

Evaluation und Gesellschaft

Bericht zur Jahrestagung der DeGEval – Gesellschaft für Evaluation e.V. 2009 in Münster

*Jan Nikolas Dicke, Nina Högbe*¹

Seit Jahren begleitet die DeGEval – Gesellschaft für Evaluation gesellschaftliche Prozesse gleichermaßen kritisch wie professionell. Die Standards für Evaluation, mit der die DeGEval seit dem Jahr 2001 Kriterien und Qualität der Evaluation beeinflusst, stehen beispielhaft für dieses Bestreben. Auf der Jahrestagung 2009 ermutigte *Professor Dr. Dr. Christiane Spiel* die Mitglieder, diesen Weg weiterzugehen. Zugleich forderte sie aber auch verstärkte Anstrengungen: Die Standards der Evaluation würden nicht hinreichend wahrgenommen, und auch gegenüber der Politik müsse die DeGEval stärker als bisher auf ihre Leistungen hinweisen.

Die Vorstandsvorsitzende der DeGEval umriss damit ein Problem, mit dem sich die rund 320 Tagungsteilnehmerinnen und -teilnehmer intensiv beschäftigten. Drei Tage lang, vom 7. bis zum 9. Oktober 2009, diskutierten sie in Münster das Verhältnis von Evaluation und Gesellschaft. Wie kann Evaluation zu einer höheren Transparenz gesellschaftlicher Prozesse beitragen? Und wie kann sie zu einer stärkeren Partizipation an politischen Entscheidungen führen? Wie viel – und welche – Evaluation braucht eine Gesellschaft?

Ein besonderes Augenmerk lag dabei auf dem Politikfeld ‚Bildung‘. Die Jahrestagung der DeGEval fand gemeinsam mit der Herbsttagung der Kommission für Bildungsorganisation, Bildungsplanung, Bildungsrecht (KBBB) der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE) statt. Deren Sprecher, *Professor Dr. Wolfgang Böttcher* (Universität Münster), fungierte zugleich auch als Ausrichter der Veranstaltung.

Als Keynote-Speaker hatten DeGEval und KBBB international renommierte Kolleginnen und Kollegen gewinnen können: *Professor Dr. Jennifer Greene* (University of Illinois at Urbana-Champaign), *Professor Dr. Frans L. Leeuw* (University of Maastricht & Netherlands Justice Research Institute) sowie *Linda Mabry, Ph.D.* (Washington State University).

Jennifer Greene unterstrich in ihrem Beitrag „Evaluation in Service of the Public Good“ die politische wie gesellschaftliche Bedeutung von Evaluation. Um dem öffentlichen Wohl zu dienen, müsse sich Evaluation nicht nur mit relevanten gesell-

¹ Institut für Erziehungswissenschaft, Abteilung 2, Qualitätsentwicklung und Evaluierung, Westfälische Wilhelms-Universität Münster

schaftlichen Themen befassen und einen Beitrag zur Klärung gesellschaftlich relevanter Fragestellungen leisten: Sie müsse auch divergierende Perspektiven berücksichtigen. Evaluatorinnen und Evaluatoren sollten auf dreifache Weise Verantwortung übernehmen: als Bürgerinnen und Bürger, als Lehrende (*educator*) sowie als Vermittelnde der Demokratie.

Als Bürgerin bzw. Bürger müssten sie sich auf das Potential und die Verpflichtung zurückbesinnen, in bedeutenden öffentlichen Fragen beratend zur Seite zu stehen und Evaluation als eine öffentliche Dienstleistung zu verstehen, aus der eine bürger-schaftliche Verantwortung erwachse. So dürften Evaluatorinnen und Evaluatoren die öffentliche Entscheidungsfindung nicht mehr nur stumm beobachten. Statt weitgehend abseits zu stehen, müssten sie ein aktiver Teil der gesellschaftlichen Ordnung werden. Als Lehrende sollten sie die Herausforderung annehmen, Evaluation aus den engen Grenzen der Rechenschaftslegung zu befreien und auf ihr Potential verweisen, bedeutendes Wissen und nützliche Einsichten zu gewinnen. Im Sinne demokratischer Verantwortung schließlich seien Evaluatorinnen und Evaluatoren gefordert, ihre Arbeit aus der „Ödnis der Wertneutralität“ zu befreien und wieder verstärkt auf das international bedeutsame Problem der Ungleichheit auszurichten.

Ausdrücklich verwies Greene in diesem Zusammenhang auf die Bedeutung, die Evaluatorinnen und Evaluatoren im Rahmen eines Evaluationsprozesses als Person zukomme. Indem sie maßgeblichen Einfluss auf die Evaluationsergebnisse nähmen, fungierten sie in gewisser Weise als Konstrukteure bzw. Konstrukteurinnen der Welt. Deshalb sei es auf der einen Seite notwendig, Standards und Richtlinien zu entwickeln, um besser mit der Unsicherheit und Individualität des Evaluationsprozesses umgehen zu können. „But, we should also encourage evaluators to ‚have a life‘ and to ‚have a dream‘.“

Auch Frans L. Leeuw beschäftigte sich mit der Rolle der Evaluation in einer demokratischen Gesellschaft. Doch fiel seine Bilanz der derzeit betriebenen Evaluationspraxis durchaus kritisch aus. Bislang beschränke sich das öffentliche Bewusstsein vor allem auf die positiven Effekte von Evaluationen; anders sei der momentan zu beobachtende „Boom“ der Evaluationsforschung und -praxis kaum zu erklären. Doch sei die Evaluation schon lange keine „baby industry“ mehr – und die Frage nach dem Verhältnis von Kosten und Nutzen, wie sie sich etwa in der Policy-Analyse oder im Auditing seit längerem stelle, gewinne möglicherweise auch für die Evaluation an Bedeutung.

Ohne Zweifel, so Leeuw, habe Evaluation zu einer verstärkten Transparenz und Rechenschaftslegung geführt. Darüber hinaus wirke sie anregend auf das organisationale Lernen von Regierungen und anderen öffentlichen Organisationen, fördere die Demokratisierung und erzeuge Dissens, was in einer demokratischen Gesellschaft durchaus zu begrüßen sei. Schließlich stimuliere sie Experimentierbereitschaft und garantiere stabile, ebenso systematische wie bedarfsorientierte Wissensströme.

Demgegenüber stünden jedoch auch deutliche Kosten. So habe Evaluation zwar durchaus zu mehr Transparenz, aber auch zu mehr Verfahrenorientierung, Regulation und Kontrolle geführt. Die derzeitige Praxis führe zu einer „Routinisierung“. Ungenaue Rekonstruktionen der Policy- bzw. Interventionstheorie verursachten darüber hinaus Fehlerkosten. Leeuw verwies auf die von Wildawsky und Sunstein formulierten Ansprüche, Evaluation solle den Mächtigen Wahrheit vermitteln und Dissens produzieren – und stellte zugleich die Frage, was von diesen Ansprüchen geblieben sei.

Zugleich thematisierte er ein Problem, dass offensichtlich auch einige seiner Zuhörerinnen und Zuhörer beschäftigte: die finanzielle Abhängigkeit zahlreicher Evaluierenden von ihren Auftraggebern. Leeuw stellte zumindest als offene Frage in den Raum, welche Auswirkungen diese auf die Qualität und die Unabhängigkeit der Ergebnisse habe. Auf die Rückfrage aus dem Plenum, welchen Ausweg er aus diesem Dilemma sehe, empfahl er, bei eindeutigen Fällen Aufträge abzulehnen.

Leeuw schloss mit drei Vorschlägen für die zukünftige Evaluationsforschung und -praxis: Zum ersten solle sich die Evaluationsforschung, statt immer weitere Prozessanalysen, Leistungsvergleiche sowie Studien zur Implementierungstreue vorzulegen, wieder stärker auf die Evaluation von Wirkungen konzentrieren. Zum zweiten solle sie ihre Kräfte bündeln und bestehende Paradigmenkonflikte entschärfen. Zum dritten müsse die Erklärungskraft von Evaluationen deutlich steigen; eine Rückbesinnung auf die Theorie sei dringend erforderlich.

Den Abschluss der Tagung bildete der Beitrag „The Responsibility of Evaluation“ von Linda Mabry, kommentiert von *Wilfried Bos*, dem Leiter des Instituts für Schulentwicklungsforschung der Technischen Universität Dortmund. Mabry griff in ihrem Vortrag auf die Entwicklung der Evaluationsforschung und -praxis in den USA sowie die dort langjährig gemachten Erfahrungen insbesondere in Bezug auf „standardized testing“ zurück und beschäftigte sich mit der Frage, welche Verantwortung Evaluatorinnen und Evaluatoren der Gesellschaft gegenüber übernehmen müssten und könnten.

Mabry kritisierte, dass derzeit ein starker Fokus auf einer technischen, datenorientierten Perspektive von Verantwortung liege. Hiernach sei es lediglich die Aufgabe von Evaluationen, Daten zu berichten und objektive Informationen bereitzustellen, um Programme und Maßnahmen, die dem Wohl der Gesellschaft dienen sollen, zu verbessern. Eine weitergehende Verantwortung der Evaluierenden würde jedoch nicht erkannt. Demgegenüber bestehe aber ebenso die Möglichkeit, Ergebnisse in einer zwar wertgeladenen, aber sozial gerechten und Fairness fördernden Weise darzustellen. So könnten Ergebnisse einer Evaluation standardisierter Tests, die laut Mabry zu Lasten armer Schülerinnen, Schüler und Schulen ausfallen, zumindest Veränderungen anregen oder zu einer negativen Einschätzung dieser Verfahren führen. Eine weitere Alternative stelle eine postmoderne Perspektive dar, nach der Machtstrukturen identifiziert und Unterdrückungen adressiert werden (affirmative postmoderne Perspektive), oder – in einer extremen Position – Evaluierende sich aufgrund bestehender Unterdrückungsprozesse weigern, das Programm oder die Maßnahme zu evaluieren.

Die Frage, ob Evaluierende sich bei ihrer Arbeit von ihren Wertvorstellungen leiten lassen und diese den Auftraggebern mitteilen sollten bzw. inwieweit bestehende Wertsysteme überhaupt außer Acht gelassen werden könnten, werde überaus kontrovers diskutiert. Dabei sei zu berücksichtigen, dass eine vielfache Verantwortung der Evaluatorinnen und Evaluatoren bestehe: gegenüber sich selbst, der Profession, den Auftraggebern, anderen Stakeholdern des Evaluationsgegenstandes sowie der Gesellschaft. Hieraus resultierten sowohl professionelle als auch soziale Aspekte, die es zu klären gelte. Zum Teil unter Rückgriff auf Positionen auch anderer Evaluationsforscher und -forscherinnen formulierte Mabry abschließend mehrere Kernfragen: Sollte professionelle Verantwortung mit der Ergebnisdarlegung und -bewertung enden (Michael Scriven)? Sollten Evaluatorinnen und Evaluatoren bei der Anwendung und Umsetzung der Ergebnisse helfen (Michael Quinn Patton)? Sind Evaluationen, die auf „Empowerment“ abzielen, überhaupt Evaluationen (David Fetterman)? Ist die Erwar-

tung von partizipativer Evaluation, dass Stakeholder bei der Evaluation adäquat einbezogen werden können, realistisch (Jennifer Greene)? Ist die Veränderung von Machtstrukturen ein angemessenes Evaluationsziel? Und: Ist eine Evaluation, die demokratische Werte verfolgt, politische Einflussnahme (Ernest House)?

Am Beispiel der American Evaluation Association (AEA) zeigte Mabry einen möglichen Umgang mit diesen Fragen auf: So spreche sich die AEA in Form von öffentlichen Statements gegen „high stakes testing“ und für vielfältige Formen der Rechenschaftslegung von Schulen aus und wirke einer Einflussnahme auf die Wahl der Evaluationsmethoden von Seiten der Regierung entgegen. Aber die Tatsache, dass unter den Vertreterinnen und Vertretern der Profession keine Einigung darüber bestehe, was genau Evaluationen seien und leisten sollen bzw. können, mache, so Mabry abschließend, die Aufgabe von Evaluatoreninnen und Evaluatoren nicht einfacher.

Als empirischer Bildungsforscher bezog sich Bos in seinem anschließenden Kommentar primär auf die von Mabry angesprochene technische Perspektive der Verantwortung von Evaluationen. Für ihn liege die Verantwortung der Forschung darin, klare methodologische Standards zu verfolgen sowie Objektivität, Reliabilität und Validität stets als notwendige Voraussetzungen von Forschungsprozessen zu berücksichtigen und – so weit möglich – einzuhalten. Diesbezüglich sei in den vergangenen Jahren ein erheblicher Fortschritt zu verzeichnen, insbesondere durch die beständige und kritische Weiterentwicklung hoher Standards für die Verknüpfung von Messtheorie und empirischer Forschung.

Dennoch, so Bos, erschöpfe sich die Verantwortung empirischer Bildungsforscherinnen und -forscher nicht in der Einhaltung methodologischer Standards. Die Fragen danach, welche Form von Forschung betrieben werden kann, welche Fragen gestellt und welche Ergebnisse kommuniziert werden (können), seien stark abhängig von gegebenen Strukturen und den Bedingungen, unter denen geforscht wird. Die Hauptverantwortung der Forschenden brachte Bos in einem kurzen Satz zum Ausdruck: „Being a community of critical friends“. Er sah eine zukünftige Aufgabe für Wissenschaftlerinnen und Wissenschaftler darin, Diskussionsforen nicht nur zur Ergebnisverbreitung, sondern verstärkt zur Reflexion der von außen gegebenen Einflussfaktoren zu nutzen.

Die näher dargestellten Keynotes wie die Abschlussbeiträge waren dabei nur ein Teil des breiten Angebotes der gemeinsamen Tagung von DeGEval und KBBB. Breit gefächerte Workshops thematisierten über 50 weitere unterschiedliche Aspekte der Evaluationsforschung und -praxis. Nachwuchswissenschaftlerinnen und -wissenschaftler hatten die Möglichkeit, sich darzustellen und weiter zu qualifizieren. Spezifische Seminarangebote beschäftigten sich mit der Planung und Steuerung von Evaluationsprojekten, mit der Wirkungsevaluation Sozialer Dienstleistungen sowie mit dem Einsatz quantitativer und qualitativer Methoden in der Evaluationsforschung. Darüber hinaus nutzten zahlreiche Forschungsgruppen sowie Evaluationspraktiker und -praktikerinnen die Gelegenheit, ihre Projekte in Posterform zu präsentieren und mit einem ebenso fachkundigen wie interessierten Publikum zu diskutieren.

Zwar hatten DeGEval und KBBB ihre Tagungsangebote jeweils eigenständig organisiert. Doch standen die Vorträge grundsätzlich allen Tagungsteilnehmenden offen. Dies trug maßgeblich zum Austausch zwischen den beiden veranstaltenden Gesellschaften und ihren Mitgliedern bei.

Alle Keynotes sowie eine Vielzahl weiterer Beiträge sind auf der Homepage der DeGEval dokumentiert und laden zum Nachlesen ein.