

EVALUATIONS-STANDARDS

DER SCHWEIZERISCHEN EVALUATIONSGESELLSCHAFT (SEVAL-STANDARDS)

Thomas Widmer, Charles Landert und Nicole Bachmann

5. Dezember 2000

EINLEITUNG

Zielsetzung der Evaluations-Standards

Die Evaluations-Standards der Schweizerischen Evaluationsgesellschaft, nachfolgend kurz SEVAL-Standards genannt, sollen einen Beitrag zur Professionalisierung der Evaluation in der Schweiz leisten. Die Einhaltung der SEVAL-Standards erhöht die Glaubwürdigkeit, Qualität und Vertrauenswürdigkeit von Evaluationen. Nur das positive Zusammenwirken aller Beteiligten, also von

- Evaluatorinnen und Evaluatoren,
- Auftraggeberinnen und Auftraggeber sowie
- weiteren, an der Evaluation beteiligten Personen

schafft die Basis für Evaluationen von hoher Qualität. Die SEVAL-Standards formulieren deshalb Kriterien, die *alle* an Evaluationen beteiligten Personen in die Pflicht nimmt.

Die SEVAL-Standards definieren die Anforderungen, welche an Evaluationen gestellt werden, nicht aber die Instrumente zur Umsetzung einer Evaluation. Im Idealfall erfüllt eine Evaluation alle in den SEVAL-Standards formulierten Anforderungen. Im konkreten Einzelfall wird es aber nicht immer möglich sein, jeden einzelnen Standard gleichermaßen zu berücksichtigen. Vielmehr gilt es, die SEVAL-Standards an die spezifische Situation anzupassen. Dies kann so geschehen, dass einzelne Standards aus bestimmten Gründen für den konkreten Fall als unwesentlich eingestuft werden, während anderen Standards ein besonders grosses Gewicht beigemessen wird. Anpassungen an die spezifischen Bedingungen einer Evaluation sollten aber wohlüberlegt erfolgen, offen dargelegt und explizit begründet werden. Die Anpassungen sollten von den an einer Evaluation Beteiligten (also Auftraggeberinnen und Auftraggeber, Evaluatorinnen und Evaluatoren, Adressatinnen und Adressaten der Evaluation sowie eventuell weitere Beteiligte) zu Beginn der Evaluation gemeinsam ausgehandelt und festgelegt werden.

Bei der Formulierung der SEVAL-Standards wurde speziell darauf geachtet, dass sie sich (mit Ausnahme der Personalevaluationen) für Evaluationen aller Art eignen. Dies hat zur Konsequenz, dass die Anwenderin oder der Anwender die SEVAL-Standards im Hinblick

auf die vorliegenden Bedürfnisse zu konkretisieren hat. Dies betrifft nicht nur die oben dargelegte Anpassungsleistung, sondern ebenso eine Konkretisierung der in den SEVAL-Standards enthaltenen Aussagen. Dieser Interpretationsvorgang soll unter Berücksichtigung von Transparenz und Nachvollziehbarkeit erfolgen. Dies erlaubt auch Aussenstehenden, die vollzogene Konkretisierung zu verstehen.

Anwendungsbereich der SEVAL-Standards

Die SEVAL-Standards gelten für Evaluationen im Allgemeinen (mit Ausnahme der Personalevaluation), unabhängig vom institutionellen Kontext, dem gewählten Vorgehen (z.B. interne oder externe Evaluationen) und dem spezifischen Themenbereich. Der Ausdruck Evaluationsgegenstand steht in den SEVAL-Standards gleichermassen für Programme, Massnahmen, Projekte, Organisationen, Institutionen, Politiken, Produkte, Materialien oder andere Evaluationsgegenstände, nicht aber für Personen.

Adressatinnen und Adressaten der SEVAL-Standards

Die SEVAL-Standards richten sich an alle Personen, die an Evaluationen beteiligt sind oder diese beeinflussen. Damit wenden sich die SEVAL-Standards namentlich an Evaluatorinnen und Evaluatoren, Auftraggeberinnen und Auftraggeber, aber auch an weitere Personen, welche die Beachtung der SEVAL-Standards unterstützen können (wie etwa jene, die sich in Aus- und Weiterbildung mit Evaluation befassen).

Entstehung der SEVAL-Standards

Die SEVAL-Standards stützen sich auf die 'Program Evaluation Standards' des 'Joint Committee on Standards for Educational Evaluation' ab (Joint Committee 1994 und 2000). Die Arbeitsgruppe 'Evaluationsstandards' der SEVAL hat sich intensiv mit der Thematik Evaluationsstandards auseinandergesetzt. Die Arbeitsgruppe, die Vertreterinnen und Vertreter aus Bund und Kantonen, aus Evaluationspraxis und aus der Wissenschaft umfasst, hat zur Formulierung der SEVAL-Standards einen Ausschuss gebildet. Die Mitglieder dieses Ausschusses haben die vorliegenden Standards unter Vernehmlassung der Mitglieder der Arbeitsgruppe formuliert. Die begleitenden Materialien enthalten Hinweise zur Vorgehensweise sowie eine Liste der an der Ausarbeitung der Standards beteiligten Personen.

Rolle der SEVAL

Die SEVAL versteht die SEVAL-Standards als Instrument zur Qualitätssicherung und –förderung. Sie verbindet damit nicht die Absicht, dass die SEVAL-Standards als Instrument im Rahmen von Akkreditierungs- oder Zertifizierungsverfahren eingesetzt werden könnten. Die SEVAL-Standards eignen sich dafür nicht.

Die SEVAL fördert den Einsatz der SEVAL-Standards in der schweizerischen Evaluationspraxis und trägt zur Weiterentwicklung der Standards bei, indem sie ein Gremium schafft, welches die aktuellen Entwicklungen laufend verfolgt und falls erforderlich der SEVAL-Mitgliederversammlung entsprechende Änderungen der SEVAL-Standards vorschlägt. Dieses Gremium hat nicht die Aufgabe, über die Einhaltung der SEVAL-Standards zu wachen. Ebensowenig verfügt es über irgendwelche Kompetenzen, um bei Verstössen gegen die SEVAL-Standards Sanktionen zu ergreifen.

Überblick über die SEVAL-Standards

Die SEVAL-Standards sind in vier Gruppen eingeteilt, die sich mit den folgenden Themenbereichen befassen: Nützlichkeit, Durchführbarkeit, Korrektheit und Genauigkeit. Diese vier Eigenschaften, werden jeweils einleitend in einem Satz umschrieben.

Nachfolgend werden die einzelnen Standards zur jeweiligen Gruppe aufgeführt. Sie umfassen eine Bezeichnung und eine Beschreibung in einem Satz. Die Standards sind jeweils begleitet von Erläuterungen, die dazu dienen, den Inhalt des einzelnen Standards zu verdeutlichen.

Bei den einzelnen Standards, wie auch bei den vier Gruppen (Nützlichkeit, Durchführbarkeit, Korrektheit und Genauigkeit) wurde auf eine Gewichtung verzichtet. Die SEVAL-Standards machen also keine Aussage dazu, welcher Standard oder welche Standard-Gruppe besonders wichtig sind. Diese Vorgehensweise wurde gewählt, weil die Bedeutung der einzelnen Standards, wie auch der vier Gruppen, von Evaluation zu Evaluation unterschiedlich ist.

Im Anschluss an die Darlegung der Standards findet sich eine Reihe von Begleitmaterialien, nämlich

- eine Übersicht – funktionale Übersicht genannt –, welche die Standards aufführt, die bei spezifischen Tätigkeiten im Evaluationsprozess im Vordergrund stehen und deshalb bei den jeweiligen Aktivitäten spezielle Beachtung verdienen. Diese Übersicht erlaubt es, in spezifischen Situationen die besonders relevanten Standards auf einfache Art und Weise aufzufinden.
- ein Text zur Vorgehensweise bei der Ableitung der SEVAL-Standards von den 'Program Evaluation Standards' des 'Joint Committee on Standards for Educational Evaluation' (1994 und 2000), der auch die Begründungen für Abweichungen vom Originaltext enthält.
- eine Transformationstabelle zu den Kurzbezeichnungen, die es erlaubt, die in den SEVAL-Standards verwendeten Kurzbezeichnungen mit den in der deutschen Fassung der 'Program Evaluation Standards' (Joint Committee 2000) verwendeten in Relation zu setzen.
- eine Liste der Mitglieder der Arbeitsgruppe 'Evaluationsstandards' der SEVAL.
- ein Literaturverzeichnis.

DIE EVALUATIONS-STANDARDS

Nützlichkeit

Die Nützlichkeitsstandards stellen sicher, dass sich eine Evaluation an den Informationsbedürfnissen der vorgesehenen Evaluationsnutzerinnen und -nutzer ausrichtet.

N1 Ermittlung der Beteiligten & Betroffenen

Die an einer Evaluation beteiligten und die von ihr betroffenen Personen werden identifiziert, damit deren Interessen und Bedürfnisse berücksichtigt werden können.

Der Personenkreis, der im Umfeld einer Evaluation berücksichtigt werden sollte, umfasst vor allem folgende Gruppen von Akteuren:

- Wer über die Zukunft des Evaluationsgegenstands zu entscheiden hat (meist die finanzkompetente Instanz)
- Wer verantwortlich ist für die Konzeption / Gestaltung des Evaluationsgegenstands
- Wer an der praktischen Umsetzung des zu untersuchenden Objekts (Projekts, Programms, Gesetzes, Produkts, etc.) beteiligt ist
- Wer durch den Evaluationsgegenstand direkt oder indirekt erreicht wird oder werden sollte (Zielgruppen und deren soziales Umfeld)
- Weitere an den Evaluationsergebnissen interessierte Kreise (z.B. Entscheidungsträger, welche ähnliche Projekte planen, Evaluatorinnen und Evaluatoren, Öffentlichkeit).

Diese Personen, Gruppen und Institutionen werden als Beteiligte & Betroffene ('stakeholders') bezeichnet.

N2 Klärung der Evaluationsziele

Alle an einer Evaluation mitwirkende Personen sind dafür besorgt, dass die Ziele der Evaluationen für alle Beteiligten & Betroffenen geklärt sind.

Klarheit aller Beteiligten & Betroffenen über die Ziele, die mit der Evaluation angestrebt werden, sind von zentraler Bedeutung für den Erfolg einer Evaluation. Die an der Evaluation beteiligten Akteure haben sich gegenseitig über die Evaluationsziele zu verständigen. Dies verhindert, dass übersteigerte Erwartungen (besonders durch die Auftraggeber) an die Evaluation gerichtet werden. Im Gegenzug werden dadurch auch die für die Ausführung der Evaluation verantwortlichen Akteure auf die Evaluationsziele verpflichtet. Die Zielklärung trägt zur Vermeidung von Missverständnissen im Evaluationsprozess bei.

Von zentraler Bedeutung ist die Klärung der Evaluationsziele auch im Hinblick auf eine zielgerichtete Vorgehensweise im Rahmen der Evaluation. Häufig ist die Klärung der Ziele einer Evaluation gleich zu Beginn der Evaluation noch nicht abschliessend möglich, sondern erfordert einen längeren Prozess, der als zentrales Element des Evaluationsprozesses zu betrachten ist.

N3 Glaubwürdigkeit

Wer Evaluationen durchführt, ist sowohl vertrauenswürdig als auch kompetent, damit bei den Evaluationsergebnissen ein Höchstmass an Glaubwürdigkeit und Akzeptanz erreicht wird.

Die Vertrauenswürdigkeit der Evaluatoreninnen und Evaluatoren beeinflusst entscheidend die Durchführbarkeit und Wirksamkeit der Evaluation. Um von den verschiedenen betroffenen

Gruppen als vertrauenswürdig beurteilt zu werden, sind vor allem folgende Eigenschaften bedeutsam: Integrität, Unabhängigkeit sowie soziale und kommunikative Kompetenzen.

N4 Umfang und Auswahl der Informationen

Der Umfang und die Auswahl der erhobenen Informationen ermöglichen die Behandlung sachdienlicher Fragen zum Evaluationsgegenstand und berücksichtigen gleichzeitig die Interessen und Bedürfnisse des Auftraggebers und anderer Beteiligter & Betroffener.

Bei der Planung eines Evaluationsprojektes muss herausgearbeitet werden, welche Informationen zur Beantwortung der Evaluationsfragen unabdingbar sind und welche Informationen zwar interessant und wünschenswert, aber für die Beantwortung der Kernfragen unerheblich sind. Es ist dabei darauf zu achten, dass die vorhandenen Ressourcen zur Datenerhebung entsprechend den Prioritäten zur Beantwortung der Kernfragen und der Bedürfnisse der wichtigsten Zielgruppen der Evaluation eingesetzt werden.

N5 Transparenz der Bewertung

Die Perspektiven und Gedankengänge, auf denen die Interpretation der Ergebnisse beruht, werden so beschrieben, dass die Grundlagen der Werturteile klar ersichtlich sind.

Die Interpretation der gesammelten Informationen und Ergebnisse stellt einen der wichtigsten und kritischsten Punkte in einem Evaluationsprozess dar. In diesem Interpretationsvorgang orientieren sich die Evaluatorinnen und Evaluatoren an theoretischen Modellen und Werthaltungen. Um diesen Interpretationsvorgang überzeugend, nachvollziehbar und beurteilbar zu machen, ist es unerlässlich, diese Grundlagen der Werturteile transparent zu machen.

N6 Vollständigkeit und Klarheit des Berichts

Evaluationsberichte beschreiben den Evaluationsgegenstand einschliesslich seines Kontextes ebenso wie die Ziele, die Fragestellungen, die Verfahren und Befunde der Evaluation, damit die wesentlichen Informationen zur Verfügung stehen und leicht verstanden werden können.

Eine Bedingung, um die Evaluationsergebnisse überzeugend kommunizieren zu können, stellt die Vollständigkeit und Klarheit des Berichts (oder irgendeiner anderen Form der Berichterstattung) dar. Die gewählte Sprache sollte präzise sein (z.B. Definition von wichtigen Begriffen, konsistente Verwendung der Begriffe) und für die vorgesehenen Leserinnen und Leser des Evaluationsberichts verständlich sein. Die Zusammenstellung der wichtigsten Ergebnisse in Tabellenform oder Graphiken sind hilfreich zum raschen Verständnis. Die Gestaltung und Art der Berichterstattung wird idealerweise im Hinblick auf eine optimale Wahrnehmung durch die angestrebte Zielgruppe geplant. Ein umfangreicher Schlussbericht stellt dabei nicht für jeden Adressatenkreis die optimale Form zur Weitergabe der Informationen dar. Je nach Adressatengruppe und Situation erreicht man mehr Beachtung mit Referaten, Workshops oder ähnlichen Kommunikationsformen.

N7 Rechtzeitigkeit der Berichterstattung

Wichtige Zwischenergebnisse und Schlussberichte werden den vorgesehenen Nutzerinnen und Nutzern so zur Kenntnis gebracht, dass diese sie rechtzeitig verwenden können.

Ein Evaluationsprojekt verliert einen Grossteil der angestrebten Wirkung, wenn sein Zeitplan nicht in Übereinstimmung mit anstehenden Entscheidungsprozessen der Adressatinnen und Adressaten gebracht wird. 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. Mitberichtsverfahren, Vernehmlassungen), bevor Entscheidungen getroffen werden können. In vielen Evaluationsprojekten ist es auch sinnvoll, vorläufige Resultate und Zwischenberichte bereits im Verlauf der Untersuchung rückzumelden, vor allem, wenn die Informationen für die Adressaten handlungsrelevant sind. Diese Feedbackschlaufen werden mit Vorteil bereits im Evaluationskonzept berücksichtigt und entsprechende Ressourcen dafür vorgesehen.

N8 Wirkung der Evaluation

Planung, Durchführung und Darstellung einer Evaluation ermuntern die Beteiligten & Betroffenen dazu, dem Evaluationsprozess zu folgen und die Evaluation zu nutzen.

Ob die Evaluationsergebnisse und -empfehlungen umgesetzt werden, hängt stark davon ab, ob die Beteiligten & Betroffenen die Erwartung haben, dass das Evaluationsprojekt ihnen und ihren Vorhaben von Nutzen ist. Eine wichtige Voraussetzung, um diese Erwartung zu fördern und auch einzulösen, ist der Einbezug der verschiedenen Akteure in die Planung und Gestaltung des Evaluationsprojektes. Ebenfalls positiv wirkt es sich aus, wenn im Verlauf der Untersuchung kontinuierlich und transparent Rückmeldungen gemacht werden.

Durchführbarkeit

Die Durchführbarkeitsstandards stellen sicher, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst ausgeführt wird.

D1 Praktikable Verfahren

Evaluationsverfahren sind dazu geeignet, die benötigten Informationen zu beschaffen ohne den Evaluationsgegenstand oder die Evaluation unnötig zu beeinträchtigen.

Bei der Planung und Durchführung der Evaluation darf nicht nur danach gestrebt werden, die aus der Sicht der Wissenschaft geeignetsten Verfahren zu verwenden. Genauso wichtig ist es, darauf zu achten, dass die gewählten Methoden und Verfahren möglichst praktisch anwendbar sind und dass dadurch weder der Evaluationsgegenstand noch die von der Erhebung betroffenen Personen unnötig belastet werden. Im Rahmen von Evaluationsprojekten können die aus wissenschaftlicher Sicht aussagekräftigsten Methoden oft nicht verwendet werden, da sie zu aufwendig, zeitraubend oder im entsprechenden Kontext ethisch nicht akzeptabel sind. Wichtig ist hier, die Vor- und Nachteile sowie die Aussagekraft der gewählten Verfahren bei der Planung des Evaluationsprojektes transparent zu machen und mit den Auftraggeberinnen und Auftraggebern und den Adressatengruppen zu diskutieren.

D2 Politische Tragfähigkeit

Evaluationen werden unter Berücksichtigung der unterschiedlichen Positionen der verschiedenen Interessengruppen geplant und durchgeführt, um deren Kooperation zu erreichen und um mögliche Versuche irgendeiner dieser Gruppen zu vermeiden, die Evaluationsaktivitäten einzuschränken oder die Ergebnisse zu verzerren respektive zu missbrauchen.

Um nicht von negativen Reaktionen auf das Evaluationsprojekt überrascht zu werden, ist es notwendig, möglichst alle Interessengruppen zu identifizieren. Neben dem engeren Kreis der Beteiligten können dies auch Personen sein, die im weiteren Umfeld des Evaluationsgegenstandes angesiedelt sind (z.B. auch Personen, die eine Art Konkurrenzprodukt zum untersuchten Gegenstand anbieten). Können die Bedürfnisse dieser verschiedenen Interessengruppen erkannt und falls möglich berücksichtigt werden, sind negative Reaktionen auf die Evaluation eher absehbar und treten u.U. gar nicht erst auf. Neben den offensichtlichen Interessen dieser Gruppen können auch sogenannte „hidden agendas“ von Bedeutung werden. Bei diesen Überlegungen zu den expliziten und impliziten Bedürfnissen verschiedener Akteure an die Evaluationsergebnisse, darf auch die auftraggebende Instanz nicht ausgeklammert werden.

D3 Kostenwirksamkeit

Evaluationen bringen Informationen mit einem Wert hervor, der die eingesetzten Mittel rechtfertigt.

Eine Evaluation ist dann kostenwirksam, wenn der zu erwartende Nutzen gleich gross oder grösser als ihre Kosten ist. Die *Kosten* bezeichnen den Wert aller benötigten Ressourcen, umfassen also auch den Zeitbedarf für die Begleitung der Evaluation oder die Kosten, die durch andere Institutionen getragen werden. Die Kosten stellen damit die Gesamtheit des (sozialen und monetären) Wertes aller zur Durchführung der Evaluation erforderlichen Ressourcen dar (Vollkosten). Der *Nutzen* bezeichnet die Summe aller Werte, die durch die Evaluation hervorgebracht werden (Wirkungsoptimierungen, Einsparungsmöglichkeiten, Kenntnis der Akzeptanz eines Programmes u.a.).

In einer Evaluation soll das Kosten-Nutzen-Verhältnis möglichst optimal ausfallen. So ist etwa unter mehreren Untersuchungsanlagen mit identischem Nutzen diejenige zu wählen, die mit dem kleinsten Aufwand verbunden ist. Bei kostenmässig ebenbürtigen Designs wäre dementsprechend dasjenige mit dem höchsten erwarteten Nutzen zu favorisieren. Bleiben die Kosten in jedem Fall höher als der zu erwartende Nutzen, ist auf eine Evaluation zu verzichten.

Korrektheit

Die Korrektheitsstandards stellen sicher, dass eine Evaluation rechtlich und ethisch korrekt durchgeführt wird und dem Wohlergehen der Beteiligten & Betroffenen gebührende Aufmerksamkeit widmet.

K1 Formale Vereinbarungen

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

Zu Beginn einer Evaluation ist das Verhältnis zwischen Auftraggeber oder Auftraggeberin und Evaluator oder Evaluatorin zumeist von gegenseitigem Respekt und Vertrauen geprägt. Dies ist ein optimales Umfeld zur Festlegung aller wichtigen Regeln und Pflichten beider Parteien in

einer schriftlichen Vereinbarung (Vertrag, Auftragsbestätigung, o.a.). Die formalen Vereinbarungen sollten zumindest in folgenden Bereichen Festlegungen treffen: Finanzen, Zeit, Personen, Publikation, Inhalt sowie Methodik und Vorgehen. Insbesondere die Rechte und Pflichten der Beteiligten sind möglichst präzise festzuhalten. Werden im Verlauf der Zeit Änderungen nötig, besteht die Möglichkeit, die Bedingungen neu auszuhandeln. Eine formelle, schriftliche Vereinbarung reduziert die Möglichkeit, dass Missverständnisse unter den Vertragsparteien auftreten bzw. erleichtert gegebenenfalls deren Bereinigung.

K2 Schutz individueller Rechte

Evaluationen werden so geplant und durchgeführt, dass die Rechte und das Wohlergehen der Menschen respektiert und geschützt sind.

Menschliche Individuen verfügen über persönliche Rechte, die sich auf Gesetze abstützen oder auf ethischer Praxis, 'common sense' und Anstand beruhen. Bei der Planung und Durchführung einer Evaluation dürfen die Rechte und das Wohlergehen der Menschen nicht tangiert werden. Personen, die in eine Evaluation involviert werden, sind über diesen Umstand zu informieren und die absehbaren Auswirkungen der Evaluation sind gründlich darzulegen. Der Auftraggeber oder die Auftraggeberin sollte es unterlassen, den Evaluators oder die Evaluatorsin zu Entscheiden zu drängen, welche die angeführten Leitlinien tangieren könnten. Führen Evaluationen zu wohlbegründeten Folgerungen, welche das Wohlergehen einzelner Personen gefährden, ist vorsichtig abzuwägen, inwiefern die Verbreitung derartiger Folgerungen gerechtfertigt ist.

K3 Menschlich gestaltete Interaktion

Evaluationen sind so angelegt, dass die Kontakte zwischen den Beteiligten von gegenseitiger Achtung geprägt sind.

Evaluatorsinnen und Evaluators sollten es vermeiden, die Würde und Selbstachtung der Personen, mit denen sie im Verlauf der Evaluation in Kontakt treten, in irgendeiner Form zu verletzen. Durch geeignetes Verhalten sollte vermieden werden, dass sich gegenüber der Evaluation feindliche Gefühle entwickeln. Dies ist nicht nur ein Gebot der Menschenwürde, sondern hat auch praktische Gründe: Personen, die sich in ihrer Würde oder in ihrer Selbstachtung eingeschränkt fühlen, verlieren nicht nur an kreativem Potential. Sie verhalten sich oft auch so, dass die Evaluation in ihren Aktivitäten eingeschränkt wird. Deshalb ist es notwendig, die kulturellen und sozialen Werte der involvierten Personen zu kennen und zu verstehen sowie zu bedenken, welche Bedeutung die einzelnen Personen der Evaluation beimessen.

K4 Vollständige und faire Einschätzung

Evaluationen sind in der Überprüfung und in der Darstellung der Stärken und Schwächen des Evaluationsgegenstandes vollständig und fair, so dass die Stärken weiter ausgebaut und die Problemfelder behandelt werden können.

Die ausgewogene Darstellung der Stärken und Schwächen eines Evaluationsgegenstandes zielt auf dessen komplette und faire Einschätzung ab. Auch wenn das primäre Ziel einer Evaluation oft eine Schwachstellenanalyse bildet, entbindet dies also nicht von der Pflicht, auch positive Aspekte ausfindig zu machen und diese zu benennen. Oft ist es nämlich möglich, bestehende Schwächen mit vorhandenen Stärken zu korrigieren. Weiter ist zu bedenken, dass die Korrektur von Schwächen des Evaluationsgegenstandes dazu führen kann, dass Stärken geschwächt werden. Es ist deshalb nützlich, den Bericht vor Abschluss der Evaluation durch externe Personen mit möglicherweise anderen Einstellungen hinsichtlich der Darlegung der positiven und negati-

ven Punkte durchsehen zu lassen. Wenn es aus irgendwelchen Gründen (z.B. wegen zeitlicher oder finanzieller Restriktionen) nicht möglich ist, gewisse Daten zu erheben, sind diese Lücken deutlich kenntlich zu machen. Auftraggeber und Auftraggeberinnen sollten Interventionen unterlassen, die eine ausgewogene Berichterstattung in Frage stellen könnten.

K5 Offenlegung der Ergebnisse

Die Vertragsparteien einer Evaluation stellen sicher, dass die Evaluationsergebnisse den betroffenen Personen zugänglich gemacht werden, ebenso wie all jenen, die einen ausgewiesenen Anspruch darauf haben.

Bei der Verbreitung der Berichts ist darauf zu achten, dass alle Personen, die in irgendeiner Form an der Evaluation beteiligt oder von dieser betroffen sind, Zugang zum Bericht erhalten. Da der Kreis der Beteiligten & Betroffenen oft sehr gross ist, muss ein Bericht in vielen Fällen öffentlich zugänglich gemacht werden. Die an einer Evaluation beteiligten Akteure (primär der Auftraggeber oder die Auftraggeberin und der Evaluator oder die Evaluatorin, aber auch weitere Personen) tragen gemeinsam die Verantwortung, dass dies gewährleistet ist. Der Bericht sollte zudem in einer Form abgefasst sein, die den Bedürfnissen der Adressaten nachkommt. Dies führt oft dazu, dass die Evaluation, angepasst an den Adressatenkreis in geeigneter Weise zu kommunizieren ist, also dass beispielsweise bei umfangreicheren Berichten Zusammenfassungen und Übersetzungen erstellt werden oder dass ein Anhang mit methodischen Hinweisen angefügt wird.

K6 Deklaration von Interessenkonflikten

Interessenkonflikte werden offen und aufrichtig behandelt, damit sie die Evaluationsverfahren und -ergebnisse möglichst wenig beeinträchtigen.

Es besteht eine Vielzahl von Konstellationen, in denen sich Evaluatoren und Evaluatorinnen mit Interessenkonflikten konfrontiert sehen, haben sie doch auch Eigeninteressen, welche die Resultate einer Evaluation unter Umständen beeinträchtigen können. So sind Evaluatorinnen oder Evaluatoren mehr oder weniger darauf angewiesen, weiterhin Aufträge zu erhalten, verfügen über spezifische philosophische, theoretische, methodische oder politische Sichtweisen und sind persönlich und organisatorisch in ein Netz eingebunden. Auch Konflikte zwischen im Kontext auffindbaren Interessengruppen können in der Evaluation aufbrechen. Dies kann zur Folge haben, dass die Evaluation in ihrem Prozess, ihren Ergebnissen und Interpretationen verzerrt wird. Interessenkonflikte sollten grundsätzlich vermieden werden. Aufgrund der angedeuteten Vielfalt von möglichen Interessenkonflikten ist es jedoch oft nicht möglich, diese gänzlich zu vermeiden. In diesen Fällen geht es darum, Wege zu entwickeln, um mit diesen umzugehen, ohne dass die Evaluation darunter leidet.

Genauigkeit

Die Genauigkeitsstandards stellen sicher, dass eine Evaluation gültige und verwendbare Informationen hervorbringt und vermittelt.

G1 Dokumentation des Evaluationsgegenstandes

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

Das Objekt einer Evaluation, sei es nun eine Massnahme, ein Programm oder eine Organisation, ist genau zu untersuchen. Es ist darauf zu achten, dass sich der Evaluationsgegenstand in unterschiedlichen zeitlichen und räumlichen Kontexten verschieden ausgestalten kann. Die Beschreibung des Evaluationsgegenstandes sollte klar machen, was präzise untersucht wurde. Dies ermöglicht den Adressaten auch den Vergleich mit anderen Evaluationsgegenständen. Weiterhin hilft eine genaue Untersuchung des Evaluationsgegenstandes, Zusammenhänge zwischen diesem und seinen Wirkungen festzustellen und bisher unbeachtete Nebenwirkungen aufzufinden. Dabei ist besonders darauf zu achten, inwiefern sich zwischen der ursprünglich vorgesehenen Form des Evaluationsgegenstandes und dessen Ausgestaltung in der Umsetzung Diskrepanzen zeigen.

G2 Kontextanalyse

Die Einflüsse des Kontextes auf den Evaluationsgegenstand werden identifiziert.

Unter dem Kontext des Evaluationsgegenstandes verstehen wir die Kombination aller Rahmenbedingungen, welche den zu evaluierenden Gegenstand umgeben, so etwa die institutionelle Einbindung, das soziale und politische Klima, die personellen Charakteristika der massgeblich Beteiligten & Betroffenen, die Struktur der Politikarena, benachbarte (und konkurrenzierende) staatliche (aber auch private) Aktivitäten oder die wirtschaftlichen Rahmenbedingungen. Diese und weitere Kontextfaktoren sind ausreichend genau zu untersuchen, um die Evaluation in Relation zu diesen planen, durchführen und vermitteln zu können. Dieses Kontextwissen ist notwendig, um die Evaluation realistisch in bezug auf die bestehenden Möglichkeiten und abgestimmt auf den Kontext ausgestalten zu können, haben doch Kontextfaktoren oft einen massgeblichen Einfluss auf die Auswirkungen eines Evaluationsgegenstandes. Eine fundierte Kontextanalyse erlaubt es auch, die Resultate einer Evaluation in bezug auf ihre Übertragbarkeit auf andere Kontexte einschätzen zu können. Es sollte vermieden werden, den Kontext eng zu definieren. Andererseits ist aber auch davor zu warnen, die Kontextanalyse zu ausführlich anzulegen, so dass die Analyse des eigentlichen Gegenstandes nicht ausreichend gründlich erfolgen kann.

G3 Beschreibung von Zielen und Vorgehen

Ziele, Fragen und Vorgehen der Evaluation werden ausreichend genau dokumentiert und beschrieben, so dass sie identifiziert und beurteilt werden können.

Die Ziele, die mit einer Evaluation angestrebt werden, die Fragen, die durch sie bearbeitet werden sowie die gewählte Vorgehensweisen sind im Verlaufe des Evaluationsprozesses genau zu dokumentieren und im Rahmen der Berichterstattung an die Adressatinnen und Adressaten der Evaluation in nachvollziehbarer Weise zu kommunizieren. Dieser Standard zielt darauf ab, den Evaluationsprozess transparent zu gestalten. Bei der Beschreibung der Ziele und der Fragen der Evaluation ist speziell darauf zu achten, dass auch abweichende Sichtweisen berücksichtigt werden. Die Dokumentation und Beschreibung des Vorgehens umfasst die detaillierte Beschreibung der Organisation, der Erhebung, der Aufbereitung, der Erfassung, der Auswertung und der Berichterstattung. Bei der Vorgehensweise ist darauf zu achten, dass sich diese im Verlaufe des

Evaluationsprozesses verändern und damit das vorgesehene und das effektive Vorgehen Differenzen aufweisen können. Es soll auch nach Aussen deutlich dargelegt werden, welche Abweichungen entstanden und welche Gründe dafür verantwortlich sind. Diese Abweichungen sind – wie die Vorgehensweise im Allgemeinen – explizit zu begründen. Ein Verzicht auf die Deklaration der Ziele, Fragestellungen und Vorgehensweisen kann eine Evaluation zu Unrecht gegen berechnigte Kritik immunisieren, aber auch unangebrachte Einwände aufkommen lassen.

G4 Verlässliche Informationsquellen

Die in einer Evaluation genutzten Informationsquellen sind hinreichend genau beschrieben, damit die Angemessenheit der Informationen eingeschätzt werden kann.

Die Beschreibung der genutzten Informationsquellen erlaubt es den Beteiligten & Betroffenen, sich selber ein Urteil über die Qualität der aus diesen Quellen hervorgegangenen Informationen zu bilden. Als Informationsquellen dienen in Evaluationen unter anderem Individuen oder Gruppen, Dokumente, audiovisuelle Materialien, Statistiken und so weiter. Der Beizug unterschiedlicher Informationsquellen erlaubt einen Vergleich der gewonnenen Informationen. Die Glaubwürdigkeit einer Evaluation kann durch eine fehlende oder unzureichende Beschreibung der Informationsquellen in Frage gestellt werden. Neben der Beschreibung der Informationsquellen sollten die aus diesen Quellen erarbeiteten Informationen auch qualifiziert werden. Die Verlässlichkeit der Informationen ist bei der Interpretation der Ergebnisse zu berücksichtigen.

G5 Valide und reliable Informationen

Die Verfahren zur Gewinnung von Informationen werden so gewählt oder entwickelt und dann eingesetzt, dass Gültigkeit und Zuverlässigkeit der gewonnenen Interpretationen für den gegebenen Zweck sichergestellt sind.

Die empirische Erfassung eines Gegenstandsbereiches ist bis zu einem gewissen Grad Fehlerquellen ausgesetzt. Validität und Reliabilität bezeichnen zwei Qualitäten des Erfassungsvorganges. Diese Qualitäten lassen sich nur im konkreten Evaluationskontext und im Hinblick auf die spezifische Zielsetzung der empirischen Erfassung einschätzen.

Unter Validität wird die Gültigkeit der Messung verstanden, also die Frage diskutiert, inwiefern das eingesetzte Instrumentarium den Aspekt, den es vorgibt zu erfassen, auch effektiv erfassen kann. Die Reliabilität eines Erhebungsinstrumentes wiederum verweist auf die Konsistenz der Erfassung in interinstrumenteller, interpersoneller und intertemporaler Hinsicht, also auf die Zuverlässigkeit und Stabilität der empirischen Erfassung. Validität und Reliabilität einer Messung hängen eng miteinander zusammen. Bei der Auswahl und Bewertung der Erhebungsinstrumente sind beide Aspekte gleichermassen zu berücksichtigen.

G6 Systematische Informationsüberprüfung

Die in einer Evaluation gesammelten, aufbereiteten und präsentierten Informationen werden systematisch auf Fehler überprüft.

Bei der Erhebung, Erfassung, Auswertung und Interpretation von Informationen bestehen vielfältige Möglichkeiten, dass Fehler entstehen. Dies können unter anderem lapidare Tippfehler bei der Datenerfassung sein, aber auch Fehlinterpretationen des erhobenen Datenmaterials. Aus diesem Grund ist es zwingend erforderlich, dass im Evaluationsprozess darauf geachtet wird, potentielle Fehlerquellen möglichst zu vermeiden. Weiter sind die gewonnenen Informationen im Hinblick auf mögliche Fehler mit geeigneten Verfahren (Plausibilitätstest, Parallelerfassung, kommunikative Validierung usw.) zu überprüfen. Mögliche Fehler und deren Konsequenzen sind im Rahmen der Berichterstattung zur Evaluation in transparenter Weise zu diskutieren.

Gelingt es nicht, Fehler zu vermeiden, kann es zu fehlgeleiteten Interpretationen und Folgerungen kommen. Darüber hinaus können fehlerhafte Angaben – auch wenn sie materiell weniger bedeutend sind – die gesamte Evaluation diskreditieren.

G7 Analyse qualitativer und quantitativer Informationen

Qualitative und quantitative Informationen einer Evaluation werden angemessen und systematisch analysiert, damit die gestellten Fragen durch die Evaluation effektiv beantwortet werden.

Im Rahmen von Evaluationen werden aufgrund der Analysen qualitativer wie quantitativer Informationen Schlüsse gezogen. Die Auswertung der erhobenen Informationen sollte in systematischer Weise erfolgen und sich nach den Regeln der qualitativen und quantitativen Methodik richten. Zumeist erweist es sich in Evaluationen als sinnvoll und zweckmässig, sowohl qualitative wie auch quantitative Informationen einzubeziehen. Für die Wahl der Informationen wie auch der Auswertungsmethodik sind Fragestellung und Datenlage ausschlaggebend. Bei diesen Entscheiden sollten andere Faktoren, etwa die Kenntnisse und Vorlieben der beteiligten Personen, keine Rolle spielen. Die Entscheide zur Methodenwahl sind transparent darzulegen und ihre Konsequenzen sind kritisch zu hinterfragen. Auf die Aussagekraft der Methoden, wie auch auf ihre Begrenzungen, ist explizit hinzuweisen.

G8 Begründete Schlussfolgerungen

Die in einer Evaluation gezogenen Folgerungen werden ausdrücklich begründet, damit die Beteiligten & Betroffenen diese nachvollziehen und einschätzen können.

Die Schlussfolgerungen einer Evaluation müssen explizit begründet und zusammen mit den zugrundeliegenden Annahmen und den eingesetzten Evaluationsverfahren transparent dargestellt werden. Dabei ist auch deutlich auf den Geltungsbereich der Schlussfolgerungen hinzuweisen. Ebenso sollte in der Berichterstattung eine Diskussion alternativer Interpretationen enthalten sein, wobei begründet werden sollte, warum diese verworfen wurden. Beim Setzen von Prämissen, welche den Folgerungen zugrunde liegen, sollte darauf geachtet werden, dass Annahmen vermieden werden, die von relevanten Akteuren nicht geteilt werden. Die Einhaltung dieses Standards erlaubt es den Nutzerinnen und Nutzern der Evaluationsergebnisse, die Aussagekraft der Folgerungen einzuschätzen. Zudem wird dadurch die Überzeugungskraft der Folgerungen gestärkt.

G9 Unparteiische Berichterstattung

Die Berichterstattung wird gegenüber Verzerrungen durch persönliche Gefühle und Vorlieben irgendeiner Evaluationspartei geschützt, so dass Evaluationsberichte die Ergebnisse fair wiedergeben.

Das Umfeld einer Evaluation ist durch vielfältige Perspektiven geprägt. Die Beteiligten & Betroffenen ('stakeholders') einer Evaluation verfügen oft über divergierende Ansichten hinsichtlich des Evaluationsgegenstandes. Für eine Evaluation besteht nun die Gefahr, dass sie sich von einer bestimmten Gruppe vereinnahmen oder instrumentalisieren lässt. Eine Evaluation sollte darauf verzichten, lediglich eine spezifische Sichtweise zu übernehmen. Vielmehr sollte sie an einer fairen Vertretung aller relevanten Interessen interessiert sein. Deshalb ist dafür zu sorgen, dass die Evaluation eine möglichst unabhängige Position einnimmt. So ist ein zu enges Verhältnis zu Auftraggeberinnen oder Auftraggebern, aber auch beispielsweise zu den für den Evaluationsgegenstand verantwortlichen Personen, zu vermeiden. Das Verhältnis der Evaluationsverantwortlichen gegenüber den Auftraggeberinnen und Auftraggebern der Evaluation und anderen relevan-

ten Gruppen ist zu Beginn des Evaluationsprozesses zu klären. Diese Rollenklärung umfasst auch Festlegungen über das Recht, Evaluationsberichte zu veröffentlichen.

G10 Meta-Evaluation

Die Evaluation selbst wird anhand der vorliegenden oder anderer wichtiger Standards evaluiert, so dass die Durchführung entsprechend angeleitet werden kann und damit die Beteiligten & Betroffenen bei Abschluss einer Evaluation deren Stärken und Schwächen einschätzen können.

Misslungene Evaluationen können Anlass zu Fehlentscheidungen geben. Evaluationen können aber auch (zu unrecht) heftiger Kritik ausgesetzt sein. Um derartige Situationen zu vermeiden, sollte auch die Evaluation selbst hinsichtlich ihrer Qualität überprüft werden. Im Rahmen einer Meta-Evaluation (Evaluation einer Evaluation) können beispielsweise die hier vorliegenden Standards dazu eingesetzt werden, um eine Evaluation zu bewerten. Eine Meta-Evaluation kann je nach Situation, summativ (bilanzierend) wie auch formativ (gestaltend) angelegt werden. Sie kann im Rahmen des Evaluationsteams (also intern) erfolgen, aber auch durch Aussenstehende (extern). Wie bei Evaluationen auch, kann eine Meta-Evaluation mehr oder weniger umfangreich ausgestaltet werden. Währenddem eine vertiefte und umfassende Meta-Evaluation nur in bestimmten Fällen zweckmässig ist, sollte auf eine knappe Selbstevaluation der Evaluation durch die an der Evaluation beteiligten Personen keinesfalls verzichtet werden. Die (in der Regel bescheidenen) Ressourcen für die Meta-Evaluation sind im Rahmen der Evaluationsplanung bereitzustellen.

BEGLEITMATERIALIEN

Funktionale Übersicht über die SEVAL-Standards

Die nachfolgende Übersicht zeigt auf, welche Standards bei den wichtigsten Evaluationsstätigkeiten besonders zu beachten sind.

Entscheid über die Durchführung einer Evaluation

- N1** Ermittlung der Beteiligten & Betroffenen
- N2** Klärung der Evaluationsziele
- N3** Glaubwürdigkeit
- N8** Wirkung der Evaluation
- D2** Politische Tragfähigkeit
- D3** Kostenwirksamkeit
- K1** Formale Vereinbarung
- K6** Deklaration von Interessenkonflikten
- G1** Dokumentation des Evaluationsgegenstandes
- G2** Kontextanalyse
- G10** Meta-Evaluation

Definition des Evaluationsproblems

- N1** Ermittlung der Beteiligten & Betroffenen
- N2** Klärung der Evaluationsziele
- G1** Dokumentation des Evaluationsgegenstandes
- G2** Kontextanalyse
- G3** Beschreibung von Zielen und Vorgehen
- G10** Meta-Evaluation

Planung der Evaluation

- N1** Ermittlung der Beteiligten & Betroffenen
- N2** Klärung der Evaluationsziele
- N4** Umfang und Auswahl der Informationen
- N5** Transparenz der Bewertung
- D1** Praktische Verfahren
- K1** Formale Vereinbarung
- K4** Vollständige und faire Einschätzung
- G1** Dokumentation des Evaluationsgegenstandes
- G3** Beschreibung von Zielen und Vorgehen
- G4** Verlässliche Informationsquellen
- G5** Valide und reliable Informationen
- G7** Analyse qualitativer und quantitativer Informationen
- G8** Begründete Schlussfolgerungen
- G9** Unparteiische Berichterstattung
- G10** Meta-Evaluation

Informationserhebung

- N3** Glaubwürdigkeit
- N4** Umfang und Auswahl der Informationen
- N5** Transparenz der Bewertung
- D1** Praktische Verfahren
- D2** Politische Tragfähigkeit

- K1** Formale Vereinbarung
- K2** Schutz individueller Rechte
- K3** Menschlich gestaltete Interaktion
- K4** Vollständige und faire Einschätzung
- G1** Dokumentation des Evaluationsgegenstandes
- G2** Kontextanalyse
- G3** Beschreibung von Zielen und Vorgehen
- G4** Verlässliche Informationsquellen
- G5** Valide und reliable Informationen
- G6** Systematische Informationsüberprüfung
- G10** Meta-Evaluation

Informationsauswertung

- N5** Transparenz der Bewertung
- D1** Praktische Verfahren
- G1** Dokumentation des Evaluationsgegenstandes
- G2** Kontextanalyse
- G7** Analyse qualitativer und quantitativer Informationen
- G8** Begründete Schlussfolgerungen
- G10** Meta-Evaluation

Berichterstattung zur Evaluation

- N1** Ermittlung der Beteiligten & Betroffenen
- N4** Umfang und Auswahl der Informationen
- N5** Transparenz der Bewertung
- N6** Vollständigkeit und Klarheit des Berichts
- N7** Rechtzeitigkeit der Berichterstattung
- N8** Wirkung der Evaluation
- K2** Schutz individueller Rechte
- K4** Vollständige und faire Einschätzung
- K5** Offenlegung der Ergebnisse
- G1** Dokumentation des Evaluationsgegenstandes
- G2** Kontextanalyse
- G3** Beschreibung von Zielen und Vorgehen
- G4** Verlässliche Informationsquellen
- G8** Begründete Schlussfolgerungen
- G9** Unparteiische Berichterstattung
- G10** Meta-Evaluation

Budgetierung der Evaluation

- N2** Klärung der Evaluationsziele
- N4** Umfang und Auswahl der Informationen
- D3** Kostenwirksamkeit
- K1** Formale Vereinbarung
- G1** Dokumentation des Evaluationsgegenstandes
- G3** Beschreibung von Zielen und Vorgehen
- G10** Meta-Evaluation

Abschluss eines Evaluationsvertrages

- N1** Ermittlung der Beteiligten & Betroffenen
- N2** Klärung der Evaluationsziele
- N3** Glaubwürdigkeit
- N4** Umfang und Auswahl der Informationen

- N7 Rechtzeitigkeit der Berichterstattung
- D2 Politische Tragfähigkeit
- K1 Formale Vereinbarung
- K2 Schutz individueller Rechte
- K5 Offenlegung der Ergebnisse
- K6 Deklaration von Interessenkonflikten
- G1 Dokumentation des Evaluationsgegenstandes
- G3 Beschreibung von Zielen und Vorgehen
- G10 Meta-Evaluation

Leiten der Evaluation / Evaluationsmanagement

- N1 Ermittlung der Beteiligten & Betroffenen
- N2 Klärung der Evaluationsziele
- N3 Glaubwürdigkeit
- N7 Rechtzeitigkeit der Berichterstattung
- D2 Politische Tragfähigkeit
- D3 Kostenwirksamkeit
- K1 Formale Vereinbarung
- K2 Schutz individueller Rechte
- K3 Menschlich gestaltete Interaktion
- K6 Deklaration von Interessenkonflikten
- G3 Beschreibung von Zielen und Vorgehen
- G6 Systematische Informationsüberprüfung
- G10 Meta-Evaluation

Personelle Ausstattung der Evaluation

- N3 Glaubwürdigkeit
- D2 Politische Tragfähigkeit
- K6 Deklaration von Interessenkonflikten
- G9 Unparteiische Berichterstattung
- G10 Meta-Evaluation

Erläuterungen zur Vorgehensweise bei der Formulierung der SEVAL-Standards

Die SEVAL-Standards sind von 'The Program Evaluation Standards' des 'Joint Committee on Standards for Educational Evaluation' (Joint Committee 1994) in der von Wolfgang Beywl und Thomas Widmer übersetzten Fassung (Joint Committee 2000) abgeleitet worden. In einer Umfrage haben sich die Mitglieder der Arbeitsgruppe Evaluationsstandards der SEVAL zu dieser Fassung der Standards geäussert. Aufgrund der eingegangenen Stellungnahmen der Arbeitsgruppenmitglieder hat sodann ein Ausschuss der Arbeitsgruppe eine Überarbeitung vorgenommen. Diese Überarbeitung wurde im Rahmen der Arbeitsgruppe wiederum diskutiert. Der Ausschuss hat aufgrund der Ergebnisse dieser Diskussion den Text angepasst. Am 26. Mai 2000 wurde diese Fassung im Rahmen der Tagung 'Qualitätssicherung in der Evaluation' in Bern, an der rund achtzig Personen teilnahmen, zur Diskussion gestellt. Aufgrund der so gewonnen Erfahrungen wurde der Text einer weiteren Überarbeitung im Rahmen der Arbeitsgruppe und des Ausschusses unterzogen, aus der schliesslich der vorliegende Text hervorging.

Dabei wurden insbesondere folgende Änderungen vorgenommen: Generell wurden die Termini 'Programm' (resp. 'Programmevaluation') durch 'Evaluationsgegenstand' (resp. 'Evaluation') ersetzt. Dies hat keine inhaltlichen Konsequenzen, wird der Begriff 'Programm' im amerikanischen Original in einem weiten Sinn verstanden. Weiter wurden zur sprachlichen Vereinfachung 'soll'- in 'ist'-Aussagen umgewandelt. Neben einer generellen sprachlichen Überarbeitung ist zudem auf die folgenden Änderungen hinzuweisen (Kurzbezeichnungen aus Joint Committee 2000):

- Zusammenlegung der Standards G5 und G6 sowie der Standards G8 und G9
- Neuformulierung des Standards N7, D1, D3, K3, K4, K6, K7, G2, G7, G10 und G12
- Ersatzlose Streichung der Standards K1 und K8
- Ergänzung um einen neuen Standard 'Klärung der Evaluationsziele' (neu: N2)

Im folgenden werden die einzelnen Änderungen erläutert und begründet.

Standards, die zusammengelegt wurden:

Die *Standards G5 und G6* befassen sich mit zwei Qualitätskriterien sozialwissenschaftlichen Arbeitens, der Validität und der Reliabilität. Da die Beurteilung dieser Aspekte für sich alleine genommen wegen der Interdependenzen wenig Sinn macht (vgl. dazu Widmer 1996: 296), wurden diese beiden Standards in einen Standard zusammengefasst.

Da sich Evaluationen in der Regel sowohl auf qualitative, wie auch auf quantitative Informationen abstützen sollten, wurden die Kriterien *G8 zur Analyse quantitativer Informationen* und *G9 zur Analyse qualitativer Informationen* in einem Standard zusammengefasst.

Standards, die neu formuliert wurden:

N7 Wirkung der Evaluation: Zur sprachlichen Vereinfachung wurde eine andere Formulierung gewählt, die jedoch zu keiner inhaltlichen Veränderung führt.

D1 Praktische Verfahren: Auch hier wurde eine rein sprachliche Vereinfachung vorgenommen.

D3 Kostenwirksamkeit: Der Standard wurde um das Erfordernis der Effizienz gekürzt, da es sich dabei um eine Anforderung handelt, die bereits im Rest des Standards enthalten und damit redundant ist.

K3 Schutz individueller Rechte: Der ursprüngliche Titel dieses Standards (Schutz individueller Menschenrechte) trifft den Gehalt des Standards nicht ganz, geht dieser doch über die Menschenrechte im engeren Sinne hinaus.

K4 Menschlich gestaltete Interaktion: Der Standard wurde hier neu formuliert, da die darin enthaltene Formulierung, dass 'Würde und Wert der Menschen' zu respektieren sei, wegen der darin zum Ausdruck kommenden Asymmetrie (zwischen Evaluatorinnen und Evaluatoren und anderen Beteiligten), nicht überzeugte.

K6 Offenlegung der Ergebnisse: Der ursprünglich (nach „Evaluationsergebnisse“) enthaltene Einschub „einschliesslich wesentlicher Einschränkungen“ wurde gestrichen, da nicht ganz klar wird, welche Einschränkungen gemeint sind und weil dies ja bereits Gegenstand anderer Standards ist (siehe N4, N5, K5 und G4).

K7 Deklaration von Interessenkonflikten: Da Beeinträchtigungen durch Interessenkonflikte nicht immer vollständig vermieden werden können, wurde die absolute Formulierung „nicht beeinträchtigen“ mit „möglichst wenig beeinträchtigen“ relativiert.

G2 Kontextanalyse: Der Standard wurde einfacher formuliert, ohne Änderungen an der inhaltlichen Aussage vorzunehmen.

G7 Systematische Informationsüberprüfung: Auf den Nachsatz, dass alle gefundenen Fehler korrigiert werden sollen, wurde verzichtet, da der explizite Hinweis auf die Korrektur als unnötig erachtet wird.

G10 Begründete Schlussfolgerungen: Der Standard wurde zur Verdeutlichung um das Erfordernis der Nachvollziehbarkeit ergänzt, die eine wichtige Voraussetzung für eine fundierte Einschätzung darstellt.

G12 Meta-Evaluation: Dieser Standard wurde in zweierlei Hinsicht leicht angepasst. Erstens wurde für ein besseres Verständnis auf die Fachtermini „formativ“ und „summativ“ verzichtet. Zweitens wurde die Formulierung „gründlich überprüfen“ durch „einschätzen“ ersetzt, um vor allem im Hinblick auf kleinere Evaluationsprojekte eine angemessene Meta-Evaluation zuzulassen.

Standards, die ersatzlos gestrichen wurden:

K1 Unterstützung der Dienstleistungsorientierung

Die Evaluation sollte so geplant werden, dass Organisationen dabei unterstützt werden, die Interessen und Bedürfnisse des ganzen Zielgruppenspektrums zu berücksichtigen und ihre Tätigkeiten danach auszurichten.

Begründung für die Streichung:

Einer der Grundsätze für die Standards sollte es sein, dass sie für möglichst alle vorgesehenen Anwendungsbereiche (wie sie in der Einleitung definiert wurden) Gültigkeit haben. Nicht alle Evaluationsgegenstände sind Organisationen mit Dienstleistungscharakter zugeordnet (z.B. Evaluationen im privatwirtschaftlichen Bereich). Es scheint auch

nicht von vornherein immer klar zu sein, dass die Berücksichtigung der Bedürfnisse des *ganzen* Zielgruppenspektrums sinnvoll ist. Es mag auch Fälle geben, in denen es wichtig ist, sich eben genau umgekehrt auf ein bestimmtes Segment zu konzentrieren.

K8 Finanzielle Verantwortlichkeit

Die Zuweisung und Ausgabe von Ressourcen durch die Evaluatorin sollte durch eine sorgfältige Rechnungsführung nachgewiesen werden und auch anderweitig klug sowie ethisch verantwortlich erfolgen, damit die Ausgaben verantwortungsbewusst und angemessen sind.

Begründung für die Streichung:

- Der Standard vermischt mehrere, wenn auch zum Teil zusammenhängende Anliegen (sorgfältige Rechnungsführung, ethisch verantwortliche Zuweisung von Ressourcen, Nachweis der Ressourcenausgabe, 'anderweitig kluge' Zuweisung und Ausgabe von Ressourcen, verantwortungsbewusste und angemessene Ausgaben ...)
- Mit dem Standard D3 ist das Kernanliegen von K8 bereits abgedeckt.
- Vom Moment des Vertragsabschlusses an ist die Beziehung zwischen Evaluator und Auftraggeber eine normale Geschäftsbeziehung, die in Vertrag und OR geregelt ist. Viel wesentlicher sind/wären die sich vor Vertragsabschluss abspielenden Vorgänge (Submissionspraxis von öffentlicher Hand und privaten Organisationen). Die Interessen der Auftraggeberinnen und Auftraggeber wären demnach auch besser geschützt durch entsprechende Formulierungen in Submissionsreglementen.
- Sehr viele Aufträge im Dienstleistungsbereich sind pauschale Abgeltungen, die auf einem Leistungsverzeichnis basieren. Vom Moment an, wo ein Vertrag unterzeichnet ist, erlischt der Anspruch auf Einsicht in die Rechnungsführung. So betrachtet ist der Standard wirklichkeitsfremd und naiv.
- Der Standard verlässt das Prinzip der Symmetrie, weil nirgendwo sonst in den Standards Auftraggeber verpflichtet werden, die für 'ethisch verantwortliche' Evaluationen und 'verantwortliche und angemessene Ausgaben' nötigen Ressourcen bereitzustellen.

Standard, der ergänzt wurde:

N2 Klärung der Evaluationsziele

Alle an einer Evaluation mitwirkende Personen sind dafür besorgt, dass die Ziele der Evaluationen für alle Beteiligten & Betroffenen geklärt sind.

Dieser Standard wurde neu in die SEVAL-Standards aufgenommen, da die Arbeitsgruppe die Auffassung vertritt, dass hier in den Standards des Joint Committee eine wesentliche Lücke vorliegt. Die Ziele der Evaluation werden zwar an verschiedenen Stellen in den Standards implizit oder auch explizit thematisiert (vgl. etwa N1 oder G3). Die Arbeitsgruppe hält es aufgrund der grossen Bedeutung, welche der Klärung der Evaluationsziele beigemessen wird, für gerechtfertigt, einen eigenen und expliziten Standard zu diesem Aspekt einzuführen.

Transformationstabelle zu den Kurzbezeichnungen

Durch die zuvor erläuterten Änderungen ergeben sich in den Kurzbezeichnungen folgende Korrespondenzen zwischen den SEVAL-Standards und den 'Program Evaluation Standards' (in der deutschen Fassung; siehe Joint Committee 2000):

<i>JC</i>	<i>SEVAL</i>	<i>JC</i>	<i>SEVAL</i>	<i>JC</i>	<i>SEVAL</i>	<i>JC</i>	<i>SEVAL</i>
N1	N1	D1	D1	K6	K5	G7	G6
fehlt	N2	D2	D2	K7	K6	G8/G9	G7
N2	N3	D3	D3	K8	entfällt	G10	G8
N3	N4	K1	entfällt	G1	G1	G11	G9
N4	N5	K2	K1	G2	G2	G12	G10
N5	N6	K3	K2	G3	G3		
N6	N7	K4	K3	G4	G4		
N7	N8	K5	K4	G5/G6	G5		

Mitglieder der Arbeitsgruppe 'Evaluationsstandards' der SEVAL

Dr. Paul Ammann, Gesundheits- und Fürsorgedirektion des Kantons Bern, Bern
 Dr. Nicole Bachmann, Gesundheits- und Fürsorgedirektion des Kantons Bern, Bern*
 Renate Cléménçon, Bundesamt für Justiz, Bern
 Valérie Eggimann-Lanz, Eidgenössische Finanzkontrolle (EFK), Bern
 Frohmüt W. Gerheuser, Büro für Politikberatung und Sozialforschung, Brugg (AG)
 Dr. Armin Kühne, Eidgenössische Finanzkontrolle (EFK), Bern
 Charles Landert, Landert Farago Davatz & Partner, Zürich*
 Dr. Marlène Läubli Loud, Office fédéral de la santé publique, Berne
 Dr. Tobias Ledergerber, Organisationsamt des Kantons Bern, Bern
 Dr. Urs Moser, Kompetenzzentrum Bildungsevaluation & Leistungsmessung, Universität Zürich
 Urs Schmidt, e&e Entwicklung & Evaluation im Sozialwesen, Zürich
 Prof. Dr. Christian Simon, Schweizerischer Wissenschafts- und Technologierat (SWTR), Bern
 Dr. Heini Sommer, Ecoplan, Altdorf (UR)
 Patrick Trees, Staatskanzlei des Kantons Bern, Bern
 Dr. Thomas Widmer, Institut für Politikwissenschaft, Universität Zürich, Zürich (Leitung)*

*: Mitglieder des Ausschusses der Arbeitsgruppe

Literatur

Joint Committee on Standards for Educational Evaluation (Ed.) (1994): The Program Evaluation Standards, 2nd Edition. Newbury Park: 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: Leske + Budrich.

Widmer, Thomas (1996): Meta-Evaluation. Kriterien zur Bewertung von Evaluationen. Bern: Haupt.

Widmer, Thomas (2000): „Qualität der Evaluation – Wenn Wissenschaft zur praktischen Kunst wird.“ In: Stockmann, Reinhard (Hrsg.): Evaluationsforschung. Opladen: Leske + Budrich: 77-102.