

Editorial

Der weltweite Evaluationsboom ist ungebrochen, wie sich leicht an einer Reihe von Indikatoren zeigen lässt: (1) In vielen Ländern ist Evaluation in die Management- und Entscheidungsprozesse von Regierungen, Ministerien, öffentlichen Einrichtungen und Nichtregierungsorganisationen integriert. Zunehmend werden gesetzliche Grundlagen für die institutionelle Verankerung der Evaluation geschaffen (vgl. Rosenstein 2013). (2) Die steigende Nachfrage nach Evaluation hat zu einem breiten Angebotsmarkt geführt, von dem vor allem Consultings profitieren (vgl. Leeuw/Toulemonde/Brouwers 1999: 487ff.). (3) Als Zeichen einer zunehmenden Professionalisierung kann die in den letzten Jahren stark gewachsene Zahl von nationalen und internationalen Evaluationsgesellschaften und -netzwerken gelten. Die International Organization for Cooperation in Evaluation (IOCE) listet auf ihrer Webseite aktuell 162 Organisationen auf. (4) Auch die weltweit offerierte Zahl von akademischen Studiengängen und Weiterbildungsangeboten expandiert (vgl. Beywl/Harich 2007: 121ff.; Rugh 2011: 586ff.). (5) Und internationale Organisationen und Netzwerke wie die Weltbank mit ihren ‚Regional Centers for Learning on Evaluation and Results‘ oder die International Development Association und EvalPartners, aber auch nationale Akteure, wie das BMZ mit seiner Evaluation Capacity Initiative, helfen beim Aufbau von Evaluationsstrukturen. (6) Die Verbreitung von Evaluationsergebnissen über das World Wide Web und über Publikationen hat zudem ein ungeahntes Ausmaß angenommen.

Dieser weltweit zu beobachtende Boom ist allerdings nicht ohne Gefahren (vgl. Stockmann, im Erscheinen) und natürlich auch nicht in allen Ländern gleich stark ausgeprägt. Einen schönen Überblick über den Ausprägungsgrad nationaler Evaluationskulturen gibt die Zusammenstellung von Speer, Jacob und Furubo (im Erscheinen). Anhand von neun Indikatoren wird aufgezeigt, wie sich die nationalen Kulturen in den letzten zehn Jahren verändert haben. Schaut man sich die Entwicklung der drei deutschsprachigen europäischen Länder an, dann fällt auf, dass sich die Schweiz im weltweiten Vergleich von einem unteren Rangplatz hinter Finnland an die Spitze aller untersuchten Länder gesetzt hat. Deutschland hingegen, 2002 nur im unteren Mittelfeld platziert, ist auch 2012 nicht vorangekommen. Österreich hingegen (leider nicht in die Auswertung von Speer, Jacob und Furubo mit aufgenommen) hat ebenfalls einen erheblichen Sprung nach vorne gemacht: Mit einem 2013 verabschiedeten Gesetz ist eine wirkungsorientierte Folgenabschätzung aller Regelungsvorhaben vorgeschrieben. Auch größere Projekte müssen auf ihre zu erwartenden

den Wirkungen ex-ante und innerhalb von fünf Jahren mit einer internen On-going- oder Ex-post-Evaluierung durch das jeweils zuständige Ressort überprüft werden (vgl. Bundeskanzleramt Österreich 2014).

Nur in Deutschland bewegt sich kaum etwas. Zwar hat der Staatssekretärsausschuss ‚Bürokratieabbau und bessere Rechtsetzung‘ im Januar 2013 eine Konzeption zur Evaluierung neuer Regelvorhaben beschlossen, nach der grundsätzlich alle wesentlichen Gesetze und Verordnungen ab einem Schwellenwert von einer Million Euro drei bis fünf Jahre nach dem Inkrafttreten evaluiert werden sollen, doch der Begriff Evaluation wird dabei so weit gefasst, dass schon eine bloße Einschätzung eines Ressortbeamten als Evaluation durchgeht. Derartige Aktionen, die Evaluationen vollkommen sinnentleeren und das Instrument zu einer Worthülse verkommen lassen, stellen für die Profession eher eine Gefahr als eine Chance dar. Deshalb ist es wohl kein Schaden, dass nach einer nicht repräsentativen Kurzumfrage durch den Autor dieses Editorials bei Evaluationsverantwortlichen in Bundesministerien dieser Beschluss vollkommen unbekannt war.

Auch das vor zweieinhalb Jahren vom BMZ gegründete Deutsche Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), das die politische Institutionenlandschaft in Deutschland hätte verändern können, da es ein Mandat hat, die deutsche Entwicklungszusammenarbeit wissenschaftlich unabhängig u.a. auf ihre Wirksamkeit hin zu überprüfen, hat noch nicht richtig Tritt gefasst. Dabei könnte dieses Institut nicht nur beispielhaft zeigen, wie ein Politikfeld fachgerecht evaluiert wird, um Entscheidungsträger mit validen Steuerungsinformationen zu versorgen, sondern auch, wie dadurch die Öffentlichkeit transparent über die Ergebnisse von Politik solide und nachvollziehbar auf der Grundlage wissenschaftlich abgesicherter Befunde informiert werden kann (vgl. Stockmann 2012: 85ff.). Mit gerade mal zwei fertiggestellten Evaluationsberichten (einer davon wird im Praxisteil hier vorgestellt) hat es bisher die hochgesteckten Erwartungen nicht erfüllen können. Der Leiter des Instituts, Prof. Dr. Helmut Asche, wurde erstaunlich geräuschlos Mitte Juli dieses Jahres mit sofortiger Wirkung entlassen. Weder das BMZ noch der Beirat des DEval haben sich zu diesem Vorgang offiziell geäußert. Das ursprünglich als ‚unabhängiges‘ Institut gegründete DEval wird jetzt kommissarisch vom BMZ selbst geleitet.

Nun, vielleicht kann das Internationale Jahr der Evaluation in 2015 ja auch in Deutschland zu einer Trendwende führen. Vor allem auf Initiative von EvalPartners, „which is an innovative global partnership movement to strengthen national evaluation capacities particularly through civil society“ (Segone, im Erscheinen), wurde das Jahr 2015 zum Internationalen Jahr der Evaluation gekürt.



Ziel ist es, Evaluation und „evidence based policy making“ auf internationaler, regionaler, nationaler und lokaler Ebene zu promoten. Hierzu werden weltweit Veranstaltungen, Konferenzen und unterschiedlichste Events geplant, auch in Deutschland. Die DeGEval plant, die Frühjahrstagungen der Arbeitskreise sowie die Jahrestagung unter einem gemeinsamen Dach zu vereinen. Dazu erhalten die Arbeitskreise die Gelegenheit, internationale Referent/inn/en einzuladen. Ein gemeinsamer Fly-

er zu allen Veranstaltungen soll auf das Internationale Jahr der Evaluation aufmerksam machen. Auch das DEval plant in 2015 Veranstaltungen, die vor allem auf die politische Entscheidungsebene zielen. Das Centrum für Evaluation (CEval) wird in Zusammenarbeit mit der Hochschule für Technik und Wirtschaft des Saarlandes (HTW) anlässlich des zehnjährigen Bestehens des Masterstudiengangs Evaluation am 11. und 12. Juni an der Universität des Saarlandes eine Tagung zur Zukunft der Aus- und Weiterbildung in der Evaluation veranstalten. Außerdem wird das CEval mit der Universidad de Costa Rica und dem CICAP (Centro de Investigación y Capacitación en Administración Pública) im internationalen Jahr der Evaluation einen Workshop für Minister und Vizeminister aller mittelamerikanischen Länder ausrichten, um vor allem den Nutzen der Evaluation für die politische Entscheidungsfindung zu vermitteln. Mit Sicherheit werden auch noch viele andere zivilgesellschaftliche Organisationen in Deutschland die Chance des Evaluationsjahres nutzen, um das Thema in Politik und Gesellschaft international und national zu verbreiten.

Die Zeitschrift für Evaluation widmet sich seit 2002 der Aufgabe, Wissenschaft und Praxis über interessante Entwicklungen, Theorien, Methoden und Ergebnisse aus diesem Bereich zu informieren. Auch dieses Heft, das einen starken methodischen Fokus aufweist:

Jan Ulrich Henses Beitrag liegt ganz im Trend der internationalen Diskussion über die Professionalisierung der Evaluation durch Zertifizierung, Aus- und Weiterbildung und durch die Definition von Kompetenzen, über die ein(e) Evaluator(in) verfügen sollte. Der Autor beschreibt die Entwicklung einer Fragebogenbatterie zur Selbsteinschätzung von Evaluationskompetenzen, die zu großen Teilen auf dem Kompetenzrahmen der DeGEval (2004) beruht, der auch dem Masterstudiengang in Saarbrücken zugrunde liegt (<http://www.master-evaluation.de/>). Mithilfe einer Faktorenanalyse können sechs Dimensionen generiert werden: Theorie und Geschichte der Evaluation, Evaluationsmanagement, quantitative und qualitative Methoden, Organisationswissen sowie Sozial- und Selbstkompetenzen. Der Autor sieht für sein Instrument vor allem Anwendungsgebiete in der Aus- und Weiterbildung sowie in der Forschung über Evaluation. So könnten Fragebatterien z.B. im Rahmen von Bedarfsanalysen dazu eingesetzt werden, selbsteingeschätzte Kompetenzdefizite aufzudecken. Die Ergebnisse könnten dann bei der Gestaltung oder Auswahl geeigneter Angebote zur Aus- oder Weiterbildung in die Entscheidung einfließen.

Der zweite Originalbeitrag von *Tobias Wolbring* beschäftigt sich mit dem empirischen Nachweis kausaler Effekte und ihrer Bewertung im Rahmen von Wirkungsevaluationen, einem genauso zentralen wie aktuellen Thema. Hierfür verbindet er Rubins Potential Outcomes Framework mit Pearls Graphenmethodologie. Da beide Verfahren komplementäre Herangehensweisen darstellen, lassen sich deren unterschiedliche analytische Schwerpunkte und Methodologien nutzen. Während anhand des Potential Outcomes Frameworks die Vorzüge experimenteller Designs verdeutlicht werden, wird mit der Graphenmethodologie gezeigt, wie die hierfür erforderlichen Kausalaussagen formuliert werden können. Ziel ist es, die Bewertungen von Interventionen auf eine solide empirische Basis zu stellen und die kausalen Effekte nicht nur zu schätzen, sondern im Sinne einer White-Box-Evaluation die Wirkungsmechanismen zu identifizieren. Hierfür sind klare programmtheoretische Aussagen

zu formulieren, so dass die für ihre Überprüfung adäquaten Untersuchungsdesigns ‚maßgeschneidert‘ entwickelt werden können. Die theoretische Modellierung mittels Pearls Graphenmethodologie macht eine Unterscheidung der Variablen nach Confounder, Collider und Mediatoren notwendig, um keine (zusätzlichen) statistischen Verzerrungen zu erzeugen. Anhand der Bologna-Reform wird diese kombinierte Methodik an einem Beispiel demonstriert.

Mit dem Thema der Wirkungsevaluation setzen sich auch zwei Praxisberichte auseinander: (1) *Sonja Possinger* und *Susanne von Jan* stellen einen vom CEval im Auftrag der GIZ entwickelten ‚alltagstauglichen‘ Evaluationsansatz zur Wirkungsevaluation dar. Ziel war es, einen methodischen Ansatz zu entwickeln, der zum einen der aktuellen wissenschaftlichen Debatte und den geforderten wissenschaftlichen Standards, zum anderen aber auch den spezifischen Rahmenbedingungen von Projekten und Programmen der Entwicklungszusammenarbeit gerecht wird, so dass robuste, empirisch fundierte Ergebnisse produziert werden können. Der Ansatz zeigt, dass sich die Ansprüche alltagstauglich und ‚rigorous‘ keinesfalls ausschließen müssen, wenn man einen gewissen Pragmatismus zulässt. So versteift sich der CEval-GIZ-Ansatz nicht allein auf experimentelle oder quasi-experimentelle Designs, sondern greift auch auf Zeitreihenanalysen, Längsschnitt- und speziell Paneldesigns zurück. ‚Alltagstauglichkeit‘ bedeutet, dass je nach Programmtyp und verfügbarer Datenlage dasjenige Evaluationsdesign ausgewählt wird, welches die höchstmögliche Datenqualität verspricht und – unabhängig vom Design – mehrere Datenerhebungs- und Analysemethoden miteinander verknüpft. In sieben Evaluationen wurde die Praxistauglichkeit des Ansatzes erfolgreich getestet. Die Ergebnisse lassen die Empfehlung zu, dass der gemeinsam entwickelte Ansatz auch für Evaluationen außerhalb des GIZ-Kontextes tauglich ist.

(2) Ebenfalls mit Wirkungen setzt sich der Praxisbeitrag von *Michael Gajo* auseinander. Er stellt den Wechsel von der Wirkungskette, die wegen ihrer Monokausalität oft kritisiert worden ist, zum Wirkungsmodell der GIZ dar. Das ‚Neue‘ am Wirkungsmodell ist zunächst, dass vom Problem als dem zu verändernden defizitären Zustand her gedacht wird. Daraus wird abgeleitet, was sich verändern muss, damit der gewünschte Zustand eintritt. Wie diese Veränderungen kausal zueinander in Beziehung stehen, aufeinander aufbauen und sich gegenseitig beeinflussen, um die als Wirkungen definierten Ziele zu erreichen, verdeutlicht graphisch das Wirkungsmodell. Zusammen mit der Wirkungsmatrix, die u.a. das Programmziel, den Outcome, die dazugehörigen Outputs und Aktivitäten (Interventionen) sowie die Indikatoren, mit denen die Veränderungen erfasst werden sollen, umfasst, gehört es mittlerweile zum Standardplanungsinstrument der GIZ. Michael Gajo zeigt, wie das Modell in der Praxis in einem Programm zur ‚Klimaanpassung und Biodiversität in Marokko‘ funktioniert.

Der dritte Praxisbericht von *Felix Gaisbauer*, *Martin Noltze*, *Thomas Schwedersky* und *Stefanie Krapp* beschäftigt sich zumindest zum Teil ebenfalls mit der Evaluation von Wirkungen. Es handelt sich dabei um die erste vom Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval) durchgeführte Evaluation. Die Darstellung dieser Studie in diesem Heft stellt somit eine Premiere dar. Ziel der summativen Evaluation war es, den Beitrag der deutschen Entwicklungszusam-

menarbeit (EZ) zum öffentlichen Gesundheitswesen über die letzten 30 Jahre in Ruanda darzustellen. Dabei sollten u.a. nicht nur erfolgreiche und weniger erfolgreiche EZ-Modalitäten und Instrumente und deren Interaktion und Anpassungsfähigkeit an veränderte Rahmenbedingungen innerhalb des Sektors und hinsichtlich sozioökonomischer und politischer Entwicklungen, sondern auch die Relevanz, Effektivität, Effizienz, der Impact, die Nachhaltigkeit, Kohärenz, Koordinierung, Komplementarität und Harmonisierung der ruandisch-deutschen Zusammenarbeit im Gesundheitssektor bewertet werden. Eine komplexe Aufgabenstellung, die der Komplexität des Evaluationsgegenstandes in nichts nachstand. Der Praxisbericht zeichnet das Vorgehen bei der Evaluation nach, stellt ein paar exemplarische Schlüsselergebnisse vor und leitet einige Empfehlungen ab.

Bevor ich zum dritten Originalbeitrag komme, möchte ich noch kurz auf einen weiteren Praxisbericht hinweisen. *Maria Akopova* und *Valeria Chernyavskaya* beschäftigten sich mit den Perspektiven und Herausforderungen der „Evaluation of Academic Science“, wobei sie über ihre Erfahrungen aus der St. Petersburg State Polytechnical University berichten. In dem Beitrag werden die Kriterien und Evaluationsstandards vorgestellt, die dort entwickelt wurden und seit 2013 angewendet werden. Die Autorinnen wollen mit ihrem Report den interdisziplinären Austausch von Informationen zwischen den Universitäten ermutigen.

Der dritte Originalbeitrag von *Christian Diller*, *Philipp Gareis*, *Martin Kohl* und *Isabel Schmidholz* ist ebenfalls stark methodisch orientiert und stellt sich der Frage, wie die Netzwerkbildung in ihrer Dynamik innerhalb von Evaluationen abgebildet werden kann. Zunächst werden allgemein methodische Fragen der Integration von Netzwerkanalysen in Evaluationen diskutiert. Danach wird eine formale Zwei-Punkt-Netzwerkanalyse des Modellvorhabens der Raumordnung zum Thema Daseinsvorsorge vorgestellt, das 2012 bis 2013 im hessischen Vogelsbergkreis durchgeführt wurde. Dabei konnten eine Reihe methodischer Probleme aufgezeigt werden, die erklären könnten, warum quantitative Netzwerkanalysen in der Evaluationsforschung bisher so wenig Anwendung erfahren haben. Während Ex-Post-Ein-Punkt-Befragungen Akteure in der Regel damit überfordern, ex-post die Entwicklung ihrer Akteurskontakte statistisch messbar einzuschätzen, liegt das Hauptproblem bei einer Zwei-Punkt-Netzwerkanalyse bei der hohen Panelmortalität. Da der Erhebungsaufwand zudem hoch ist, vor allem, wenn in einem längeren Zeitabstand die Nachhaltigkeit der Förderimpulse für die Bildung von Netzwerken abgeschätzt werden soll, stufen die Autor/inn/en die Tauglichkeit des Evaluationsinstruments einer formalen quantitativen Zwei-Punkt-Netzwerkanalyse eher zurückhaltend ein.

Neben den gereviewten Originalbeiträgen und den Praxisberichten liefert auch die DeGEval ...Info wieder interessante Beiträge aus dem Innenleben der Gesellschaft. Verschiedene Arbeitskreise berichten von ihren Frühjahrstagungen. Der Arbeitskreis Aus- und Weiterbildung in der Evaluation hat sich unter dem dramatischen Titel „Wie weit können, dürfen und müssen Evaluierende gehen, um die Welt zu retten?“ mit dem Rollenverständnis von Evaluierenden auseinandergesetzt. Der Arbeitskreis Forschungs-, Technologie- und Innovationspolitik widmete sich dem Thema Wirkungsanalysen nach § 7, Abs. 2 der Bundeshaushaltsordnung (BHO). Das klingt auf den ersten Blick nicht spektakulär, ist es aber, denn es geht um die

Frage, wie ernst die Bundesministerien mit dieser verpflichtenden Anforderung, finanzwirksame Maßnahmen zu evaluieren, umgehen. Wie eingangs in diesem Editorial betont, ist es damit in vielen Ministerien nicht weit her und der Bundesrechnungshof hat dies immer wieder beklagt. Der Arbeitskreis Berufliche Bildung hat sich bei seiner Frühjahrstagung mit den Standards in der Evaluation beruflicher Bildung auseinandergesetzt und Anspruch und Wirklichkeit miteinander verglichen. Der Arbeitskreis Entwicklungspolitik beschäftigte sich mit den Fragen, welche Methoden sich zur Evaluation von Politikberatung eignen, anhand welcher Kriterien Advocacy-Arbeit evaluiert werden kann und welche Daten erforderlich sind, um die Vorhaben von Good-Governance-Vorhaben zu bewerten. Weitere Beiträge wie z.B. eine Darstellung der Ukrainian Evaluation Association (UEA) sowie der Österreichischen Plattform für Forschungs- und Technologiepolitikevaluierung sowie ein Einwurf gegen PISA runden den Info-Teil ab.

Schließlich möchte ich es nicht versäumen, darauf hinzuweisen, dass mit dieser Ausgabe ein langjähriger Förderer der Zeitschrift für Evaluation aus dem Editorial Board ausscheidet. Herr Univ.-Prof. i.R. Dr. Helmut Kromrey hat sich nicht nur um die Etablierung des Themas Evaluation in Deutschland verdient gemacht, er war vielmehr auch einer der Gründungsherausgeber, dessen Engagement erheblich dazu beigetragen hat, dass die Zeitschrift für Evaluation aus der Taufe gehoben werden und sich erfolgreich entwickeln konnte. Gerade zu Beginn war es besonders wichtig, dass die Idee, eine Zeitschrift für Evaluation zu gründen, von herausragenden und renommierten Persönlichkeiten unterstützt wurde. Deshalb bin ich Herrn Prof. Dr. Helmut Kromrey auch zu persönlichem Dank verpflichtet, denn der Weg von der Idee bis zur Umsetzung war nicht ohne Hürden. Er hat tatkräftig und mit Zuversicht dazu beigetragen, diese Schwierigkeiten aus dem Weg zu räumen, so dass die Zeitschrift für Evaluation mit einer Auflage von 1.400 Exemplaren heute nicht mehr aus der deutschsprachigen Evaluationslandschaft wegzudenken ist. Bis 2011 hat er dem Herausbergremium angehört und seither die Zeitschrift auch als Mitglied des neu etablierten Editorial Board mit Rat und Tat unterstützt. Im Namen der Herausgeberinnen und Herausgeber sowie der Redaktion möchte ich Herrn Prof. Dr. Kromrey für sein erfolgreiches Wirken ganz herzlich danken!

Bleibt mir nur noch, Ihnen viel Spaß bei der Lektüre dieser von unserem Redakteur, Dr. Hansjörg Gaus, wieder einmal professionell zusammengestellten Ausgabe der ZfEv zu wünschen und natürlich viel Vorfriede auf das Internationale Jahr der Evaluation in 2015.

Reinhard Stockmann

Für die Herausgeberinnen und Herausgeber

Literatur

- Beywl, Wolfgang/Harich, Katja (2007): University-Based Continuing Education in Evaluation. In: *Evaluation*, 13 (1), S. 121-134.
- Bundeskanzleramt Österreich (2014): Bericht über die wirkungsorientierte Folgenabschätzung. Bericht gemäß § 68 Abs. 5 BHG 2013 iVm §§ 6 und 9 Abs. 1 Wirkungscontrollingverordnung. Wien: Bundeskanzleramt Österreich. Verfügbar unter: https://www.oeffentlicherdienst.gv.at/wirkungsorientierte_verwaltung/dokumente/WFA-Bericht_2014.pdf?4e3592 [19.08.2014].
- DeGEval – Gesellschaft für Evaluation e.V. (Hg.) (2004): Empfehlungen für die Aus- und Weiterbildung in der Evaluation: Anforderungsprofile an Evaluatorinnen und Evaluatoren. Köln: DeGEval.
- Leeuw, Frans L./Toulemonde, Jacques/Brouwers, André (1999): Evaluation Activities in Europe: A Quick Scan of the Market in 1998. In: *Evaluation*, 5 (4), S. 487-496.
- Rosenstein, Barbara (2013): Mapping of National Evaluation Policies. Commissioned by Parliamentarians Forum on Development Evaluation in South Asia Jointly with EvalPartners. Verfügbar unter: http://api.ning.com/files/saEvSFZMI0naAMulhi89Gvi-w8IUqOKc*yzLZZ4Hif3BpGzHtm3IXVES8mjmCWstNTOTFWVCvysHjO2kOuLPLlcd5a3BmzOh/MappingNEPReportDecember2013.pdf [18.08.2014].
- Rugh, Jim (2011): The Expanding World Scene in Evaluation. In: *American Journal of Evaluation*, 32 (4), S. 586-593.
- Segone, Marco (im Erscheinen): Foreword. In: Meyer, Wolfgang/Stockmann, Reinhard (Hg.): *The Future of Evaluation. Global Trends, New Challenges and Shared Perspectives*. New York: Palgrave Macmillan.
- Speer, Sandra/Jacob, Steve/Furubo, Jan-Eric (im Erscheinen): Different Paths for Institutionalizing Evaluation. Updating the International Atlas of Evaluation 10 Years Later. In: *Evaluation – The International Journal of Theory, Research and Practice*.
- Stockmann, Reinhard (2012): Von der Idee zur Institution. Institut für deutsche Entwicklungsevaluierung gegründet. In: *Zeitschrift für Evaluation*, 11 (1), S. 85-93.
- Stockmann, Reinhard (im Erscheinen): The Future of Evaluation in Modern Societies. In: Meyer, Wolfgang/Stockmann, Reinhard (Hg.): *The Future of Evaluation. Global Trends, New Challenges and Shared Perspectives*. New York: Palgrave Macmillan.