

## Sind Evaluationen grau und langweilig? Veröffentlichung, Verwertung und Verbreitung von Entwicklungsevaluierungen

Bericht über die Frühjahrstagung des Arbeitskreises  
Entwicklungspolitik 2015 am 13. und 14. April 2015  
bei Brot für die Welt, Berlin

*Daniela Richter,<sup>1</sup> Jens Koy,<sup>2</sup> Susanne Neubert,<sup>2</sup> Susanne von Jan<sup>2</sup>*

### Hintergrund und Tagungsprogramm

Sind unsere Evaluierungen viel zu technisch und zu langweilig, als dass die Öffentlichkeit sie wahrnimmt? Zur Erfüllung von Evaluationsaufträgen sind in aller Regel auch Datenbearbeitungen und Bewertungen zu technischen Details notwendig – kann man da erwarten, dass sie in Gänze tatsächlich durch die Öffentlichkeit wahrgenommen werden? Ist die Öffentlichkeit überhaupt an der Wirksamkeit von Entwicklungsvorhaben im Einzelnen interessiert? Wenn ja, soll an der Vermarktung der Ergebnisse stärker gearbeitet werden? Diesen und weiteren Fragen zum Thema Verwertung, Verbreitung und Veröffentlichung von Entwicklungsevaluierungen widmete sich der AK Entwicklungspolitik auf seiner diesjährigen Frühjahrstagung. Wie im vorigen Jahr wurden Präsentationsformate mit interaktiven Werkstattgesprächen kombiniert.

Vor dem Einstieg in das Thema der Frühjahrstagung stellte die AG Effizienz, vertreten durch Barbara Winker (Kindernothilfe), Bernhard Hoeper (Welthungerhilfe) und Jens Koy (freier Gutachter), den aktuellen Stand der Handreichung „Effizienzbewertung in Evaluationen“ vor. Anschließend gab Alexander Kocks (DEval) einen Input zur *Cost-Utility-Analysis*. Der freie Journalist Dirk Asendorpf betrachtete daraufhin die Frage, warum Medien kaum über Evaluierungen berichten. In vier parallelen Werkstattgesprächen wurde dieser Frage weiter nachgegangen. Der zweite Tag begann mit einem Vortrag über die Organisation von Wahrnehmung von Evalu-

---

1 Seminar für ländliche Entwicklung (SLE), Berlin

2 Sprecherteam des Arbeitskreises Entwicklungspolitik: Susanne Neubert (SLE – Seminar für Ländliche Entwicklung, Berlin), Jens Koy (Freier Gutachter und Evaluator), Susanne von Jan (HSS-Hanns-Seidel-Stiftung)

ationen durch Martina Vahlhaus (GIZ). Anschließend berichtete Eva Terberger (Kreditanstalt für Wiederaufbau (KfW) Entwicklungsbank) über die Farbigkeit von EZ-Evaluierungen und die dennoch oft graue Berichterstattung.

## Bericht der AG Effizienz: Vorstellung der Handreichung „Effizienzbewertung in Evaluationen“

Die drei Vertreter(innen) der Arbeitsgruppe Effizienz des Arbeitskreises Entwicklungspolitik der DeGEval stellten die zentralen Elemente des Entwurfs der Handreichung „Effizienzbewertung in Evaluationen“ vor, die noch in 2015 als „Arbeitspapier des Arbeitskreises Entwicklungspolitik“ verabschiedet werden sollte. Die Handreichung richte sich an Gutachter(innen) und Auftraggebende, die die Analysen im DAC-Kriterium<sup>3</sup> Effizienz in Projekt- und Programmevaluationen weiter qualifizieren wollen. Im Zentrum der Betrachtung stünden „durchschnittliche“ Projektevaluierungen und Programme der Entwicklungszusammenarbeit. Kernstück der Handreichung sei das Kapitel „Effizienzbewertung“, in dem die Erarbeitung von spezifischen Evaluationsfragen, die Erfassung der Kostenseite, Maßstäbe für die Effizienzbewertung sowie Verfahren der Effizienzbewertung vorgestellt würden. Wichtig sei, dass Evaluator(inn)en die Fragestellungen der Effizienzbewertung in ihr Methodenrepertoire integrieren könnten. Die Handreichung verschaffe außerdem einen Überblick über theoretische Grundlagen und die aktuelle Diskussion zum Thema. Dabei würden das Effizienzverständnis im entwicklungspolitischen Kontext, Perspektiven der Effizienzbetrachtung sowie Stärken und Herausforderungen des DAC-Kriteriums Effizienz dargestellt.

Eine wichtige Grundlage der vorliegenden Handreichung sei die Diskussion der Studie „Tools and Methods for Evaluating the Efficiency of Development Interventions“ (Palenberg 2011), die im Auftrag des BMZ erstellt wurde und einen Überblick über angewandte Methoden im Bereich der Effizienzbewertung gibt. Sie kann als Referenzstudie im deutschsprachigen Raum angesehen werden. Für viele Evaluator(inn)en sei diese Studie ein theoretisch gut fundierter Einstieg in die Fragestellungen der Effizienz. Für die Umsetzung in die Praxis der Projektevaluationen fehlten jedoch in vielen Bereichen weitere Hilfestellungen und Anregungen. Mit dem hier vorgelegten Beitrag solle diese Lücke zwischen Theorie und Praxis in der Effizienzbewertung weiter geschlossen werden. Die Kolleg(inn)en der AG haben zu diesem Zweck Erfahrungen ausgetauscht, recherchiert und aufbereitet. Die Handreichung verkünde daher keine abschließenden Weisheiten, sondern ermuntere nur zur maßvollen Anwendung von Methoden und Vorgehensweisen. Sehr wichtig sei jetzt die Anwendung in der Evaluationspraxis.

---

3 DAC: Development Assistance Committee der OECD (Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung)

## Effizienzmessung von Entwicklungsmaßnahmen mit multiplen Outcomes – Praktikabilität der Kosten-Nutzwert-Analyse in der Entwicklungsvaluierung

Herr Kocks beschäftigte sich mit der Frage, wie der Arbeitskreis Entwicklungspolitik die Erkenntnisse des Palenberg-Papiers nutzen könne. Die Kosten-Nutzwert-Analyse (*Cost-Utility-Analysis*, CUA), ein Verfahren zur relativen Effizienzmessung, sei geeignet für die Ermittlung von Produktions- und Allokationseffizienz. Über den Vergleich von Maßnahmen ließe sich z.B. feststellen, ob die Ziele im gleichen Umfang mit geringerem Mitteleinsatz erreicht würden (Kostenminimierung) oder einen höheren Zielumfang mit demselben Mitteleinsatz erreichten (Zielmaximierung). Die CUA gehöre zu den anspruchsvolleren Maßnahmen, bedürfe aber keiner ökonomischen Kenntnisse. Außerdem wäre sie auf Interventionen mit multiplen, nicht monetären Outcomes anwendbar. Sie sei in der EZ-Evaluierungspraxis daher besonders geeignet. Die Auswahl des Verfahrens zur relativen Effizienzmessung bliebe aber abhängig vom Evaluierungsgegenstand, die CUA und andere Verfahren wie die *Cost-Benefit-Analysis* oder die *Cost-Effectiveness-Analysis* stünden daher nicht in Konkurrenz zueinander. Da die Outcomes eines Vorhabens in unterschiedlichen Einheiten vorlägen, müssten sie in Nutzwerte transferiert werden, bevor sie in Beziehung gesetzt werden könnten. Diese Transformation impliziere eine Bewertung, die am besten in Absprache mit der Zielgruppe vorgenommen würde.

Vorteile seien die vergleichsweise einfache Durchführbarkeit und die Anwendung auf Vorhaben mit multiplen Outcomes ohne eine ethisch umstrittene Monetarisierung (wie im Falle der Kosten-Nutzen-Analyse). Darüber hinaus leiste die CUA durch den Einbezug der Stakeholder einen Beitrag zur Konsensbildung und partizipativen Entscheidungsfindung. Nachteile seien aber, dass die CUA nur so valide wie die eingesetzte Submethode zur Nutzenmessung und -gewichtung sei. Außerdem setze die relative Effizienzmessung die Vergleichbarkeit von Vorhaben voraus – die Kontextfaktoren, die die Kosten und Zielerreichung beeinflussen, seien aber in jedem Vorhaben unterschiedlich.

## Abstrakt, gesichtslos, selbstverliebt – Warum über Evaluationen nicht berichtet wird

Herr Asendorpf beschäftigte sich in seiner Präsentation mit der Frage, warum über Evaluationen nicht berichtet wird. Aufmerksamkeit sei die Währung des Journalismus, Evaluationen seien durch ihren abstrakten Gegenstand, durch ihre „gesichtslose Sprache und ihre selbstverliebten Inhalte“ jedoch kein Gegenstand der Medien. Zum einen fänden Evaluationen ungeachtet ihres breiten Anwendungsbereichs nach wie vor sehr wenig Aufmerksamkeit im gesamten öffentlichen Raum. Die Kriterien für gute Medienthemen seien folgende: neu, relevant, originell, menschlich, tragisch, komisch oder merkwürdig, ungeschlossen oder widersprüchlich und kontraintuitiv. Evaluationen hingegen müssten gründlich, fair, erschöpfend, transparent, hilfreich und vertraulich sein. Sie erfüllten daher einzig die Relevanz-An-

forderung. Allein aufgrund der Vertraulichkeit der Informationen in Evaluationen und dem Datenschutz der involvierten Stakeholder seien sie nicht zur Veröffentlichung geeignet. Medienveröffentlichungen mit Eigen- oder Projektnamen tauchten jahrelang ganz oben in Google-Trefferlisten auf. Werden Evaluationsergebnisse doch als Quelle für PR-Material in Rechenschafts- und Jahresberichten als Information für Spender(innen) veröffentlicht, böte sich eine unabhängige Aufbereitung von Journalist(inn)en an, in der Projekte und Personen aber anonym blieben – wie etwa MISEREOR es in seinem Jahresevaluierungsbericht tue.

Medien mit unabhängigen Redaktionen, wie die großen Tageszeitungen, könnten vor allem durch originelle, kontraintuitive Themen begeistert werden. Da sie allerdings unabhängig in ihrer Recherche und Berichterstattung seien, ließen sie sich in der letztendlichen Themenwahl und Schwerpunktsetzung sowie der Auswahl ihrer Quellen nicht beeinflussen. Folglich könne kein Einfluss auf die Endfassung der Berichterstattung genommen werden. Die Veröffentlichung in diesen Medien berge also auch ein gewisses Risiko. Somit bliebe ein großer Teil des Verwertungspotenzials von Evaluationen ungenutzt. Die Gutachter(innen) und EZ-Organisationen sollten sich daher überlegen, mit welcher Geschichte (von Menschen handelnd) und mit welcher klaren Botschaft sie in die Medien wollten.

## Werkstattgespräche

Der freie Journalist Dirk Asendorpf beschäftigte sich mit der Frage, wie die Aufmerksamkeit der Medien erreicht werden könnte. Der Bezug zu Deutschland sei ein relevantes Thema, um überhaupt in die Medien zu kommen. So hätte der zivile Friedensdienst die Konfliktbewältigung als zentrales Feld, die deutschen Erfahrungen würden hier als Positivbeispiel dargestellt. Die Gruppe diskutierte, wie solche Erfahrungen für die Öffentlichkeitsarbeit genutzt werden könnten. Sie kamen zu dem Schluss, dass verschiedene Schwerpunkte und Sprachen je nach Zielgruppe unterschieden werden müssten. Auch negative Ergebnisse sollten nicht verändert, sondern mithilfe von Rahmenbedingungen begründet werden.

Katja Maurer von medico international referierte zur Perspektive der Humanitären Hilfe. Anhand eines selbstgedrehten Films, der ein kritisches Bild auf die Humanitäre Hilfe in Haiti wirft, zeigte sie die Schwierigkeit auf, kritisch zu sprechen. Gegenstand des Films waren Interviews der Projektpartner vor Ort, die die Auswirkungen der Hilfe als Ganzes beschrieben. Sie erwähnte, dass die Filmerstellung ergänzend zur klassischen Evaluierung vor Ort ablief. Durch die Interviews würde der Kontext besser beschreibbar, Fallstudien stünden im Mittelpunkt, statt im Anhang zu verschwinden. Weiter diskutiert werden müsste allerdings, welche Ergebnisse und Informationen überhaupt weitergegeben werden sollten. Die zentrale Botschaft des Films sei die Zusammenarbeit mit den Akteuren vor Ort, die durch dieses Medium gut vermittelt werden könnte.

Matthias Lanzendörfer von MISEREOR, Thorsten Baer von World Vision Deutschland und Bernhard Hoepfer von der Welthungerhilfe behandelten die Veröffentlichung von Evaluierungsergebnissen und ob diese eine Gefahr oder Chance für

spendensammelnde Organisationen sei. Die Anwesenden sehen die Gutachter(innen) und EZ-Organisationen in der Bringschuld gegenüber der Öffentlichkeit. Spender(innen) wollten und sollten wissen, was mit ihrem Geld geschieht. Auch wenn Evaluierungen nicht gelesen würden, könnten sie als Sicherheitsnetz dienen, da ihre bloße Verfügbarkeit Vertrauen in der breiten Öffentlichkeit schaffe. Wichtig sei in jedem Fall eine klare, kurze Botschaft, um die Leser(innen) nicht zu überfrachten. Abgewogen werden müsse der Aufwand der Aufbereitung. Auch schlechte Ergebnisse müssten veröffentlicht werden, um die Glaubwürdigkeit der Organisation zu erhöhen und den eigenen Lernprozess darzustellen. Vorsicht sei hier aber geboten, damit gute Ergebnisse nicht neben schlechten Ergebnissen verschwinden und keine Spender(innen) abgeschreckt würden.

Gregor Jungheim vom unabhängigen Magazin DIE STIFTUNG fragte sich in seinem Vortrag, welche Geschichten ein(e) Redakteur(in) braucht, um von Evaluationen zu erzählen. Anschließend führte er realitätsnahe Übungen mit den Teilnehmer(inne)n durch. Die Gruppe stellt fest, dass alle fachlich tief im Thema drin seien, sich jedoch nicht immer über ihre klare Botschaft bewusst seien. Bei journalistischen Anfragen würde daher oft eher die wissenschaftliche Repräsentativität und Methodengenauigkeit betont, die allerdings nicht an erster Stelle im öffentlichen Interesse stünden.

## Die Organisation von Wahrnehmung ist der Schlüssel für den Erfolg Veröffentlichung von Evaluierungsergebnissen

Frau Vahlhaus bestätigte Herrn Asendorpf in dem Medien-Leitspruch ‚*bad news are good news*‘. Skandale und Gerüchte ließen sich zwar verhindern, wenn die/der Betroffene schlechte Nachrichten selbst zum Thema mache, trotzdem sei ein gewisses Risiko vorhanden. Die GIZ veröffentliche auch schlechte Ergebnisse und setze diesen belastbare erfolgreiche Ergebnisse entgegen. Die GIZ stelle einerseits standardisierte Kurzberichte aller Evaluierungen im BMZ-Geschäft ins Netz. Andererseits würden daraus gezogene und journalistisch aufbereitete Konsequenzen in Form von Broschüren und Querschnittsauswertungen veröffentlicht. Ziele der Veröffentlichung seien die Legitimation und Transparenz der Arbeit der GIZ, die Vertretung der GIZ als lernende Organisation, die Herstellung von Glaubwürdigkeit und auch die Darstellung komplexer Zusammenhänge. Zielgruppen seien Auftraggebende und Partner, interne Verantwortliche, die fachliche und die externe Öffentlichkeit.

Als zentrale Herausforderungen erwähnte Frau Vahlhaus einerseits Sprache und Länge der Berichte. Ein weiteres Dilemma ergebe sich aus den vorgeschriebenen Evaluierungskriterien des Development Assistance Committees (DAC) der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD): Die Relevanz von Vorhaben werde oft gut bewertet, aber als banal wahrgenommen. Effizienz werde immer wichtiger, bereite aber Schwierigkeiten in der Messung und stehe manchmal im Widerspruch zur Nachhaltigkeit (zwar effizient, aber ohne nachhaltige Wirkung). Die Nachhaltigkeit wiederum werde von der Öffentlichkeit als das Wichtigste wahrgenommen, werde aber mit dem Impact oft am schlechtesten bewertet.

Die seit Jahren bleibende Durchschnittsnote 2 bis 3 wirke zudem wenig glaubwürdig. Die Last bzw. Lust an der Veröffentlichung könne aber auch eine Chance zur Schärfung der Methodik darstellen. Eine Alternative zu den DAC-Kriterien könnten unternehmensstrategische Evaluierungen zu bestimmten relevanten Fragen sein. Als Erfolgsfaktoren für erfolgreiche Veröffentlichung erwähnte Frau Vahlhaus die wissenschaftliche Begleitung als Türöffner zu bestimmten Veranstaltungen, die Einbindung der Unternehmenskommunikation in den Evaluierungsprozess und die Vorstellung von Ergebnissen auf zielgruppenorientierten Veranstaltungen.

### FZ-Evaluierung: Inhalte faszinierend farbig – Vermarktung grau in grau?

Frau Terberger stellte, unter anderem am Beispiel der Demarkierung von Indianergebieten im brasilianischen Amazonasgebiet, die spannenden und außergewöhnlichen Inhalte vor, die Evaluierungen aufweisen. Trotz der in den Augen der Evaluierenden faszinierend farbigen Inhalte gebe es aber nur ein geringes Interesse der (Fach-)Öffentlichkeit. Die KfW hätte als Reaktion darauf ein neues Berichtsdesign entworfen, lege mehr Wert auf eine verständliche Sprache und nachvollziehbare Überlegungen sowie die zielgruppengerechte Aufbereitung der Ergebnisse. Institutionelles Lernen und FZ-interne Adressat(inn)en müssten anders angesprochen werden als die Wissenschaft oder die breite Öffentlichkeit. Die KfW entsende außerdem operative Projektmanager(innen) auf Evaluierungsmissionen, um ihnen die „passende Brille“ aufzusetzen. Deren Erfahrungen könnten wiederum in das eigene Projektmanagement eingebracht werden. In abhängigen Medien wie dem KfW-Magazin sei eine Aufbereitung sehr zeitaufwändig, die Endredaktion könne die KfW aber mitbestimmen. In den freien Medien hingegen sei nicht steuerbar, wie und ob Bericht erstattet würde. Abschließend sollten Evaluierungen vielleicht auch nicht *zu* wichtig genommen werden. Es dürfe nicht vergessen werden, dass Entwicklungspolitik nicht nur von Evaluationsergebnissen, sondern auch von gegebenen Rahmenbedingungen, politischen Prioritäten, globalen Agenden und den Prioritäten der Partnerländer beeinflusst wird. Daher sollten sich die Anwesenden mit ihrer spannenden, farbigen Aufgabe zufrieden zeigen und mehr Bescheidenheit in der Reichweite der Ergebnisse walten lassen.

### Abschlussdiskussion

Zum Abschluss des inhaltlichen Teils der Tagung fand im Plenum eine Fishbowl<sup>4</sup>-Diskussion statt. Dazu wurden die aufgetauchten Fragestellungen der letzten zwei Tage wieder aufgegriffen.

---

4 Bei der *Fishbowl*-Methode werden Publikumsbeiträge systematisch in Podien einbezogen. Vor oder innerhalb des Publikums wird ein kleiner Stuhlkreis errichtet. Möchte jemand etwas sagen, nimmt sie/er entweder auf einem der freien Stühle Platz oder löst eine vorangegangene Sprecherin/einen vorangegangenen Sprecher auf einem anderen Stuhl ab.

Zentral war einerseits, wie sehr Evaluierungsergebnisse in die Öffentlichkeit gebracht werden sollten. Politische Stiftungen beispielsweise seien keine Spendenwerber und agierten darüber hinaus teils in stark politisch autoritären Umfeldern, in denen sich kurzfristig nur schwer positive Ergebnisse darstellen lassen. Auch passten die Wahlperiode, die eine bestimmte Agenda setze, und der Wirkungszeitraum nicht immer zusammen. Die Ergebnisse müssten vor allem intern verbreitet werden, was sich oft schon schwierig gestalte. Es ginge daher um einen maßgeschneiderten Austausch, dem immer die Frage nach der zu platzierenden Botschaft vorausginge. Vorsicht sei allerdings geboten, wenn man sich auf einzelne Abschnitte konzentriere, um die Komplexität zu reduzieren. Andere wichtige Aspekte und der Gesamtzusammenhang gingen leicht verloren. Eventuell sei ein Fokus auf die Holschuld der Interessierten zu legen, anstatt Ergebnisse unaufgefordert aufzubereiten. Freie Kapazitäten könnten sich im Gegenzug auf den Ausbau des Austauschs organisationsintern und unter den Evaluator(inn)en konzentrieren. Offen blieb die Frage, inwieweit der Anspruch, dass Evaluierungsergebnisse das globale Agenda Setting und die globale Policygestaltung beeinflussen können, gehalten werden könne. Sollte nicht besser betrachtet werden, wie das Lernen im eigenen System organisiert wird? Im heutigen globalen Setting stelle sich die Frage, ob beispielsweise im Zuge des Diskurswechsels von den *Millennium Development Goals* zu den *Sustainable Development Goals* Ergebnisse und Schlussfolgerungen mitgenommen worden wären. Außerdem sei zu beantworten, wie das Spannungsfeld zwischen zeitnaher Ergebnislieferung und dem Genügen wissenschaftlicher und methodischer Ansprüche aufgelöst werden könne.

## Literatur

- Palenberg, Markus A. (2011): Tools and Methods for Evaluating the Efficiency of Development Interventions. Evaluation Working Papers. Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung. Verfügbar unter: [http://www.bmz.de/en/zentrales\\_downloadarchiv/erfolg/BMZ\\_WP\\_Tools\\_Methods\\_Evaluating\\_Efficiency.pdf](http://www.bmz.de/en/zentrales_downloadarchiv/erfolg/BMZ_WP_Tools_Methods_Evaluating_Efficiency.pdf).