

Tagungsbericht zum Transdisziplinären Kolloquium „Zwischen Objektivität und Normativität: Wie kommt die Evaluation zu(m) Werten?“

Dörte Schott,¹ Anja Lietzmann,² Sandra Schopper,³
Susanne Zuschlag,⁴ Jan Hense⁵

Eröffnung und Einstiegsreferate

Als lokaler Ausrichter der Veranstaltung begrüßte Jan Hense die Teilnehmenden am Freitag, den 27. Februar 2015, an der Justus-Liebig-Universität Gießen und ging in seinem einführenden Vortrag auf die Bedeutung von Werten in der Evaluation sowie verschiedene Problematiken im Zusammenhang mit dem Bewerten ein. Er zeigte anhand von Beispielen aus konkreten Evaluationsberichten auf, dass obwohl Werte und das Werten zumindest dem Wortsinn nach elementare Bestandteile der Evaluation sein müssten, Evaluationen sich oft auf deskriptive Befunde beschränken, den Bezugsrahmen für Bewertungen im Dunkeln lassen oder die eigentliche Bewertung an Dritte delegieren. Im Ausblick auf die Tagung warf er die Frage auf, ob dies wirklich Defizite sind – oder legitime Praxis in bestimmten Evaluationskontexten.

Einstiegsreferat 1: Verena Friedrich (Universität Bern) legte in ihrem Beitrag „Die Bedeutung der Evaluationsfragestellung für das Bewerten“ dar, dass die Bewertung (Was ist gut/schlecht?) ebenso wie die Operationalisierung bereits in der Evaluationsfragestellung angelegt sein sollte, da auf dieser Grundlage die Kriterien und Indikatoren entwickelt werden, welche für die Messung und Bewertung benötigt werden. Sie veranschaulichte dies anhand eines konkreten Beispiels einer Fragestellung zur Veränderung der Luftqualität in öffentlichen Räumen nach Einführung des Nichtrauchergesetzes.

1 Freie Evaluatorin, Bonn

2 k.o.s GmbH, Berlin

3 Universität des Saarlandes und Hochschule für Technik und Wirtschaft des Saarlandes, Saarbrücken

4 Bildungs- und Beschäftigungsberatung, Rosenheim

5 Justus-Liebig-Universität, Gießen

Einstiegsreferat 2: Christine Heuer (Bundesamt für Gesundheit, Bern) bezog sich auf den Input ihrer Vorrednerin, als sie auf Basis ihrer empirischen Masterarbeit über die „*Verwendung von Bewertungskriterien in der Praxis*“ feststellte, dass das Thema „Bewertungskriterien“ in der Evaluationspraxis vernachlässigt wird. Daraus ergäben sich Folgerungen für Auftraggebende, Evaluierende und für die Aus- und Weiterbildung. So sollten Auftraggebende Bewertungskriterien v.a. bei Ausschreibungen oder beim Formulieren von Fragestellungen konkret thematisieren. Evaluierende sollten Bewertungskriterien im gesamten Evaluationsprozess vom Abgeben einer Offerte über die Datenerhebung und -auswertung bis hin zur Berichterstattung berücksichtigen. Schließlich sollten sich Bewertungskriterien und der Umgang mit ihnen als „roter Faden“ durch die gesamte Ausbildung von Evaluierenden ziehen.

Im Anschluss an die beiden Einstiegsreferate wurden viele interessierte Nachfragen gestellt und rege diskutiert. Ein Fazit aus der Diskussion war, dass sich das Thema Bewertungskriterien durch die Aus- und Weiterbildung in der Evaluation ziehen sollte.

Fallberichte aus der beruflichen Bildung

Fallbericht 1: Miguel del Estal (InterVal GmbH, Berlin) forderte in seinem Beitrag „*Die Vergleichsperspektive als Triebfeder einer Bewertungskultur*“ einen „Fahrplan zu einer Bewertungskultur“. Das Potenzial von Evaluationen werde ohne s.m.a.r.t.e. Ziele für den Binnenvergleich oder eine Referenzbasis für einen Außenvergleich aus seiner Sicht nicht ausgeschöpft. In der Diskussion wurde unter anderem infrage gestellt, ob Bewertungskriterien in jedem Fall politisch machbar seien: Geben Gesetze beispielsweise eine ausreichende Grundlage dafür her?

Fallbericht 2: Veronika Alz (MedienCampus Bayern, München) beschrieb in ihrem Beitrag „*Das Qualitätssiegel des MedienCampus Bayern als Beispiel zur Evaluation von Bildungsangeboten*“ konkrete Schwierigkeiten bei der Bewertung im Rahmen des Qualitätssiegels, welche beispielsweise durch die hohe Komplexität oder Heterogenität der Evaluationsgegenstände, in Ermangelung einer Referenzbasis für den Binnen- oder Außenvergleich o.ä. hervorgerufen werden sowie den Umgang mit dieser Problematik beim MedienCampus Bayern.

Parallelsession A: Wer bewertet letztendlich? Evaluation oder Rezipient(inn)en?

Dörte Schott führte als Moderatorin die Positionen der beiden folgenden Beiträge ein, in denen es darum geht, wer – Evaluierende oder Rezipient(inn)en – an welchen Stellen Bewertungen vornimmt oder Werte vorgibt. Während Felix Albrecht in seinem Beitrag aufzeigte, wie er die Rezipient(inn)en anstatt von Evaluierenden der Lehrveranstaltungsevaluation die Bewertung vornehmen lässt, legte Christian Diller dar, dass im Feld der regionalen Siedlungsentwicklung in den USA die Evaluation am Ende eine Bewertungslücke füllt, welche von der Politik offen gelassen wird.

Beitrag 1: Felix Albrecht (Karlsruher Institut für Technologie) stellte in seinem Vortrag „*Evaluation einer Lehrveranstaltung als Bewertungsgrundlage für verschiedene Interessensgruppen*“ die Evaluation einer Lehrveranstaltung am KIT vor, die zur Semestermitte hin getätigt wird. Ziel dieser Lehrevaluation ist, noch während des Semesters gegensteuern zu können, falls sich Problemstellungen ergeben sollten. Als nutzenorientierte Evaluation soll sie allen Beteiligten bzw. Interessent(innen) eine verwertbare Rückmeldung zur Lehrveranstaltung geben können. Da Hinweise auf Verbesserungsmöglichkeiten in den klassischen Standardverfahren zur Lehrevaluation aus seiner Sicht fehlen, schlug er einen qualitativen Lösungsansatz vor: Auf der Grundlage eines Leitfadens führen Hilfskräfte (keine in den Lehrbetrieb involvierte Personen) Gruppendiskussionen mit den Studierenden der jeweiligen Lehrveranstaltung zur Bewertung der Ergebnisse durch, welche anschließend inhaltsanalytisch ausgewertet werden. Hieraus würden sich ganz konkrete Verbesserungsideen ergeben, welche anschließend von den Dozent(inn)en zur Kenntnis genommen und umgesetzt werden können. Die Evaluation kann von den Dozent(inn)en selbst angeregt und selbstevaluativ durchgeführt werden.

Beitrag 2: Christian Diller (Universität Gießen) gab zunächst einen Ein- und Überblick über die Raum- und Siedlungsplanung in Deutschland und den USA. Sein Beitrag „*Bewertung in der Evaluation regionaler Siedlungsentwicklung: Ein Vergleich Deutschland – USA*“ beschäftigte sich mit dem antiproportionalen Zusammenhang zwischen Planungsichte und Evaluationstätigkeit in beiden Ländern. In Deutschland würde die Stadt- und Raumplanung aufgrund der dichten Besiedlung im Gegensatz zu den USA viel stärker gesteuert; gleichzeitig gäbe es kaum Evaluationen derselben. Anders ist es in den USA, wo die Stadt- und Raumplanung in einigen Staaten kaum gesteuert wird, aber zugleich eine große Anzahl von Evaluationen zu diesem Thema vorliege, welche meist von Hochschulen durchgeführt würden. Die Evaluationskriterien würden hier im Rahmen von Evaluationen und durch die Evaluierenden (unter Einbezug der betroffenen Bevölkerung) entwickelt. Er berichtete in diesem Rahmen außerdem über den Versuch, einen Satz von Standardkriterien für die Stadt- und Raumplanung in Deutschland zu entwerfen, welcher in der Fachcommunity als Qualitätsstandard diskutiert würde.

Den beiden Beiträgen der Session A schloss sich eine rege Frage- und Diskussionsrunde an, in welcher es abschließend um die Frage ging, ob es Aufgabe der Evaluation sei, Lücken in den Bewertungskriterien der Politik zu schließen.

Parallelsession B: Wessen Werte? Werte(n) in sozialen und komplexen Kontexten

Beitrag 1: Dirk Groß (Agentur für Qualitätssicherung, Evaluation und Selbstständigkeit von Schulen, Bad Kreuznach) stellte in seinem Beitrag „*Der Einsatz der FIWE-Methode im Kontext der externen Schulevaluation*“ ein Beispiel externer Evaluation in staatlichen Schulen in Rheinland-Pfalz vor. Das hierbei zum Einsatz kommende „wertepriorisierende Evaluationsmodell“ (vgl. Beywl/Speer/Kehr 2004) wurde darauf geprüft, inwiefern die Werte der Beteiligten und Betroffenen im Sinne

des DeGEval-Standards „N5 Transparenz von Werten“ sinnvoll identifiziert und thematisiert werden können. Exemplarisch wurde dies aufgezeigt am Einsatz der FIWE-Methode (Fakten-Interpretation-Werte-Empfehlungen; vgl. Beywl/Schepp-Winter 2000).

Beitrag 2: Maria Gutknecht-Gmeiner (Impulse – Evaluation und Organisationsberatung, Wien) ging in ihrem Beitrag „*Developmental Evaluation: Was bedeutet die Evaluation von Innovationen in komplexen, dynamischen Umfeldern für das Bewerten?*“ auf Aspekte der „Developmental Evaluation“ (dt. „evolutive Evaluation“, vgl. Patton 2010) ein – ein spezifisches Vorgehen für die Evaluation von innovativen, sich ständig anpassenden Maßnahmen unter veränderlichen Rahmenbedingungen. Im Zentrum ging sie der Frage nach, welchen Einfluss ein „evolutives“, entwicklungsorientiertes Vorgehen in der Evaluation auf das (Be)Werten hat. Anhand des Projektes „Technik zu den Mädchen bringen“ stellte sie dar, in welchen Bereichen die Prämissen für das Bewerten sowie der Bewertungsvorgang in der Developmental Evaluation sich von gängigen Vorstellungen vom Bewerten in Evaluationen unterscheiden. Dabei ging sie auf typische Aspekte der Developmental Evaluation ein, wie die Instabilität und der geringe Ausdifferenzierungsgrad des evaluierten Programms, das Ziel der Entwicklungsunterstützung, die direkte Handlungsrelevanz, die intensive Zusammenarbeit mit den Programmumsetzenden sowie insgesamt die enge Verflechtung von Evaluation und Intervention.

Keynote: Wie kommt die Evaluation zu(m) Werten?

In seiner abschließenden Keynote ging Wolfgang Beywl (Pädagogische Hochschule FHNW Brugg/Univention Köln) noch einmal grundsätzlich auf die Frage des Bewertens ein. Er stellte die These auf, dass eine Evaluation bewerten muss – und zwar systematisch. Michael Scriven zitierend gab er zudem eine Antwort auf die Frage, wer bewerten muss: die Evaluierenden.

Grundsätzlich unterschied er drei unterschiedliche Bedeutungen des Wortes „Wert“: (a) Güte, Tauglichkeit, Bedeutsamkeit (worth/merit), (b) sozialer Wert im normativen oder moralischen Sinne, (c) statistischer Wert. Die Bestimmung des Wertes eines Evaluationsgegenstandes (a), d.h. die Bewertung, sei zentrale und unverzichtbare Aufgabe jeder Evaluation. Kriterien seien dabei die Bezugsmaßstäbe für die Bewertung, und damit der Dreh- und Angelpunkt der Bewertung. Für Bewertungsakte in Evaluationen seien in der Regel operationalisierte Kriterien erforderlich. Als vier generische Kriteriendimensionen für Evaluationen, das heißt für die Bewertung einer Evaluation (Metaevaluation), stellte er die aus den DeGEval-Standards bekannten Dimensionen Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit vor.

Zum Umgang mit Werten im Vergleich zwischen Forschung einerseits und Evaluation andererseits stellte er die These auf, dass die Evaluation der Sozialforschung aktuell „um eine Nasenlänge voraus“ sei, da sie sich systematisch mit der Bewertungsfrage auseinandersetze anstatt sie zu vermeiden. Schließlich ging er noch auf generische Kriteriendimensionen für Gegenstände der Bildungsevaluation und Quel-

len für wertebasierte Kriterien in Gleichstellungsevaluationen ein. Er schloss seinen Vortrag ab mit der Beantwortung zweier Fragen, welche aus den Reihen des Nachwuchsnetzwerkes im Rahmen des Kolloquiums am Vortrag an ihn gerichtet worden waren.

Zum Abschluss des Transdisziplinären Kolloquiums dankte Jan Hense im Namen der Gastgebenden allen Anwesenden für die rege Beteiligung. Er würde sich auf weitere Diskussionen freuen, die im Rahmen der Tagung nur angestoßen, jedoch nicht abgeschlossen werden konnten.

Im Rahmen des Transdisziplinären Kolloquiums ist die Frage des Bewertens aus ganz unterschiedlichen Blickwinkeln und vor verschiedenen Hintergründen beleuchtet worden. Einigkeit bestand gleichwohl darin, dass die Aufgabe der Bewertung in der Evaluation eine zentrale ist – und zwar nicht nur aufgrund der Begrifflichkeit (evaluate = bewerten), sondern weil sie ein besonderes Potenzial im Hinblick auf die Verwendung der Evaluationsergebnisse entfalten kann. Deutlich wurde währenddessen, dass Bewertungen eine gewisse Grundlage brauchen, auf der sie fußen: transparente Werte, klare und operationalisierte Kriterien – aber auch, dass diese von den Auftraggebenden (z.B. von politischen Entscheidungsträger(inne)n) nicht immer formuliert oder vorgegeben werden können. Hier ist es an der Evaluation, diese Grundlage systematisch und unter Einbezug der relevanten Stakeholder bzw. Interessensgruppen herzuleiten. Wie aus der Praxis berichtet wurde, können dabei verschiedene Schwierigkeiten auftreten, die beispielsweise durch eine hohe Komplexität oder Heterogenität der Evaluationsgegenstände oder in Ermangelung einer Referenzbasis für den Binnen- oder Außenvergleich entstehen.

Alles in allem fand auf der Tagung ein sehr vielseitiger und fruchtbarer Austausch statt, der den Beginn einer konstruktiven Debatte über die Bedingungen und Methoden sowie Möglichkeiten und Grenzen der Bewertung in der Evaluation markieren könnte.

Literatur

- Beywl, Wolfgang/Schepp-Winter, Ellen (2000): Zielgeführte Evaluation von Programmen. Ein Leitfaden. In: Bundesministerium für Familie, Senioren, Frauen und Jugend (Hg.): Materialien zur Qualitätssicherung in der Kinder- und Jugendhilfe, 29.
- Beywl, Wolfgang/Speer, Sandra/Kehr, Jochen (2004): Wirkungsorientierte Evaluation im Rahmen der Armuts- und Reichtumsberichterstattung. Perspektivstudie. Bonn: Bundesministerium für Gesundheit und Soziale Sicherung (BMGS). Verfügbar unter: http://www.univation.org/download/Evaluation_der_Armuts-_und_Reichtumsberichterstattung.pdf.
- Patton, Michael Q. (2010): Developmental Evaluation. Applying Complexity Concepts to Enhance Innovation and Use. New York: Guilford Press.

„Zwischen Nutzung, Einfluss und Nachhaltigkeit – Wie wirken Evaluationen in unterschiedlichen Systemen?“

**19. Jahrestagung der DeGEval –
Gesellschaft für Evaluation e.V.
vom 21. bis 23. September 2016 in Salzburg**

Call for Proposals

Die langfristige Wirkung von Evaluationen in unterschiedlichen sozialen Systemen ist Gegenstand einer intensiven Debatte, die wir anlässlich der 19. Jahrestagung der Gesellschaft für Evaluation in Salzburg aufnehmen wollen. Die Bandbreite erstreckt sich dabei von Fragen zu den Voraussetzungen und Bedingungen für eine gewinnbringende Nutzung von Evaluation, also die Verwendung oder den Gebrauch von Evaluation durch die Stakeholder in einem geplanten Sinne, bis hin zu direkten und indirekten Einflüssen und Wirkungen von Evaluationen. Mit dem Begriff der Nachhaltigkeit wird verbunden, dass durch die Nutzung von Evaluationen und ihren Ergebnissen eine langfristig positive Wirkung entfaltet wird. Dabei sollen gesellschaftliche, ökologische und ökonomische Aspekte gleichermaßen Berücksichtigung finden.

Die Jahrestagung 2016 will Nutzung, Einfluss, Wirkung und Nachhaltigkeit von Evaluation mit übergreifenden Fragestellungen in den unterschiedlichen Systemen Bildung, Politik, Kultur, Gesundheitswesen, Wirtschaft und Verwaltung sowie aus unterschiedlichen Perspektiven beleuchten. Folgende Fragestellungen öffnen das Feld:

- Wie kann eine intendierte, an den Zwecken der Evaluation orientierte Nutzung von Evaluation erreicht werden? Wie greifen Nutzung des Evaluationsprozesses und der Evaluationsergebnisse ineinander? Wie lässt sich die Nachhaltigkeit von Evaluation bestimmen? Wie hängen Nutzung und Nachhaltigkeit von Evaluation zusammen? Braucht es ‚windows of opportunity‘ oder existieren weitere Gestaltungsspielräume?
- Wann und unter welchen Bedingungen wird aus Nutzung Nützlichkeit? Hat die zeitliche Nutzung von Evaluation – vor der Evaluation, währenddessen, kurz danach, lange nach der Evaluation – einen Einfluss auf die Nachhaltigkeit von Evaluation?
- Welche Rollen kommen den Akteuren im ‚tricky rectangle‘, den Auftraggebenden, den Evaluierenden, den Vertreter(inne)n der Nutznießenden und Bürger(innen) sowie den politischen Entscheidungsträger(inne)n zu, wenn Nutzung und Nachhaltigkeit von Evaluation erreicht werden sollen?
- Im Jahr 2015 werden von den Vereinten Nationen die Ziele für nachhaltige Entwicklung verabschiedet (Sustainable Development Goals). Wie lässt sich bestimmen, ob diese Ziele erreicht werden? Sind die verfügbaren Monitoring- und Evaluationssysteme hierfür bereits geeignet?
- Welche Rolle spielt Evaluation für eine nachhaltige Entwicklung? Welchen Einfluss hat Evaluation im demokratischen Prozess? Wie kann eine lediglich vordergründige Nutzung von Evaluation vermieden werden, die über die Interessen der Auftraggebenden hinaus ein gesellschaftliches, soziales, ökonomisches Wohl im Auge hat?

- Inwiefern und wie sind nachhaltige Wirkungen in Evaluationen zu erfassen? Was sind Indikatoren und messbare Bedingungen von Nachhaltigkeit? Wie ist Nachhaltigkeit im inzwischen alltagsprachlichen Sinne zu operationalisieren im Gegensatz zum ursprünglichen Nachhaltigkeitsbegriff?
- Kann man in Bezug auf die Fragestellungen von ähnlichen Prozessmustern und Mechanismen in den unterschiedlichen Systemen ausgehen oder gibt es deutliche Unterschiede? Welchen Mehrwert haben Erkenntnisse über Ähnlichkeiten bzw. systembedingte Unterschiede der Nutzung und Wirkungen von Evaluationen für das jeweils eigene Forschungs- und Handlungsfeld?

Wir laden Sie herzlich ein, an der Diskussion in Salzburg teilzunehmen und Ihre Arbeit auf der Jahrestagung vorzustellen. Eingeladen sind Beiträge, die sich mit den oben genannten Fragen wissenschaftlich, reflexiv oder praktisch auseinandersetzen. Die Beitragseinreichung wird ab dem 1. November 2015 über die Konferenzseite der DeGEval (<https://www.conftool.com/degeval2016/>) möglich sein. Die Tagung bietet folgende Beitragsmöglichkeiten:

- **Blitzvortrag:** In einem Blitzvortrag wird ein Forschungs- oder Praxisprojekt mit seinen innovativen Aspekten oder neuartigen Fragestellungen auf möglichst allgemein verständliche Weise dargestellt. Eine Blitzvortragssession besteht aus zehn fünfminütigen Blitzvorträgen, an die sich je drei Minuten Fragen anschließen. Die Möglichkeit des Blitzvortrags richtet sich insbesondere an Studierende, Doktorand(inn)en und Berufseinsteiger(innen), aber auch an Evaluationspraktiker(innen), welche ihre Arbeit schlaglichtartig einem möglichst breiten Publikum präsentieren möchten.
- **Einzelvortrag:** In Einzelvorträgen wird das Thema Evaluation entweder praktisch, reflexiv oder wissenschaftlich betrachtet: (a) Evaluationen als Projekte mit praktischen Herausforderungen (Planung, Steuerung, Durchführung etc.), (b) Evaluationen als Grundlage zur Reflexion und zum Generieren von (Meta-)Wissen über Evaluation (methodische, theoretische, methodologische Aspekte) oder (c) Forschungsergebnisse über Evaluation (Evaluation als Forschungsgegenstand). Für den Vortrag stehen 20 Minuten und 10 Minuten Diskussionszeit zur Verfügung. Einzelvorträge werden zu thematischen Sessions zusammengestellt (max. drei Vorträge pro Session).
- **Komplette Session:** Komplette Sessions (90 Minuten) können als Vortragssession, Podiumsdiskussion, Workshop oder in anderer Form konzipiert sein. Abstracts umfassen die Beschreibung der geplanten Session sowie aller beteiligter Beiträge und Personen. Bei einer Vortragssession sind kurze Inhaltsangaben der einzelnen Vorträge mit einzureichen sowie die/der Moderator(in) und ggf. Diskutant(inn)en zu benennen. Bitte reichen Sie diese dann gesammelt in einem Dokument ein. Komplette Sessions können in deutscher und englischer Sprache stattfinden.
- **Poster:** Poster präsentieren entweder innovative Evaluationsvorhaben oder Forschungsarbeiten zu Fragestellungen, welche empirisch, konzeptionell, theoretisch oder methodisch behandelt werden. Es wird eine Posterführung bzw. eine moderierte Postersession stattfinden.

Bitte reichen Sie ihr Abstract mit max. 2200 Zeichen inklusive Leerzeichen (für komplette Sessions mit max. 4000 Zeichen) bis zum **11. März 2016** über unsere Konferenzhomepage (<https://www.conftool.com/degeval2016/>) ein und beachten Sie bitte die dort angeführten Hinweise zur Einreichung. Alle Beiträge werden in einem Review-Verfahren begutachtet. Sie werden bis zum 29. Mai 2016 über die Beitragsannahme informiert. Nähere Informationen zur Tagung finden Sie unter www.degeval.de/veranstaltungen/jahrestagungen/salzburg-2016/.

Kontakt:

DeGEval – Gesellschaft für Evaluation e.V.
 Frau Natalie Salf
 Johannes Gutenberg-Universität Mainz
 c/o Zentrum für Qualitätssicherung und
 -entwicklung (ZQ)
 Colonel-Kleinmann-Weg 2, SBII, 04-543
 D-55099 Mainz

Tel.: +49 (0) 6131 / 39-2 68 69
 Fax: +49 (0) 6131 / 39-2 68 68
 E-Mail: info@degeval.de
<http://www.degeval.de>

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,¹ Jens Koy,² Susanne Neubert,² Susanne von Jan²

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-Kriterium³ 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 Fishbowl⁴-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.

Methodische Herausforderungen der Wirkungsanalyse bei knappen Ressourcen

Frühjahrstagung 2015 des AK Soziale Dienstleistungen

*Stefanie Reiter,¹ Stefan Schmidt,² Rainer Strobl,³ Florentina
Astleithner,⁴ Benjamin Froncek,⁵ Peter Stepanek⁴*

Tagungsthema und Format

Wirkungsanalysen sind eine zentrale Evaluationsaufgabe, aber auch eine große methodische Herausforderung, weil die Effekte oft klein sind, unterschiedliche Ursachen haben können und häufig nicht eindeutig auf bestimmte Maßnahmen zurückzuführen sind. Die Schwierigkeiten nehmen drastisch zu, wenn die zeitlichen und materiellen Ressourcen für die Planung und Entwicklung der Instrumente sowie für die Erhebung, Datenanalyse und Ergebnisaufbereitung knapp sind. Auf Evaluationen trifft das fast immer zu. Was können Evaluatorinnen und Evaluatoren also tun, wenn es keine passende Skala für ein bestimmtes Phänomen gibt und eine aufwändige Skalenentwicklung mitunter aus Ressourcengründen nicht infrage kommt? Oder wie soll man damit umgehen, wenn eine Randomisierung nicht machbar ist und der Datenschutz die Erhebung wichtiger Kontrollvariablen in einem quasi-experimentellen Design unmöglich macht? Wie kann man Ergebnisse aus standardisiert erhobenen Daten z.B. durch die Erhebung zusätzlicher qualitativer Daten absichern? Kann man mit qualitativen Interviews, Fokusgruppen und Gruppendiskussionen die Wirkungsfrage beantworten? Und wie lässt sich der Aufwand dabei in Grenzen halten? Geht das dann zu Lasten der methodischen Standards (oder auch der Evaluationsstandards)? Bis zu welchem Grad können oder müssen diese Standards bei Evaluationen aus dem Bereich der sozialen Dienstleistungen eingehalten werden?

Die Frühjahrstagung des Arbeitskreises Soziale Dienstleistungen der DeGEval am 24. Mai 2015 bot Evaluatorinnen und Evaluatoren eine Gelegenheit, diese

1 Deutsches Jugendinstitut e.V., Halle an der Saale

2 schmidt evaluation, Köln

3 proVal, Gesellschaft für Sozialwissenschaftliche Analyse – Beratung – Evaluation, Hannover

4 FH Campus Wien

5 FernUniversität in Hagen

und ähnliche Fragen zu diskutieren und gemeinsam über Lösungen nachzudenken.⁶ Rund 30 Personen aus Deutschland und Österreich nahmen im Internationalen Jahr der Evaluation an der Frühjahrstagung in Wolfsburg teil.

Berichte aus der Evaluationspraxis zu Herausforderungen von Wirkungsanalysen

Den Einstieg der Tagung bildeten zwei *Impulsreferate*, die Herausforderungen und Probleme bei der Wirkungsanalyse anhand von konkreten Beispielen aus der Evaluationspraxis darstellten.

Im ersten Vortrag „Hat das Training Ihren Erwartungen entsprochen?“. Herausforderungen und Probleme bei der Wirkungsanalyse im Kontext von Diversity Trainings“ (Autor(inn)en: Benjamin Froncek, Verena Piper, Dr. Agostino Mazziotta und Prof. Dr. Anette Rohmann, FernUniversität in Hagen) wurden Herausforderungen bei der Erfassung von Wirkungen im Kontext von Maßnahmen zur Förderung interkultureller Kompetenzen betrachtet. Benjamin Froncek stellte vor, dass evidenzbasierte Praxis auch im Rahmen von Diversity Trainings zusehends gefordert wird.

Neben wenigen veröffentlichten Evaluationsstudien, die sich einzelnen Fragestellungen von Diversity Trainings widmen, blieb bislang jedoch die Frage offen, wie sich die derzeitige Evaluationspraxis im Feld von Diversity Trainings konkret ausgestaltet. Um diese Frage beantworten zu können, wurden Trainerinnen und Trainer von Maßnahmen zur Förderung interkultureller Kompetenzen (N=172) im deutschsprachigen Raum in einem Online-Survey zu ihrer Evaluationspraxis befragt. Dabei zeigte sich, dass Trainerinnen und Trainer der Wirksamkeitsanalyse bei der Evaluation von Trainings durchaus eine hohe Bedeutung beimessen. Das Ausmaß des Einsatzes von Wirksamkeitsanalysen und die Art der methodischen Umsetzung spiegeln in der Praxis diese Bedeutung nicht immer entsprechend wider: Als Maß für den Trainingserfolg steht die Zufriedenheit der Trainingsteilnehmenden im Fokus der Evaluationen. Die Trainings werden überwiegend selbst evaluiert. Es kommen meist schriftliche Befragungen mit Selbstaussagen und einem Messzeitpunkt – im Anschluss an ein Training – zum Einsatz. Ein solches Vorgehen ist ggf. nützlich für die Qualitätssicherung, aber wenig aussagekräftig für die Wirksamkeitsüberprüfung. Der Survey zeigte also eine sichtbare Diskrepanz auf zwischen dem Anspruch, Wirkungen nachweisen zu wollen auf der einen und der berichteten Evaluationspraxis auf der anderen Seite.

Als mögliche Begründung dieser Diskrepanz wurden Herausforderungen bei der Umsetzung von Wirksamkeitsanalysen auf verschiedenen Dimensionen diskutiert: Erschwerende Bedingungen für den Einsatz von Wirkungsanalysen können unter anderem durch *Kontextfaktoren*, wie die Art und die organisationale Einbindung des Trainings, aber auch durch zur Verfügung stehende Ressourcen entstehen. Vonseiten der *Auftraggebenden* beeinflusst möglicherweise das vorhandene Wissen über

6 Die Frühjahrstagung des AK Soziale Dienstleistungen wurde von Olaf Lobermeier und Rainer Strobl (proVal), Stefanie Reiter (DJI) und Stefan Schmidt (schmidt evaluation) organisiert. Die letztgenannten drei Personen bilden zugleich das Sprecher(innen)team des AK.

Wirksamkeitsanalysen, ob Wirksamkeitsnachweise tatsächlich eingefordert werden. Aufseiten der *Trainerinnen und Trainer* stellt sich diesbezüglich die Frage nach der Haltung gegenüber evidenzbasierter Praxis sowie nach den Kompetenzen, Wirksamkeitsanalysen beurteilen oder gegebenenfalls selbst durchführen zu können. In Hinblick auf die *Teilnehmenden* eines Trainings spielen Aspekte der zeitlichen, kognitiven und motivationalen Beanspruchung für die Durchführung von Wirksamkeitsanalysen eine Rolle. Insbesondere im Zusammenspiel der aufgeführten Dimensionen ergibt sich für Evaluatorinnen und Evaluatoren die Herausforderung, ein kontextsensibles Design zu finden, das trotz der gegebenen Umstände in der Lage ist, Wirkungsdimensionen angemessen zu operationalisieren und objektive multiperspektivische Instrumente und Maße zu verwenden, um (nachhaltige) Wirkungen auf das Training zurückführen zu können.

Im zweiten Impulsvortrag „Zur Produktivität knapper Ressourcen in der Verbindung von Lehre – Praxis – Forschung: Wirkungsanalysen auf den Weg bringen“ (Autor(inn)en: Florentina Astleithner und Peter Stepanek, FH Campus Wien) erläuterte Florentina Astleithner anhand von Beispielen aus der eigenen Evaluationspraxis, dass sich methodische und praktische Herausforderungen auch bereits vor der eigentlichen Durchführung von Wirkungsstudien stellen: Neben der Gestaltung des konkreten Forschungsdesigns rückt – auch aufgrund von steigenden Forderungen nach Wirkungsanalysen und Modellen der wirkungsorientierten Steuerung seitens verschiedener Stakeholder – die Frage der Finanzierung einer umfassenden Wirkungsstudie für Evaluierende sowie für Auftraggebende zusehends in den Fokus. Knappe Ressourcen tragen dazu bei, dass sich Organisationen als Träger von Programmen und Maßnahmen im Zuge der Vorbereitung von Wirkungsanalysen vermehrt mit diesem Anliegen auch an Universitäten und Fachhochschulen wenden. Die dadurch entstehende besondere Situation der Universitäten und Fachhochschulen, an der Schnittstelle zwischen Lehre und Forschung unter Einbindung von Studierenden und in enger Kooperation mit der Praxis Forschungsprojekte initiieren zu können, birgt innovative Ideen und hohe Nutzenpotenziale für alle Beteiligten (insbesondere Lehrende, Studierende und Praxiseinrichtungen) und wirft zugleich zusätzliche Herausforderungen auf, welche im vorgestellten Beitrag kritisch hinterfragt wurden: Unter anderem anhand eines Kooperationsprojekts zwischen einem Studiengang (Sozialwirtschaft und Soziale Arbeit) und einem Gemeinwesenprojekt in Wien wurde aufgezeigt, wie Masterarbeiten, begleitende Lehrveranstaltungen und Workshops zur Exploration des Forschungsfeldes als Vorbereitung einer Wirkungsstudie beitragen können. Basierend auf den vorliegenden Erfahrungen wurde diskutiert, inwieweit eine solche Kooperation tatsächlich die vielfältigen Ziele erreichen kann und z.B. sinnvoll sowie lehrreich für Studierende ist, das gewünschte Ergebnis für die Praxiseinrichtung liefert und aus Sicht der Forschungsverantwortlichen mit Blick auf fachliche Standards vertretbar ist. Außerdem wurde in diesem Kontext die Frage aufgeworfen, ob Kooperationen dieser Art eine Basis für die Durchführung eines umfassenden Auftragsforschungsprojektes bieten können bzw. wo die Grenzen (ökonomisch, methodisch, organisatorisch) liegen. In der bisherigen Umsetzung des Kooperationsprojekts erwies es sich als essenziell, Komplexität im Laufe des Forschungsprozesses aufzubauen, indem unter anderem unterschiedliche Perspekti-

ven berücksichtigt, das Design theoriegeleitet und methodisch durchdacht und vorab auch geklärt wurde, welche Folgen in der Praxiseinrichtung entstehen können, wenn Wirkungen systematisch in den Blick genommen werden. Zugleich war die schrittweise Reduktion von Komplexität (beispielsweise im Sinne einer Klärung methodischer und ressourcenbedingter Grenzen sowie der Abstimmung von Rollen und Forschungs-, Beratungs- und Lehrendenperspektiven) ein wesentlicher Erfolgsfaktor für einen für alle Beteiligten gewinnbringenden Prozess. Der Ansatz erwies sich nach den Erfahrungen von Florentina Astleithner und Peter Stepanek insgesamt als gut geeignet, um verschiedene Perspektiven ineinander zu verschränken sowie Zeit und Raum für Auseinandersetzungen zu schaffen, um den Boden für Wirkungsanalysen vorbereiten zu können.

In drei *Arbeitsgruppen* ging es im Anschluss an die Impulsvorträge darum, pragmatische und – aus rein wissenschaftlicher Sicht vielleicht suboptimale Lösungen – für die effektive Durchführung von Wirkungsanalysen unter den Bedingungen knapper Ressourcen, eines engen Zeitrahmens und einer möglichst geringen Störung der Praxis zu diskutieren. Ein wichtiger Aspekt war dabei die kaum diskutierte Frage, ob die gefundenen Lösungen – auch mit Blick auf die Standards der DeGEval – noch vertretbar sind oder ob sie die Qualität der Evaluationsergebnisse gefährden. In diesem Zusammenhang konzentrierte sich die AG 1 auf methodische Herausforderungen bei einem primär quantitativen Vorgehen, die AG 2 richtete ihr Augenmerk auf methodische Herausforderungen bei einem vorwiegend qualitativen Vorgehen⁷ und die AG 3 befasste sich mit methodischen Herausforderungen bei der Kombination von quantitativen und qualitativen Methoden. Den Arbeitsgruppen war es freigestellt, ob sie an die Eingangsreferate anknüpfen oder andere Probleme besprechen wollten. Die Ergebnisse der Arbeitsgruppen wurden anschließend im Plenum vorgestellt, zusammengeführt und diskutiert.

Ergebnisse der Arbeitsgruppen

In den Arbeitsgruppen wurde durchaus mit Bezug auf die Eingangsvorträge anhand von Evaluationsbeispielen aus dem Kreis der Teilnehmenden diskutiert, wie Wirkungen unter den oft schwierigen Rahmenbedingungen einer Evaluation seriös nachgewiesen werden können. Die eigentliche Schwierigkeit besteht nach den Ergebnissen der Diskussion darin, die Wirkungen auf die realisierten Maßnahmen zurückzuführen. Das Problem der knappen Ressourcen stellt sich dabei aus Sicht der Tagungsteilnehmenden oftmals sowohl im Vorfeld von Wirkungsuntersuchungen als auch in allen Phasen des eigentlichen Evaluationsprozesses. Die Teilnehmenden diskutierten verschiedene spezifische Probleme und Lösungswege in den einzelnen Phasen. Nachstehend sind die zentralen Ergebnisse wiedergegeben.

7 Nach Erörterung der von den Anwesenden praktizierten Erhebungsmethoden bzw. -quellen beschränkte AG 2 die Diskussion auf die aus der Erfahrung der Anwesenden in der Evaluation am häufigsten angewendeten qualitativen Methoden, auf Interview und Gruppendiskussion. Weitere Methoden, z.B. Beobachtung und Skype-Interview, wurden in der Diskussion nur am Rande besprochen.

Planungsphase:

Bei der Konzeption und Planung einer Evaluation werden unter anderem die grundlegenden Weichen für eine effektive Erhebung gestellt. Insbesondere bei qualitativen Erhebungen muss den befragten Personen ein Rahmen ermöglicht werden, in dem sie ungestört, ohne Zeitdruck und frei von sozialer Erwünschtheit ihrem Wissen und ihren Haltungen zu den jeweiligen Fragestellungen Ausdruck verleihen können. Wer bereits bei der Planung Ressourcen sparen möchte, läuft große Gefahr, keine bzw. kaum verwertbare Ergebnisse zu generieren. Gerade am Anfang können die Voraussetzungen für ein ressourcenarmes Arbeiten geschaffen werden. Als wesentliche „Stellschrauben“ wurden die Zielexplication mit den Auftraggebenden und die Wahl des Designs, aber auch die Planung der Erhebung genannt:

- *Klärung des Zwecks/Ziels der Erhebung, der Nutzenerwartungen und der Fragestellungen bzw. Hypothesen mit den Auftraggebenden:* Als Voraussetzung für eine gelingende Wirkungsanalyse wurden hinreichend konkrete Ziele betrachtet, die gemeinsam mit den Auftraggebenden festgelegt werden. Dabei sollte man diese – auch in Abwägung mit den zur Verfügung stehenden Ressourcen – auf das Wichtigste reduzieren, beispielsweise nur den Zweck „Optimierung“ verfolgen und nur zwei Fragestellungen/Hypothesen untersuchen. In dieser frühen Phase ist es besonders wichtig, zu prüfen, welches Design für welche Zielstellung angemessen ist und realisiert werden kann und welche Aussagen zur Wirksamkeit sich daraus ableiten lassen bzw. welche Grenzen ggf. bestehen. Die Auftraggebenden sollten bereits in dieser Phase dafür sensibilisiert werden, welche Schwierigkeiten darin bestehen, Wirkungen auf die realisierten Maßnahmen zurückzuführen, welche Aussagen sich bei welchem Design (und anderen Designalternativen) generieren lassen und welcher Wirkungsbegriff angemessen ist (Erwartungsmanagement). Kritisch wurde insbesondere die Tendenz einer weitverbreiteten, unzureichenden Unterscheidung zwischen Zufriedenheitsmessungen und Wirkungsnachweisen gesehen. Gerade im Bereich der Evaluation sozialer Dienstleistungen muss (selbst-)kritisch geprüft werden, ob der Nachweis „knallharter Fakten“ durch quasi-experimentelle Designs realisierbar und im Einzelfall gegenstandsangemessen ist. Dabei sollten sich Evaluatorinnen und Evaluatoren auch vor Augen halten, dass Auftraggebende oftmals wiederum Legitimationszwängen von außen unterliegen. Diese zweite (Erwartungs-) Ebene kann dadurch berücksichtigt werden, dass beispielsweise den Auftraggebenden Argumentationshilfen für den Umgang mit Anforderungen und Nachfragen von außen zur Verfügung gestellt und Machbarkeitsgrenzen offen kommuniziert werden.
- *Wahl des Designs:* In diesem Zusammenhang nahm, insbesondere in der Arbeitsgruppe zu quantitativen Verfahren, die Diskussion über die Bildung von Kontrollgruppen einen breiten Raum ein. Die Teilnehmenden waren sich einig, dass ein sogenanntes RCT-Design (eine randomisierte, kontrollierte Studie) im Bereich sozialer Dienstleistungen kaum umsetzbar ist. In diesem Bereich werde stattdessen häufig argumentiert, dass der Ausschluss der Kontrollgruppenmitglieder von der Teilnahme an einem Angebot ein ethisches Problem darstellt.

Als realisierbare Möglichkeit wurden in diesem Zusammenhang sogenannte Warte-Kontrollgruppen genannt. In diesem Fall nehmen die Mitglieder der Warte-Kontrollgruppe zu einem späteren Zeitpunkt an den Maßnahmen teil. Anspruchsvoller wäre ein sogenanntes Crossover-Design, bei dem zwei Gruppen an zwei Maßnahmen in unterschiedlicher Reihenfolge teilnahmen. Wenn es gelingt, Zeitreihen zu erfassen, dann kann einem Diskussionsbeitrag zufolge unter Umständen auch ganz auf eine Kontrollgruppe verzichtet werden. Unter Umständen sei es auch möglich, eine ausgearbeitete Programmtheorie, beispielsweise mit einem Strukturgleichungsmodell, zu überprüfen. Allerdings würden hierfür dann wieder erhebliche Ressourcen benötigt.

- *Planung der Erhebung:* Eine sorgfältige Planung einer Erhebungssituation (beispielsweise eine gelungene Terminierung von telefonischen Interviews und Vereinbarung des benötigten Zeitfensters, kriteriengeleitete Zusammenstellung der Personen für Gruppendiskussionen, ungestörte Räumlichkeiten etc.) ist essenziell. Es wurde angeregt, dass Evaluatorinnen und Evaluatoren prüfen sollten, ob Instrumente mit „Doppelfunktion“ (Nutzen für das Projekt/Programm und die Evaluation) einsetzbar sind. Ist eine Erhebung verzichtbar, wenn z.B. Memos von Mitarbeitenden und Teilnehmenden eines Projekts zu bestimmten Projektaktivitäten erstellt werden? Für die Projektbeteiligten kann das der Reflexion dienen, und das Evaluationsteam kann die Memos in einer Dokumentenanalyse auswerten. Ein weiteres Beispiel ist das Beschreiben und Auswerten von Papiertischdecken oder Flipcharts von Kleingruppen bei Großveranstaltungen, z.B. im Rahmen von Open-Space-Veranstaltungen. Zudem wurde überlegt, Informationen, die in neuen Medien generiert werden und die ohnehin verfügbar sind, als Datengrundlage mit einzubeziehen.
- *Instrumentenentwicklung:* Aus Diskussionen der Teilnehmenden ging hervor, dass der Leitfaden/die Frageroute erst nach der Klärung des Zwecks/Ziels, der Fragestellungen, der Stichprobe, der Erhebungssituation und der Auswertungsmethode entwickelt werden sollte. Den Auftraggebenden sollte transparent gemacht werden, welche Fragen zur Untersuchung welcher Fragestellungen und Hypothesen gestellt werden. Das vermeidet bzw. reduziert Missverständnisse und Erklärungsbedarfe.

Erhebungsphase:

Für die Erhebungsphase wurde in den Arbeitsgruppen diskutiert, welche ressourcenschonenden Zugänge gefunden und welche Personengruppen in die Erhebung einbezogen werden können.

- *Zugang über Communities:* Um den Aufwand zur Ansprache von Proband(innen) und die Erschließung von Zugängen zur Zielgruppe zu verringern, wurde in einer Arbeitsgruppe vorgeschlagen, Personen und Communities als „Türöffner“ zu nutzen, die ohnehin bereits Zugang zu den anvisierten Erhebungspersonen haben (z.B. Sozialarbeiter(inne)n).
- *Einbeziehung der Praxis in die Erhebung:* Diskutiert wurde, ob es sinnvoll sein könnte, Praktiker(innen) für die Durchführung von Wirkungsanalysen zu

- qualifizieren, beispielsweise indem entsprechende Inhalte als Module in die Ausbildung von Sozialarbeiter(inne)n und Multiplikator(inn)en eingespeist werden. Hier wurde jedoch tendenziell – vor allem in Abhängigkeit von den eingesetzten Methoden und Themen/Zielen der Evaluation – eine (zeitliche und fachliche) Überforderung der Praxis gesehen. Praktiker(innen) könnten dennoch systematisch auf Wirkungsindikatoren achten und Hinweise für Wirkungen protokollieren. Selbstbeobachtung und Reflexion könnten dabei in Maßnahmen und Programmen als Bestandteil von Kurskonzepten/Schulungsinhalten etc. noch stärker integriert werden. In der Arbeitsgruppe zu Mixed Methods wurde angeregt, Erhebungsdaten, die von Praktikerinnen und Praktikern generiert werden, mit Expertenerhebungen zu triangulieren.
- *Erhebung durch Peers*: Ferner wurde vorgeschlagen, beispielsweise bei der Evaluation von Bildungsmaßnahmen für Jugendliche, Peers in die Erhebung einzubeziehen. Aufgrund fehlender Erfahrungswerte wurde dieser Punkt nicht abschließend diskutiert.

Dokumentations- und Auswertungsphase:

Für die Phase der Ergebnisdokumentation, Analyse und Aufbereitung wurden folgende Ideen für eine ressourcenschonende Umsetzung gesammelt und mit Blick auf die Einhaltung von fachlichen Standards erörtert.

- *Dokumentation*: Der Spielraum für die Dokumentation ist zum einen abhängig vom Zweck/Ziel der Erhebung und zum anderen davon, wie die Erhebungsergebnisse genutzt werden sollen. Auch diesbezüglich ist eine Klärung mit den Auftraggebenden anzuraten (siehe Planungsphase). Gerade beim Einsatz qualitativer Verfahren stellt sich die Frage nach dem Umfang der Dokumentation. Die Spanne kann durch zwei Beispiele verdeutlicht werden:
 - Sollen die Ergebnisse *nur* dazu genutzt werden, um mit einem Projektteam Optimierungsmöglichkeiten für die Durchführung einer zweistündigen Aktivität zu erarbeiten, reicht ggf. ein Protokoll mit den zentralen Aussagen der Befragten, die in einem Workshop mit dem Projektteam diskutiert und interpretiert werden.
 - Sollen die Ergebnisse genutzt werden, um die Wirkungen eines umfangreichen und kostenintensiven Programms zu beurteilen, damit Auftraggebende über dessen Fortführung entscheiden können, sind vollständige Transkripte (von mehreren Erhebungen zur Realisierung einer Datentriangulation) zu empfehlen, die von mindestens zwei Forscherinnen bzw. Forschern getrennt voneinander ausgewertet werden. Die Analyseergebnisse werden dann anschließend zusammengeführt (Forschertriangulation).
- In der AG zu qualitativen Verfahren wurde zudem diskutiert, dass bestimmte Auswertungsmethoden nach spezifischen Dokumentationen verlangen. Beispielsweise erfordert die dokumentarische Methode nach Bohnsack ein Volltranskript mit Kennzeichnungen zur Lautstärke, Betonung, Unterbrechungen etc. Die Evaluationspraxis der Mitdiskutierenden wird den Erfahrungsberichten zufolge aber wesentlich geprägt von der inhaltsanalytischen Methode nach Mayring. Nach den Erfahrungen

der Mitdiskutierenden sind die folgenden zwei Fälle am häufigsten anzutreffen und folgende Vorgehensweisen vertretbar:

- Interview: Das Telefonat wird auf Tonband mitgeschnitten. Die/der Interviewende fertigt während des Interviews stichwortartige Notizen des Verlaufs sowie der zentralen Aussagen an und erstellt anschließend ein Gedächtnisprotokoll. Abschließend wird das Tonband komplett abgehört, die für die untersuchten Fragestellungen wichtigen Aussagen werden ausführlich in eigenen Worten wiedergegeben oder vollständig transkribiert und damit das Gedächtnisprotokoll überarbeitet, ergänzt sowie korrigiert. Wörtliche Zitate werden immer vollständig transkribiert.
- Gruppendiskussion: Hier empfiehlt sich nach der Erfahrung der Anwesenden eine zweite Person zur Protokollführung. Die Moderation einer Gruppenerhebung erfordert eine hohe Aufmerksamkeit und Beobachtung auch nonverbaler Kommunikation in der Gruppe, was kaum zulässt, nebenbei Ergebnisnotizen anzufertigen. Die Dokumentation im Anschluss an die Erhebung entspricht denen des Interviews (siehe oben).
- *Auswertung*: Bezüglich der Auswertungsmethoden wurden in der AG qualitative Verfahren Erfahrungen bei der Anwendung der inhaltsanalytischen Auswertung nach Mayring und der dokumentarischen Methode nach Bohnsack ausgetauscht. Die am weitesten verbreitete Software zur Textanalyse ist MAXQDA, zur Transkription f4.

Phase der Ergebnisaufbereitung und Rückspiegelung der Evaluationsergebnisse:

In den Arbeitsgruppen wurden ferner einige Anregungen für eine ressourcenschonende Aufbereitung und Weitergabe von Resultaten gesammelt.

- *Ergebnisaufbereitung*: Bezüglich der Aufbereitung von Evaluationsergebnissen für die Auftraggebenden und ggf. weitere Personengruppen wurden als Alternativen für einen ausführlichen schriftlichen Bericht Präsentationen sowie Schaubilder an Moderationswänden mit unterschiedlichen Visualisierungsmethoden angesprochen.
- *Rückspiegelung der Ergebnisse*: Von Teilnehmenden wurde es als relevant erachtet, den Mut zu haben, „kleine“ Ergebnisse darzustellen und zu erklären, wie die Qualität und Aussagekraft von Ergebnissen ist. Wie bereits in der Planungsphase ist es vonseiten der Evaluatorinnen und Evaluatoren sinnvoll, wiederum auch die Anforderungen von außen an die Auftraggebenden (beispielsweise aus dem politischen Raum) zu reflektieren und Argumente für die Weitergabe der Ergebnisse an zusätzliche Stakeholder zu liefern. Auch hier gilt es, Sensibilisierungsarbeit zu leisten und gegenüber Auftraggebern, weiteren Stakeholdern – und auch anderen Evaluatorinnen und Evaluatoren – zu klären, dass harte Wirkungsnachweise nicht immer erbracht werden können und ggf. auch „kleine“ Ergebnisse für die Legitimation von Maßnahmen und Programmen ausreichen. Als Idee wurde zudem geäußert, neben dem Evaluationsteam weitere Personengruppen in die Kommunikation von Ergebnissen einzubeziehen, um eine Er-

gebnisbetrachtung aus verschiedenen Perspektiven zu ermöglichen und das Verständnis für die Ergebnisinhalte abzusichern.

Für eine ausführliche Erörterung der Themen Auswertung und Aufbereitung der Evaluationsergebnisse fehlte in den Arbeitsgruppen leider die Zeit.

Fazit

Die Idee zu dem vielleicht etwas unüblichen Tagungskonzept mit einem hohen Anteil an Arbeitsgruppen entstand während der DeGEval-Jahrestagung 2014 in Zürich. Dort hatte das AK-Sprecher(innen)team den Eindruck, dass es oft weniger die großen Themen als vielmehr die alltäglichen Probleme im Evaluationsalltag sind, die Evaluatorinnen und Evaluatoren unter den Nägeln brennen. Die Tagung war in gewisser Weise ein Experiment, das nur gelingen konnte, wenn die Teilnehmenden bereit waren, in den Arbeitsgruppen auch ein wenig ‚aus dem Nähkästchen zu plaudern‘, Dinge zu fragen und Lösungen vorzuschlagen, die sie bisher eher für sich behalten hatten.

Rückblickend war das Konzept ein voller Erfolg. Das große Interesse im Vorfeld der Tagung zeigte, dass das AK-Sprecher(innen)team mit dem Tagungsthema und dem Tagungsformat einen Nerv getroffen hatte. Nach den Eingangsreferaten, die den Rahmen der Tagung noch einmal verdeutlichten, ist es in den Arbeitsgruppen gelungen, eine vertrauensvolle und offene Diskussionsatmosphäre zu schaffen, in der neue und ungewöhnliche Ideen vorgestellt und diskutiert wurden. Die überaus positiven Rückmeldungen am Ende der Tagung bestätigen die Einschätzung, dass es einen erheblichen Bedarf an einem Tagungsformat gibt, das einen offenen und vertrauensvollen Austausch fördert. Zudem ist dieses Format mit einem relativ begrenzten Aufwand zu realisieren. Das AK-Sprecher(innen)team kann es daher als Ergänzung zur üblichen Form einer Fachtagung uneingeschränkt empfehlen.

Danksagung

Nicht nur für die inhaltliche Unterstützung und die Organisation der Tagungsräume, sondern auch für die Zubereitung eines köstlichen Chilis gebührt Olaf Lobermeier ein besonderer Dank. Herzlich danken möchten wir auch Jana Klemm (proVal) für ihre organisatorische Unterstützung während der Tagung, Edith Halves (HAW) für die Unterstützung bei der Arbeitsgruppenmoderation sowie unserem Kooperationspartner, dem Zentrum Demokratische Bildung in Wolfsburg, der uns die Tagungsräume zur Verfügung gestellt hat. Ein herzliches Dankeschön geht schließlich an Florentina Astleithner (FH Campus Wien), Benjamin Froncek (FernUniversität in Hagen) und Peter Stepanek (FH Campus Wien), die sich als Vortragende auf das Tagungskonzept eingelassen und während der gesamten Tagung engagiert mitgearbeitet haben.

Gewinnbringende Anwendung von Evaluation in der unternehmerischen Praxis – eine Fallstudienanalyse

Martin Rost,¹ Julia Hapkemeyer,² Sonja Kind,³ Dina Weiler⁴

Im Rahmen dieses Artikels sollen der Nutzen und die Einsatzmöglichkeiten von Evaluation in der Wirtschaft dargestellt werden. Einsatzmöglichkeiten und Nutzen wurden dafür anhand von sechs Fallstudien analysiert, die der Arbeitskreis „Evaluation in der Wirtschaft“ der DeGEval gesammelt hat. Diese sind ausführlich in der Broschüre „Praxisbeispiele: Wie Evaluation gewinnbringend in der unternehmerischen Praxis angewandt wird“ dargestellt (siehe Hapkemeyer/Kind/Rost/Weiler 2014). Die Fallstudien wurden von den Evaluierenden verfasst und wissenschaftlich aufbereitet. Die Autorinnen und Autoren dieses Beitrags wählten die Fallbeispiele aus und generalisierten die Aussagen, so dass die Fallstudienammlung eine Übersicht über zentrale Einsatzmöglichkeiten von Evaluation in der Wirtschaft bietet.

Aber was bedeutet der Begriff Evaluation in der Wirtschaft überhaupt? Westermann (2002: 4) versteht unter „Evaluation [...] die explizite und systematische Verwendung wissenschaftlicher Forschungsmethoden zur Beschreibung und Bewertung bestimmter Gegenstände[...]“. Ausgehend von früheren Publikationen des Arbeitskreises „Evaluation in der Wirtschaft“ (z.B. Arbeitskreis „Evaluation in der Wirtschaft“ 2008) wurde folgende Beschreibung der Anwendungsmöglichkeiten von Evaluation in der Wirtschaft herausgearbeitet:

Mit Hilfe von Evaluation werden die Wirkung, der Erfolg und die Qualität von in Unternehmen durchgeführten Maßnahmen geprüft. Diese Methode ermöglicht einen ganzheitlichen Blick auf die betrieblichen Strukturen, Prozesse und deren strategische Ausrichtung. Dabei können auch nichtbeabsichtigte Effekte erkannt und umfassende Lernprozesse in den Unternehmen initiiert werden. Die Ziele der Organisation werden transparenter, Verbesserungspotenziale in Bezug auf Strukturen und Prozesse identifiziert und Handlungsempfehlungen abgeleitet (Hapkemeyer/Kind/Rost/Weiler 2014: 4).

1 Universität Stuttgart
2 EO Institut GmbH, Berlin
3 iit – Institut für Innovation und Technik, Berlin
4 Simdustry – S(t)imulate your Business, Berlin

Folgende Ziele von Evaluationen wurden im Rahmen der Fallstudienanalyse identifiziert: Überprüfung der Wirksamkeit von Personalentwicklungsmaßnahmen, Überprüfung von Organisationsentwicklungsmaßnahmen, Begleitung von strategischen Neuausrichtungen und Fusionen, Erfassung der Wirkung von Öffentlichkeitsarbeit, Erfassung der Qualität externer Dienstleistungen sowie Initiierung von kontinuierlichen Lern- und Verbesserungsprozessen (siehe Tabelle 1). Diese letzte Anwendungsmöglichkeit findet sich implizit in fast allen analysierten Fallstudien wieder. Evaluierende können den Organisationsmitgliedern durch ihre Unabhängigkeit den Spiegel vorhalten und damit Reflexionsprozesse auslösen, die die Lern- und Anpassungsfähigkeit der Organisation insgesamt erhöhen (vgl. Kraus/Rost 2012). Diese Möglichkeit zu nutzen und explizit zu kommunizieren, stellt eine große Herausforderung für Evaluierende dar.

In Bezug auf die wissenschaftlichen Methoden bildeten Dokumentenanalysen verbunden mit Befragungen (Interviews, Fragebögen, Gruppendiskussionen) den Kern der meisten Evaluationen. Diese Methoden werden je nach Fragestellung durch Methoden aus dem ganzen Spektrum der empirischen Sozialforschung, der Betriebswirtschaftslehre und den Ingenieurwissenschaften ergänzt.

Die in der folgenden Tabelle im Überblick dargestellten Evaluationsbeispiele finden sich ausführlich in der Broschüre „Praxisbeispiele: Wie Evaluation gewinnbringend in der unternehmerischen Praxis angewandt wird“ des Arbeitskreises „Evaluation in der Wirtschaft“ der DeGEval.

Tabelle 1: Wofür Evaluation in Unternehmen einsetzen

Unternehmerische Fragestellung	Lösungsansatz
...die Wirksamkeit einer Organisationsentwicklungsmaßnahme überprüfen und steigern.	Eine prozessbegleitende Evaluation unterstützt den Organisationsentwicklungsprozess, indem sie Hinweise auf Fortschritte, Anpassungsbedarfe und Lernpotenziale der Organisation liefert.
...die strategische Ausrichtung des Gesamtunternehmens einer umfassenden, unabhängigen Prüfung unterziehen.	Eine externe Evaluation durch eine unabhängige Kommission aus erfahrenen Führungskräften anderer Unternehmen stellt Ihr Unternehmen auf den Prüfstand. Die wichtigsten Optimierungsbedarfe im Hinblick auf Strategie, Organisation und Prozesse im Unternehmen werden erfasst.
...den Nutzen Ihrer Personalentwicklungsmaßnahmen bewerten.	Die Evaluation dient der Überprüfung, ob die Personalentwicklungsmaßnahme die Ziele erreicht und der Praxistransfer gelingt. Förderliche und hemmende Einflussfaktoren sowie unbekannte Wirkungszusammenhänge werden aufgedeckt.
...die Qualität und den Nutzen von Beratungsleistungen beurteilen.	Eine bilanzierende Evaluation zeigt auf, inwieweit die angebotene Beratungsleistung ihre Ziele und Zielgruppe erreicht. Förderliche und hemmende Einflussfaktoren sowie Verbesserungsmöglichkeiten werden erfasst.
...die Wirksamkeit Ihrer Öffentlichkeitsarbeit bewerten und optimieren.	Eine quantitative und qualitative Medienresonanzanalyse dient der systematischen Bewertung und Steuerung der Öffentlichkeitsarbeit.

Unternehmerische Fragestellung	Lösungsansatz
...einen kontinuierlichen Verbesserung- und Lernprozess initiieren.	Eine Selbstevaluation durch ein betriebsinternes Team ermöglicht es, Verbesserungspotenziale zu identifizieren und Maßnahmen abzuleiten, die Ihre Arbeitsabläufe im Betrieb verbessern und zur Effizienzsteigerung beitragen.
...die Fusion von zwei Organisationseinheiten unterstützen und das Ergebnis des Fusionsprozesses bewerten.	Eine begleitende Evaluation erfasst die Einflussfaktoren des Fusionsprozesses und gibt kontinuierliches Feedback zum Fortschritt. Nach Abschluss des Fusionsprozesses erfolgt eine Gesamtbewertung im Hinblick auf betriebswirtschaftlichen Nutzen, Synergien, Akzeptanz und Arbeitszufriedenheit.

Quelle: eigene Darstellung, modifiziert nach Hapkemeyer, Kind, Rost und Weiler 2014

Literatur

- Arbeitskreis „Evaluation in der Wirtschaft“ in der DeGEval – Gesellschaft für Evaluation e. V. (2008): Leitfaden Evaluation in der Wirtschaft. Verfügbar unter: http://www.degeval.de/fileadmin/users/Arbeitskreise/AK_Wirtschaft/Leitfaden_Evaluation_Wirtschaft_DeGEval.pdf [12.06.2014].
- Hapkemeyer, Julia/Kind, Sonja/Rost, Martin/Weiler, Dina (2014): Praxisbeispiele: Wie Evaluation gewinnbringend in der unternehmerischen Praxis angewandt wird. Mainz: Arbeitskreis „Evaluation in der Wirtschaft – Wirkung, Erfolg und Qualität messen“ in der DeGEval – Gesellschaft für Evaluation e. V. Verfügbar unter: http://www.degeval.de/fileadmin/users/Arbeitskreise/AK_Wirtschaft/Broschuere_Evaluation_in_der_unternehmerischen_Praxis_g.pdf [12.06.2014].
- Kraus, Rafaela/Rost, Martin (2012): Evaluation großangelegter Veränderungsprojekte. In: Rosenstiel, Lutz/von Hornstein, Elisabeth/Augustin, Siegfried (Hg.): Change Management Praxisfälle. Veränderungsschwerpunkte Organisation, Team, Individuum. Berlin, Heidelberg: Springer, S. 13-28.
- Westermann, Rainer (2002): Merkmale und Varianten von Evaluationen: Überblick und Klassifikation. In: Zeitschrift für Psychologie. 210 (1), S. 4-26.

UNSERE BUCHEMPFEHLUNG



Wolfgang Böttcher, Christiane Kerlen, Peter Maats, Oliver Schwab, Sonja Sheikh, (DeGEval-Vorstand) (Hrsg.)

Evaluation in Deutschland und Österreich
Stand und Entwicklungsperspektiven in den Arbeitsfeldern der DeGEval – Gesellschaft für Evaluation

2014, 220 Seiten, br., 29,90 €, ISBN 978-3-8309-3149-2
E-Book: 26,99 €, ISBN 978-3-8309-8149-7



WAXMANN

www.waxmann.com
order@waxmann.com

Profile of the Czech Evaluation Society

Jiri Remr¹

General Introduction

The Czech Evaluation Society (hereinafter referred to as CES) is a non-profit, non-governmental and independent organization staffed only by volunteers without any paid employees. The main goal of CES is to associate individuals and institutions interested in evaluations. Members of CES are driven by their concern in building evaluation capacities within the Czech Republic, increasing the awareness of evaluations among the relevant target groups, improving the reputation of evaluations as well as advancing the professional skills of all evaluators, i.e. practitioners and commissioners. CES represents a platform supporting the cooperation with international associations and networks of evaluators, especially with NESE (Network of European Evaluation Societies) and IDEAS (International Development Evaluation Association).

CES was founded in 2007 by a group of individuals who were interested and engaged in evaluations. The current CES board consists of five members: Jiri Remr (Chairman) and Vladimir Sodomka (Vice-Chairman), representing the statutory body. Further board members are Oto Potluka, Inka Pibilova and Daniel Svoboda.

Up to 2015, there was no formal working group; CES only had an informal task force that evolved spontaneously as a consequence of the division of labor. However, since 2015 CES has three official working groups:

a) DEV

is focused on evaluations of development cooperation programs. CES members cooperate with the Ministry of Foreign Affairs and the Czech Development Agency that establishes an institutional framework of the Czech development cooperation and participates in the creation of development cooperation programs. On a long-term basis the CES members perform peer review of evaluation reports, assist with drafting the terms of reference (ToR) and help with the implementation of evaluation conclusions.

1 Czech Evaluation Society, Prague

- b) ESIF
is concerned with issues associated with evaluating the European Structural and Investment Funds. CES members provide feedback to official and binding guidelines and manuals for evaluation and submit comments of legislation relevant to evaluation (e.g. Financial Control Act, etc.).
- c) LEC
which is engaged in lecturing. CES members organize sessions and seminars with various stakeholders in order to teach and improve the professional skills in performing evaluations and in utilization of evaluation conclusions. Current activities cover not only ad hoc training but also higher (university) education, especially lecturing at the Charles University, Faculty of Social Sciences, Prague, even as part of the ERASMUS program when lectures on evaluation research are open to all students coming from other EU Member States.

Apart from these working groups, the new statutes (came into force in 2015) brought also the Cassation and Ethics Committee. It is expected that in the near future such a committee will also serve as a fourth working group. Its aim is to interpret the provisions of the Ethical Code and the Standards for Evaluations, to find and comment the case studies accused of breaching the rules or customs and to provide mentoring to members and other stakeholders outside the CES. The Cassation and Ethics Committee also resolves and reconciles the disputes among members and provides opinions on professional and ethical matters concerning the evaluations performed.

Membership

CES currently has more than 30 members who are mainly evaluation practitioners. However, among the members there are also academicians (approximately 20%) and evaluation commissioners/clients – usually civil servants and representatives of the managing authorities of the operational programs financed by the European Structural and Investment Funds (approximately 17%). CES currently relies on individual membership. However, the new statutes will enable institutions to become members from 2016 on. The membership base is stable and moderately growing: Approximately five to eight new members apply for membership each year.

As a matter of fact, it is not easy to become a CES member. The entrance procedure is a multistage process that is strictly based on a referral system, i.e. each new applicant must gain the recommendations from at least two current members. Moreover, she/he has to fill in the application form and must explicitly confirm the compliance with the Ethical Code and the Standards for Evaluation. There are basically two routes how one can become a member:

- a) In case current CES members know the applicant, they know her/his work and are aware of the applicant's past and current activities, they do approve the integrity of the applicant and provide recommendations to the Board that (after considering the facts in the application form, the provided recommendations and other credentials of the applicant) approves or disapproves the membership.

- b) In case none of the current CES members knows the applicant, two Board members conduct independently interviews with the applicant focusing on the applicant's professional activities, her/his effort in improving evaluations and building evaluation capacity. The two Board members also inform the applicant about CES activities, opportunities for the applicant's involvement (in case of admission) and principles of ethical and professional conduct. Applicants must show the commitment to comply with the Ethical Code and the Standards for Evaluations. The interviewers then report on the results of the interviews to the Board that makes the decision.

At the end of this process, new members receive an official letter confirming their membership and specifying the names of the two CES members who provided recommendations or performed interviews (such members are considered as 'godfathers'). Full membership obviously requires paying the annual fee. New entrants are introduced to the current CES members in the Newsletter and their profiles are also published on the CES website.

Key Activities of the CES

Annual Conference

In 2015, the Conference was held for the fifth time. Since 2011, this one-day conference has become the most important evaluation event in the Czech Republic. Each year, more than 100 participants attend approximately 15 lectures by Czech and foreign speakers, carefully selected by the Conference Steering Committee. Usually, the opportunities, prospects and alternative routes of future development of evaluations are discussed. However, substantial attention is always paid to evaluation methodology and its advances. Especially, new methods and techniques in evaluation are popular among the auditorium. The conference is also open to those who prefer to present a poster on their case studies or evaluations that used unique design or inquiry that might inspire others.

Peer-Reviewed Journal

Under the auspices of CES, the professional peer-reviewed journal "Evaluation Theory and Practice" is issued. It is the only journal in the Czech Republic that focuses on evaluations and is recognized by the Czech Government as a scientific journal. "Evaluation Theory and Practice" is released twice a year and is mostly written in Czech language (in order to help developing a proper terminology). However, all abstracts and even some articles from foreign authors are published in English. There are 500 copies per issue which are distributed to practitioners, public administration, NGOs, academicians and libraries. All members of CES receive a copy of the journal as part of their membership.

Training

CES is accredited by the Ministry of the Interior to perform courses on evaluation for civil servants in municipalities. Beside this, CES organizes other ad hoc trainings, for instance, in January 2015 two of our members performed a two-day training in participative evaluations that was much appreciated. For some training events CES invites foreign lecturers – e.g. in fall 2015 CES organizes a training in conducting the focus groups and in qualitative data analysis led by Janet Mancini Billson, PhD.

EPDET

CES, together with Development Worldwide and the Slovak Evaluation Society, is co-organizing the European Program for Development Evaluation Training (EPDET). It is a comprehensive, intensive, five-day training program that is open to evaluators and managers of development projects and programs from the whole world. It is the European offer of the summer school IPDET (International Program for Development Evaluation), organized by the Carleton University, Ottawa, Canada, and that is why its structure and content follow a successful scheme of the IPDET.

Evaluation Café

In 2015, CES introduced a new format of meetings for evaluators (CES members are invited guests as well). These sessions in an informal atmosphere provide the great opportunity for approximately seven to ten people to discuss important issues concerning the current evaluation practice. For instance, the title of our last Café was “Evaluations that really had impact”.

Newsletter

Each member bi-monthly receives the newsletter that has a form of e-zine. The newsletter summarizes the key events of the past two months, informs about planned conferences in Europe and other parts of the world, about the CES Board decisions and activities of working groups, and introduces new members.

Documents of CES

There are two documents that CES considers crucial, apart from our statutes and other legal documents:

Ethical Code

CES has its own Ethical Code since 2011. This document reflects the values that CES considers important in the course of evaluation conduct – proficiency, integri-

ty and accountability. Each value is further elaborated in detail and a set of specific provisions is defined.

Standards for Evaluation

Beside the Ethical Code, CES implemented the Joint Committee Standards for Educational Evaluation. In this respect CES decided not to create an own set of standards because they wanted to express commitment and adherence to the norms that were adopted by many evaluation societies around the world. In doing so, CES shows its appreciation to principles defined by evaluators from countries with longer tradition in evaluations.

Both documents are generally binding in the Czech Republic because they were approved by the government and released as an appendix of governmental decree (No. 597/2013) on guiding documents for evaluation.

UNSERE BUCHEMPFEHLUNG



Marie-Luise Schütt

E-Learning als Baustein im inklusiven Unterstützungs- und Beratungssystem in Deutschland

Konzeption, Implementierung und Evaluation des Onlineangebots „MIT BISS“ für Regelschullehrerinnen und Regelschullehrer

Internationale Hochschulschriften, Bd. 617, 2015, 252 Seiten, br., 34,90 €, ISBN 978-3-8309-3218-5 E-Book: 30,99 €, ISBN 978-3-8309-8218-0

Angesichts des (inter-)nationalen Bestrebens, inklusive sowie qualitativ hochwertige Bildungsstrukturen zu verwirklichen, müssen die spezifischen Belange von Kindern und Jugendlichen mit Beeinträchtigung in das Zentrum des Handelns aller professionellen Akteure gestellt werden. Im Rahmen dieser Arbeit wurde das notwendige Know-how im Umgang mit Schülerinnen und Schülern mit Sehbeeinträchtigung an der allgemeinen Schule empirisch geprüft. In Verbindung mit lerntheoretischen und mediendidaktischen Erkenntnissen kam es zur Konzeption des Onlineangebots MIT BISS (Methodisches • Informatives • Theoretisches – Basics zur inklusiven Beschulung sehgeschädigter Schülerinnen und Schüler), welches die eigenständige Weiterqualifizierung ermöglicht.

Basierend auf den Forschungsergebnissen zu MIT BISS werden Nutzenpotentiale digitaler Anwendungen für den zukünftigen Einsatz im inklusiven Unterstützungs- und Beratungsprozess herausgestellt und relevante Empfehlungen für das Handlungsfeld Schule unter Bezugnahme auf theoretische Konzepte gegeben.



WAXMANN

www.waxmann.com
order@waxmann.com