus, sie geben jedoch alle einen allgemeingültigen Rahmen für professionelle Evaluationen und legen damit einen Grundstein für eine planvolle und allgemein akzeptable Vorgehensweise. Heute werden im deutschsprachigen Raum dazu vier Qualitätsdimensionen von Evaluationen aufgeführt: *Nützlichkeit*, *Durchführbarkeit*, *Fairness* und *Genauigkeit* (DeGEval 2017). Darüber hinaus bieten die Hinweise für Evaluationen in der Wissenschaftskommunikation der *Impact Unit* spezifische Unterstützung in dem Themenbereich dieses Bands. Das *Factsheet* Evaluationsstandards bietet eine konkrete Hilfestellung, indem die Evaluationsstandards der *DeGEval* für die Wissenschaftskommunikation gerahmt und erläutert werden.

2 Standards für Evaluation

Die Standards für Evaluationen können auch eine Richtschnur zur Bewertung von Evaluationsplänen und -berichten sein und sollten bewusst nicht als ein restriktives Regelwerk gelesen werden (Sanders 2013). Es ist wichtig anzuerkennen, dass nicht alle Standards immer vollumfänglich Berücksichtigung finden können. Die Standards sind zudem als Maximalstandards formuliert worden, dabei können die Einzelstandards teilweise sogar in Konkurrenz zueinander stehen (DeGEval 2017). Daher sollten Evaluator:innen und/oder damit betraute Personen immer eine angemessene sowie zweckdienliche Nutzung der Standards sicherstellen. Die im jeweiligen Kontext anwendbaren Standards müssen dazu zunächst identifiziert und bei der Planung und Durchführung der Evaluation berücksichtigt werden. Gleichmaßen sollte eine bewusste Nichterfüllung von (Einzel-)Standards offen und nachvollziehbar dokumentiert werden. Nachfolgend werden die Standards für Evaluationen der *DeGEval* benannt und anschließend kurz erläutert.

2.1 Nützlichkeit

Die Nützlichkeitsstandards beschreiben zunächst die Orientierung einer Evaluation an den Informationsbedürfnissen der Nutzer:innen. Die *DeGEval* hat dazu acht Teilstandards formuliert (DeGEval 2017).

2.1.1 Identifizierung der Beteiligten und Betroffenen

Standard N1

„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“ (DeGEval 2017, S. 34).

Zumeist gibt es eine Vielzahl an unterschiedlichen Stakeholder:innen und damit an Personen und/oder Personengruppen, die in ein Evaluationsvorhaben einbezogen werden sollten oder von ihm betroffen sein könnten (Sanders 2013). Entscheidend ist hier die Ermittlung der Informationsbedürfnisse, um das Evaluationsvorhaben entsprechend passgenau ausrichten zu können. Dabei sollte bereits zu Beginn festgelegt werden, welche Bedeutung den Informationswünschen der Beteiligten und Betroffenen beigemessen werden kann – und wie diese im Rahmen der zeitlichen und finanziellen Ressourcen realisierbar sind.

2.1.2 Klärung der Evaluationszwecke

Standard N2

„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“ (DeGEval 2017, S. 35).

Die Festlegung von Evaluationszwecken ist eine notwendige Grundlage für die Planung und Durchführung von Evaluationsvorhaben. Evaluationen sollen Orientierung geben können. Im Sinne von formativen Evaluationen kann daher der Hauptzweck, bspw. in der Bereitstellung von Informationen zur weiterführenden Gestaltung des Evaluationsgegenstandes liegen (siehe auch Volk in diesem Band). Demgegenüber kann der Hauptzweck bei summativen Evaluationen in der Informationsbeschaffung für grundlegende Entscheidungen

liegen, bspw. um die Ausweitung oder Einstellung eines Programms bzw. einer Maßnahme mit empirischen Informationen zu unterstützen. Evaluationen können aber auch Diskussionen im öffentlichen, politischen oder wissenschaftlichen Raum anstoßen und stellen zu diesem Zweck systematisch gewonnene Erkenntnisse bereit (DeGEval 2017).

2.1.3 Kompetenz und Glaubwürdigkeit des Evaluators/der Evaluatorin

Standard N3

„Wer Evaluationen durchführt, soll fachlich und methodisch kompetent sein, damit für die Evaluation und ihre Ergebnisse ein Höchstmaß an Glaubwürdigkeit und Akzeptanz erreicht wird“ (DeGEval 2017, S. 36).

Sach- und Fachwissen, Ausbildung, technische Kompetenz, Erfahrung, Integrität von Evaluator:innen oder damit betrauten Teams stellen Eigenschaften dar, die eine Glaubwürdigkeitszuschreibung von Dritten beeinflussen. Da selten eine Person all den Anforderungen genügt, ist es immer empfehlenswert als Kollektiv gemeinsam über all diese Kompetenzen zu verfügen. Eine frühzeitige Überprüfung der Glaubwürdigkeit wird den Teams später dabei helfen die Nützlichkeit, Relevanz und Angemessenheit der Befunde bekräftigen zu können. Die *DeGEval* hat darüber hinaus auch sog. Evaluationskompetenzen für Evaluator:innen in Form von Anforderungsprofilen ausgearbeitet (detaillierte Ausführung hierzu s. DeGEval 2004b).

2.1.4 Auswahl und Umfang der Informationen

Standard N4

„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“ (DeGEval 2017, S. 36).

Es ist in der Regel nicht möglich, alle potenziell verfügbaren Informationen zusammenzutragen und zu analysieren. Die gewonnenen Informationen müssen jedoch so umfangreich sein, dass alle wichtigen Dimensionen des zu evaluierenden Gegenstandsbereichs erfasst und dokumentiert werden können. Um auf Unwichtiges bewusst verzichten zu können, müssen daher Abwägungen

getroffen werden. Dazu müssen Fragestellungen herausgefiltert und gewichtet werden (Sanders 2013). Da sich Informationsbedürfnisse auch im Laufe der Zeit verändern können, sollte der Austausch über eine mögliche Aktualisierung regelmäßig gesucht werden (Sanders 2013). Themenbezogene Literatur, frühe Evaluationsberichte, Austausche in Netzwerken und moderierte Diskussionen können diesen Prozess unterstützen (DeGEval 2017).

2.1.5 Transparenz von Werthaltungen

Standard N5

„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“ (DeGEval 2017, S. 37).

Wie bereits im Einleitungskapitel herausgestellt wurde, ist das Bewerten immanenter Bestandteil von Evaluationen. Eine Bewertung beschreibt somit die Einschätzung einer Sache nach ihrer Nützlichkeit und ihrem allgemeinen Wert (Sanders 2013). Die Interpretation der gesammelten Informationen stellt den wichtigsten Punkt im Evaluationsprozess dar (DeGEval 2017). Alle gewonnenen Informationen – quantitativ, qualitativ, prozess- oder produktbezogen, formativ oder summativ – müssen daher anhand zuvor festgelegter, transparenter und nachvollziehbarer Kriterien interpretiert und bewertet werden. Für eine solche angemessene Ergebnisinterpretation muss dargelegt werden, wer die Werturteile vornimmt und welche Verfahren dazu Verwendung finden (Sanders 2013). Evaluator:innen oder damit betraute Personengruppen handeln jedoch nicht per se objektiv und wertfrei – vielmehr müssen die Bewertungskriterien vorab sorgfältig erarbeitet werden.

2.1.6 Vollständigkeit und Klarheit der Berichterstattung

Standard N6

„Evaluationsberichte sollen alle wesentlichen Informationen zur Verfügung stellen und für ihre Adressatinnen und Adressaten verständlich und nachvollziehbar sein“ (DeGEval 2017, S. 37 f.).

Zwischenergebnisse und Schlussberichte sollen allen Nutzer:innen dienlich sein, weshalb der Vermittlung der Evaluationsergebnisse in Form von Berichten besondere Aufmerksamkeit zu schenken ist. Es gilt zu antizipieren, zu welchen Zeitpunkten die spezifischen Informationen am besten genutzt werden können. Daher empfiehlt es sich, alle vorgesehenen Nutzer:innen der Berichte zu identifizieren und passende Ansätze für verschiedene Nutzer:innengruppen zu erarbeiten. Hier wird besonders deutlich, dass das Evaluationsteam zusammen mit den Auftraggeber:innen der Informationspflicht nachkommen sollte. Eine enge Abstimmung und ggf. schriftliche Vereinbarung können möglichen Interessenskonflikten bei der Identifikation von potenziellen Nutzer:innen entgegenwirken. Insbesondere in sog. Selbstevaluationsberichten sollten die wesentlichen Informationen (Beschreibung des Evaluationsgegenstandes, des Kontextes, der Ziele, der Fragestellungen, der Verfahren sowie der Befunde) in verständlicher Weise allen Beteiligten und Betroffenen zugänglich gemacht werden (DeGEval 2004a).

2.1.7 Rechtzeitigkeit der Evaluation

Standard N7

„Evaluationsvorhaben sollen so rechtzeitig begonnen und abgeschlossen werden, dass ihre Ergebnisse in anstehende Entscheidungs-, Verbesserungs- oder sonstige Nutzungsprozesse einfließen können“ (DeGEval 2017, S. 38 f.).

Hier geht es vor allem um Zeitpläne, die die notwendigen Abstimmungsprozesse, Vorbereitungsphasen, bestenfalls sogar Pufferzeiten sowie Zeiten für eine sorgfältige Fehlerprüfung und Auswertung realistisch abbilden. Die *DeGEval* weist darauf hin, dass in vielen Fällen eine beträchtliche Vorlaufzeit eingeplant werden muss, da interne Abstimmungsprozesse (z. B. Stellungnahmen oder Mitzeichnungsverfahren) vor der Veröffentlichung eingehalten werden müssen. Rückmeldungsschleifen können auch so angesetzt werden, dass Nutzer:innen bspw. bereits durch Zwischenberichte handlungsrelevante Informationen beziehen können. (DeGEval 2017).

2.1.8 Nutzung und Nutzen der Evaluation

Standard N8

„Planung, Durchführung und Berichterstattung einer Evaluation sollen die Beteiligten und Betroffenen dazu ermuntern, die Evaluation mitzutragen und ihre Ergebnisse zu nutzen“ (DeGEval 2017, S. 38 f.).

Der mögliche Nutzen von Evaluationen ist nicht immer allen Beteiligten und Betroffenen von Anfang an ersichtlich. Daher empfiehlt es sich, bei der Kommunikation über das Vorhaben auch zu klären, in welcher Weise die Ergebnisse genutzt werden können. Besonders bei der Berichtslegung empfiehlt es sich, genügend Ressourcen einzuplanen, um über den Bericht und die erzielten Ergebnisse in den Austausch treten zu können.

2.2 Durchführbarkeit

Die Durchführbarkeitsstandards sollen den Rahmen für realistische, gut durchdachte sowie kostenbewusste Evaluationsvorhaben geben. Dazu zählen neben vertraglichen Vereinbarungen vor allem die Auswahl der Datenquellen, die Entscheidung für Methoden der Datensammlung, aber auch die Speicherung und Aufbereitung und Analyse der Daten sowie die Ergebnisaufbereitung. Gleichzeitig sollten Alternativen entwickelt werden, die einem ein rasches Ausweichen bei unvorhergesehenen Problemen ermöglichen. Die *DeGEval* hat dazu drei Durchführbarkeitsstandards formuliert, die im Folgenden erläutert werden.

2.2.1 Angemessene Verfahren

Standard D1

„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“ (DeGEval 2017, S. 39 f.).

Evaluationen bedeuten für alle beteiligten Personen – auch außerhalb des Evaluationsteams – einen Mehraufwand. Daher sollten möglichst wohl-durchdachte Abwägungen zwischen den Anforderungen wissenschaftlicher

Güte und minimaler Störung der beteiligten Personen und/oder Organisationen getroffen werden. Die so erarbeiteten Vor- und Nachteile sowie die Angemessenheit und Aussagekraft der gewählten Herangehensweisen müssen dabei nachvollziehbar begründet werden (DeGEval 2017).

2.2.2 Diplomatisches Vorgehen

Standard D2

„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“ (DeGEval 2017, S. 40).

Evaluationen können durchaus politische Tragweite haben, wenn z. B. Ergebnisse genutzt werden, um Entscheidungen zur Umverteilung von Ressourcen herbeizuführen. Verschiedene Standpunkte, aber auch Bedenken einzelner Interessensgruppen sind ernst zu nehmen. Es gilt das Vertrauen der Beteiligten und Betroffenen zu gewinnen und so versichern zu können, dass die Evaluation auf faire Weise erfolgt, d. h. ohne einzelnen Gruppen womöglich Vorteile zu verschaffen.

2.2.3 Effizienz von Evaluation

Standard D3

„Der Aufwand für Evaluation soll in einem angemessenen Verhältnis zum Nutzen der Evaluation stehen“ (DeGEval 2017, S. 40).

Das Aufwand-Nutzen-Verhältnis sollte bereits in der Planung von Evaluationen reflektiert und nachvollziehbar dargestellt werden. Ziel ist es, herauszustellen, welcher Aufwand entsteht und welcher Nutzen erwartet wird. Dabei kann der Nutzen als Wirkung der Evaluation verstanden werden, was direkte und indirekte, intendierte und nicht-intendierte Wirkungen umfasst (Sanders 2013).

2.3 Fairness

Die Fairnessstandards thematisieren den respektvollen und fairen Umgang mit allen beteiligten und betroffenen Personen und Gruppen (DeGEval 2017). Die *DeGEval* hat dazu die vier Standards formuliert.

2.3.1 Formale Vereinbarungen

Standard F1

„Die Rechte und Pflichten der an einer Evaluation beteiligten Parteien (was, wie, von wem, wann getan werden soll und darf) sollen schriftlich festgehalten werden“ (DeGEval 2017, S. 41).

Eine schriftliche Vereinbarung stellt bestenfalls eine gemeinsame Referenz dar und hilft so dabei, dass sich alle Verantwortlichen an die vereinbarten und gemeinsam getragenen Bedingungen halten können. Mit formalen Vereinbarungen kann zumindest versucht werden, mögliche Missverständnisse frühzeitig zu vermeiden oder diese rechtzeitig zu bereinigen sowie Rechte und Pflichten transparent auszuhandeln (DeGEval 2017). Meist handelt es sich um folgende Bereiche: Finanzen, Zeit, Methodik, Veröffentlichungsrechte sowie mitwirkende Personen und Gruppen (DeGEval 2017).

2.3.2 Schutz individueller Rechte

Standard F2

„Evaluationen sollen so geplant und durchgeführt werden, dass Rechte, Sicherheit und Würde der in eine Evaluation einbezogenen Personen geschützt sind“ (DeGEval 2017, S. 41 f.).

Der Schutz individueller Rechte steht über dem Interesse an Informationen (DeGEval 2017). Evaluationsteams sollten daher die gesetzlichen, behördlichen und/oder organisationsbezogenen Bestimmungen zum Datenschutz kennen und einhalten.

2.3.3 Umfassende und faire Prüfung

Standard F3

„Evaluationen sollen die Stärken und die Schwächen des Evaluationsgegenstandes möglichst fair und umfassend prüfen und darstellen“ (DeGEval 2017, S. 42).

Stärken und Schwächen sollten gleichermaßen offengelegt werden. Dazu ist die Identifikation von unterschiedlichen Sichtweisen verschiedener Stakeholder:innen

unerlässlich. Ziel ist es, umfassende und gleichzeitig ausgewogene Evaluationen sicherzustellen (DeGEval 2017).

2.3.4 Unparteiische Durchführung und Berichterstattung

Standard F4

„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“ (DeGEval 2017, S. 42 f.).

Die vielfältigen Sichtweisen können durchaus verschieden sein. Evaluator:innen sollten daher eine möglichst unparteiische Position beziehen. Dies kann beispielsweise durch eine regelmäßige und planvolle Reflexion gelingen (DeGEval 2017).

2.3.5 Offenlegung von Ergebnissen und Berichten

Standard F5

„Evaluationsergebnisse und -berichte sollen allen Beteiligten und Betroffenen soweit wie möglich zugänglich gemacht werden“ (DeGEval 2017, S. 43).

Die Offenlegung von Berichten sollte bereits zu Beginn von Evaluationen sorgsam thematisiert werden. Die Evaluationsergebnisse und -berichte sollten i. d. R. allen Beteiligten und Betroffenen zugänglich gemacht werden, jedoch ist eine vollständige Offenlegung aus unterschiedlichsten Gründen nicht immer möglich (DeGEval 2017). Die *DeGEval* empfiehlt, bereits in der Planungsphase über die Art und den Umfang der Offenlegung der Ergebnisse zu entscheiden. Der Spielraum für eine kontextspezifische Lösung liegt dabei zwischen einer vollumfänglichen Veröffentlichung und einer begründeten Auswahl von (Teil-) Ergebnissen.

2.4 Genauigkeit

Die Genauigkeitsstandards thematisieren vor allem die wissenschaftlichen Gütekriterien und setzen die Gültigkeit und Nachvollziehbarkeit der Ergebnisse und Schlussfolgerungen in den Fokus (DeGEval 2017). Die *DeGEval* hat dazu acht Standards formuliert.

2.4.1 Beschreibung des Evaluationsgegenstandes

Standard G1

„Sowohl das Konzept des Evaluationsgegenstand[es] als auch seine Umsetzung sollen genau und umfassend beschrieben und dokumentiert werden“ (DeGEval 2017, S. 44).

Herausuarbeiten, welchen Geltungsbereich Evaluationen haben, bedeutet eine lückenlose und sorgfältige Beschreibung des Evaluationsgegenstandes. Dabei soll eine solche Beschreibung die beteiligten Akteur:innen, Ziele und Zwecke, Rahmenbedingungen und Strukturen sowie die bereitgestellten Ressourcen umfassen. Gleichzeitig sind auch die notwendigen Änderungen und Anpassungen wichtige Bestandteile einer solchen lückenlosen Beschreibung. Ziel ist, eine fehlerhafte Interpretation zu vermeiden und die wesentlichen Bedingungen für den möglichen Erkenntnistransfer transparent aufzeigen zu können (DeGEval 2017).

2.4.2 Kontextanalyse

Standard G2

„Der Kontext des Evaluationsgegenstandes soll ausreichend umfassend und detailliert analysiert sowie bei der Interpretation von Ergebnissen berücksichtigt werden“ (DeGEval 2017, S. 44 f.).

Um die Ergebnisse von Evaluationen hinsichtlich ihrer Gültigkeit und Übertragbarkeit einschätzen zu können, sind Erkenntnisse über mögliche Bedingungsfaktoren notwendig. Dies betrifft politische, ökonomische, soziale, technologische oder ökologische Rahmenbedingungen, die bei Evaluationen nicht ausgeblendet werden sollten (DeGEval 2017). Um mögliche Faktoren zu identifizieren, die einen maßgeblichen Einfluss auf den Evaluationsgegenstand und/oder die Ergebnisse haben können, sind Kontextanalysen ein wichtiger Bestandteil von Evaluationsvorhaben, denen genügend Aufmerksamkeit geschenkt werden sollte.

2.4.3 Beschreibung von Zwecken und Vorgehen

Standard G3

„Zwecke, Fragestellungen und Vorgehen der Evaluation, einschließlich der angewandten Methoden, sollen so genau dokumentiert und beschrieben werden, dass sie nachvollzogen und beurteilt werden können“ (DeGEval 2017, S. 45).

Eine genaue und vollständige Dokumentation beinhaltet die Beschreibung von Zwecken, Fragestellungen, Zeitplänen, Vorgehensweisen und Methoden sowie mögliche Abweichungen.

2.4.4 Angabe von Informationsquellen

Standard G4

„Die im Rahmen einer Evaluation genutzten Informationsquellen sollen hinreichend genau dokumentiert werden, damit die Verlässlichkeit und Angemessenheit der Informationen eingeschätzt werden können“ (DeGEval 2017, S. 45 f.).

Um die Qualität und Glaubwürdigkeit von Evaluationen einschätzen zu können, müssen die genutzten Informationsquellen lückenlos beschrieben werden (DeGEval 2017).

2.4.5 Valide und reliable Informationen

Standard G5

„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“ (DeGEval 2017, S. 46).

Die Methoden und Verfahren, die bei Evaluationen genutzt werden können, sind immer kontextabhängig. In diesem Band ist eine ganze Fülle von methodischen Herangehensweisen gesammelt worden. Unabhängig von den individuell gewählten methodischen Zugängen steht immer der Anspruch an eine möglichst hohe Validität der Schlussfolgerungen an erster Stelle. Dies betrifft die Güte und die Glaubwürdigkeit der Schlüsse, die aus den Ergebnissen des

Informationsgewinnungsprozesses gezogen werden können (Sanders 2013). Ziel ist es, konsistente, reproduzierbare, intersubjektiv nachvollziehbare, zulässige und valide Informationen bereitzustellen (DeGEval 2017).

2.4.6 Systematische Fehlerprüfung

Standard G6

„Die in einer Evaluation gesammelten, aufbereiteten, analysierten und präsentierten Informationen sollen systematisch auf Fehler geprüft werden“ (DeGEval 2017, S. 46 f.).

Empirisches Arbeiten, d. h. das Sammeln, Verarbeiten, Analysieren und Dokumentieren von Daten ist ein Prozess, der durchaus Fehlermöglichkeiten birgt (Sanders 2013). Mit einer systematischen Informationsprüfung soll sichergestellt werden, dass die Informationen so weit wie möglich fehlerfrei und abgesichert sind (Sanders 2013). Darum empfiehlt es sich, systematisch nach Fehlern zu forschen und das Thema der Qualitätssicherung ernst zu nehmen.

2.4.7 Angemessene Analyse qualitativer und quantitativer Informationen

Standard G7

„Qualitative und quantitative Informationen einer Evaluation sollen nach fachlichen Maßstäben angemessen und systematisch analysiert werden, damit die Fragestellungen der Evaluation beantwortet werden können“ (DeGEval 2017, S. 47 f.).

Die Analyseverfahren sind so auszuwählen, dass sie für die Fragestellungen der Evaluationen sowie für die Art der Daten angemessen sind. Dabei können sich qualitative und quantitative Zugänge gegenseitig ergänzen und die Interpretationen gemeinsam tragen (Sanders 2013).

2.4.8 Begründete Bewertungen und Schlussfolgerungen

Standard G8

„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“ (DeGEval 2017, S. 48).

An Schlussfolgerungen wird der Anspruch gestellt, sowohl die Fragestellungen wahrheitsgemäß zu beantworten, als auch das Vorgehen und die Ergebnisse wahrheitsgetreu widerzuspiegeln (Sanders 2013). Mögliche Limitationen der Evaluationsverfahren und der gewonnenen Datengrundlage sollten gleichermaßen thematisiert werden. Schlussfolgerungen sollten sich daher niemals auf unzureichende und ungenaue Informationen stützen, auch wenn der Wunsch nach der Beantwortung der Fragestellungen noch so groß erscheint (DeGEval 2017).

2.4.9 Meta-Evaluation

Standard G9

„Meta-Evaluationen evaluieren Evaluationen. Um dies zu ermöglichen, sollen Evaluationen in geeigneter Form dokumentiert, archiviert und soweit wie möglich zugänglich gemacht werden“ (DeGEval 2017, S. 48 f.).

Auch wenn Meta-Evaluationen nicht im Fokus dieses Beitrags stehen, sind Systematisierungen von Evaluationen nur möglich, wenn einzelne Evaluationsvorhaben den Standards genügen. Daher sei an dieser Stelle angemerkt, dass eine angemessene Dokumentation und Berichtslegung Meta-Evaluationen, Evaluationssynthesen und gegebenenfalls auch Meta-Analysen möglich macht (DeGEval 2017).

3 Fazit

Standards sind nicht nur Hilfsmittel bei der Konzeption von Evaluationen, sie dienen auch als Grundlage für Bewertungen und sind zentral bei der Qualitätskontrolle. Evaluationsstandards können auch – durch eine angemessene Berücksichtigung in einzelnen Evaluationen – den Wissenstransfer im Themen- und Forschungsfeld der Evaluationen fördern. Es ist grundsätzlich möglich, dass einzelne Aspekte durch Evaluationsteams oder damit betrauten Personen – bspw. bei Selbstevaluationen – angepasst werden müssen oder sich nicht anwenden lassen (DeGEval 2004a). Da die Standards als sog. Maximalstandards formuliert sind, kann es in der Praxis bspw. zu Konflikten zwischen zwei oder mehr dieser Standards kommen, sodass begründete Einschränkungen erfolgen müssen. Jedoch sind Auslassungen und Anpassungen von Einzelstandards immer sorgfältig abzuwägen und schriftlich festzuhalten (DeGEval 2004a). Nützliche, durchführbare,

faire und genaue Evaluationen zu ermöglichen, ist jedoch nicht allein die Aufgabe von Evaluator:innen. „Für die Qualität von Evaluationen sind viele mitverantwortlich“ (Hense 2021, S. 5). Die fachlichen Standards für Evaluationen richten sich gleichermaßen an Auftraggebende, Entscheidungstragende, Projektbeteiligte, Zielgruppenangehörige und je nach Konstellation auch an weitere Akteur:innen (Hense 2021). Evaluationen können zudem Grundlagen für Fort- und Weiterbildungen sein und unterstützen eine nachvollziehbare Argumentation (DeGEval 2004b). Insbesondere im Handlungsfeld der Selbstevaluation hat die *DeGEval* weitere spezifische Konkretisierungen und Interpretationshilfen erarbeitet und empfiehlt die Anwendung einer besonderen Gewichtung einzelner Evaluationsstandards, die insbesondere die (Doppel-)Rolle der Evaluator:innen und die daraus abzuleitenden feldspezifischen Anforderungen betreffen (DeGEval 2004a).

Literatur

- Beywl W (2019) Evaluationsstandards–Orientierungshilfen für Evaluationen in Schule und Unterricht. In: Buhren CG, Klein G, Müller S (Hrsg) Handbuch Evaluation in Schule und Unterricht. Beltz, Weinheim, S 30–44
- DeGEval [Gesellschaft für Evaluation] (2004a) Empfehlungen zur Anwendung von Standards für Evaluation im Handlungsfeld der Selbstevaluation. DeGEval – Gesellschaft für Evaluation, Alfter
- DeGEval [Gesellschaft für Evaluation] (2004b) Empfehlungen für die Aus- und Weiterbildung in der Evaluation. Anforderungsprofile an Evaluatorinnen und Evaluatoren. DeGEval – Gesellschaft für Evaluation, Mainz
- DeGEval [Gesellschaft für Evaluation] (2017) Standards für Evaluation. Erste Revision 2016. DeGEval – Gesellschaft für Evaluation, Mainz
- Hense JU (2021) Nicht jede Evaluation ist eine gute Evaluation. Warum gute Evaluationen fachliche Standards berücksichtigen sollten. FORUM Sexualaufklärung und Familienplanung: Informationsdienst der Bundeszentrale für gesundheitliche Aufklärung (BZgA) 1:3–5
- Sanders JR (2013) Handbuch der Evaluationsstandards: Die Standards des Joint Committee on Standards for Educational Evaluation. VS Verlag, Wiesbaden
- Westermann R (2016) Methoden psychologischer Forschung und Evaluation. Kohlhammer, Stuttgart

Vanessa van den Bogaert ist wissenschaftliche Mitarbeiterin am Lehrstuhl für Lehr-Lernforschung der Ruhr-Universität Bochum. Sie widmet sich in ihren Forschungsschwerpunkten der wissenschaftlichen Begleitforschung von Citizen-Science-Projekten sowie der Grundlagenforschung zur Interessengengese an außerschulischen Lernorten. Sie leitet die Arbeitsgruppe *Science of Citizen Science* in Zusammenarbeit mit *Bürger schaffen Wissen*.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

