

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

Bericht zur 19. Jahrestagung der DeGEval –
Gesellschaft für Evaluation e.V. in Salzburg

Ulrike Greiner,¹ Andreas Paschon,² Franz Riffert²

Vom 21. bis 23. September 2016 waren die *School of Education* und der *Fachbereich Erziehungswissenschaft* der Universität Salzburg Gastgeber der 19. Jahrestagung der DeGEval – Gesellschaft für Evaluation e.V. in Salzburg. Über 320 Teilnehmer(innen) aus Deutschland, der Schweiz, Russland und Österreich setzten sich in fünf Sessionblöcken und einer moderierten Postersession intensiv und auf vielfältige Weise mit dem Thema „Zwischen Nutzung, Einfluss und Nachhaltigkeit – Wie wirken Evaluationen in unterschiedlichen Systemen“ kritisch auseinander. Neben der Möglichkeit zur Präsentation eigener Forschungsprojekte und dem Austausch von Forschungsergebnissen konnte die Tagung auch zur Weiterbildung genutzt werden.

Das Tagungsthema signalisierte, dass „Evaluation“ bereits in Phase 3 ihrer Entwicklung als professionelle Praxis und als Forschungsprogramm angekommen ist: Nach der „Gründungsphase“, nämlich der Installierung von Evaluation als professionellem Blick auf (organisationales) Handeln, und der Optimierung von Evaluation mittels Qualitätsstandards in Phase 2 folgte nun der Blick auf Kontexte und Wirkungen, und dies zumeist in komplexen Mehrebenensystemen mit teils widersprüchlichen Nutzungserwartungen.

Die vielfältigen damit verbundenen Ansprüche wurden in unterschiedlichen Formaten wie z.B. Diskursen, Projektdarstellungen, Entwicklungsberichten und forschungsmethodischen Workshops lebhaft diskutiert. Evaluation wurde sozusagen einem Selbstevaluationsprozess unterzogen. Wie sehen die Verarbeitungsmuster von Resultaten externer und interner Evaluationen in den unterschiedlichen sozialen Teilsystemen Bildung, Kultur, Wirtschaft, Verwaltung, Politik, etc. aus? Wie hängen die Ergebnisse von Evaluationen, deren Präsentation und deren Kommunikation mit ihrer Nutzung zusammen? Wie werden Evaluationsergebnisse in Entscheidungspro-

1 School of Education, Universität Salzburg

2 Fachbereich Erziehungswissenschaft, Universität Salzburg

zessen gewinnbringend berücksichtigt? Welche unterschiedlichen Qualitätsinteressen können die beteiligten Akteure haben und wie können diese gegebenenfalls aufeinander abgestimmt werden? Und haben wir adäquate Instrumente, um auch langfristige Auswirkungen zu erfassen?

Die Angebots- und Austauschstruktur der Tagung zeugte von gelebter Vielfalt und intelligenter Differenzierung von Fragestellungen, Akteurskonstellationen und methodischen Formaten.

Am Mittwochnachmittag wurde mit parallelen Weiterbildungsseminaren der Fokus, wie in Vorlaufveranstaltungen sehr oft üblich, auf anwendungsorientierte Auseinandersetzung mit einem Schwerpunkt auf Forschungsmethodik (Grounded Theory, computergestützte qualitative Inhaltsanalyse, Wirksamkeitsmessung im Rahmen von Evaluation) gelegt.

Die Mitgliederversammlung an diesem ersten Abend stellte zugleich den inoffiziellen Start des gemeinschaftlichen Unternehmens ‚Jahrestagung‘ dar.

Am Donnerstagvormittag wurde beim Treffen der 16 Arbeitskreise die komplexe Querstruktur der DeGEval. sichtbar: von klassischen Bildungsthemen wie Schulen, berufliche Bildung, Hochschulen, über Gesundheitswesen, soziale Dienstleistungen, Kultur und Kulturpolitik, Stadt- und Regionalentwicklung, Verwaltung, Forschungs-, Technologie- und Innovationspolitik, Strukturpolitik, Entwicklungspolitik, Umweltbereich sowie Wirtschaft, bis hin zu Methoden, Professionalisierung und Gender Mainstreaming.

Daran anschließend wurde die Tagung formell durch Herrn Dekan a.o. Univ.-Prof. Dr. Weichbold von der Kultur und Gesellschaftswissenschaftlichen Fakultät der Universität Salzburg eröffnet.

Die Keynote im Anschluss, gehalten von o. Univ.-Prof. Dr. Herbert Altrichter, war in diese offizielle Tagungseröffnung eingebettet. Altrichters engagierte Auseinandersetzung mit den aktuellen Entwicklungen des österreichischen Bildungswesens an der Schnittstelle von Steuerung und Evaluation zeigte das Bildungssystem als paradigmatischen Ort der Spannungen (Akteure, Inklusion, Kommunikation, Paradoxien, nicht intendierte Wirkungen, usw.). Es gab kontroverse Debatten zum Evaluationsbegriff. Nicht alle Forscherkollegen und –kolleginnen teilten den teils ernüchterten Blick auf den möglichen Nutzen von makroanalytisch ausgerichteter Evaluation auf das Mikrosystem (Einzelschule).

Ab Donnerstag am Nachmittag begannen nach den Posterpräsentationen die themengebundenen Sessions (mit den Sessionblöcken A, B und C – mit jeweils sechs Parallelgruppen und jeweils wiederum je drei Beiträgen). Die Gruppierung ‚Meet the Authors‘ war eine (neue) erfolgreiche Session des Nachwuchsnetzwerkes: Autor(inn)en präsentierten neue Veröffentlichungen und gingen ins Gespräch.

Am Freitagvormittag wurden die Sessionblöcke D und E durchgeführt; pointiert mit den Themenperspektiven ‚Erfüllte oder enttäuschte Wirkungserwartungen‘, ‚Methodische Standards‘ und ‚Nutzung in politischen Prozessen‘.

Vor dem Abschlussgespräch wurden die Verleihungen der Nachwuchs- und Posterpreise vorgenommen. In der rege besuchten Preisverleihung wurden folgende Gewinner(innen) mit folgender Themenstellung ausgezeichnet: Dr. Carolin Frei-

er und Dr. Peter Kupka für ihr Poster „Der ‚professionelle Fremde‘ artikuliert das Unsagbare“.

Die abschließende Podiumsdiskussion brachte Wissenschaftlerinnen und Wissenschaftler mit unterschiedlichen methodischen Perspektiven an einen Tisch, aus der klassisch sozialwissenschaftlich quantitativen Forschung und aus eher kulturwissenschaftlich-rekonstruktiven Zugängen: Wolfgang Böttcher, Reinhard Stockmann, Herbert Altrichter, Philipp Mayring, den Vorstandsvorsitzenden, sowie Ulrike Greiner von der *School of Education* der Universität Salzburg. Es wurde durchaus kontrovers debattiert. Geht das wissenschaftlich fundierte Evaluationsverständnis von einer Handlungsrationalität aus, die so in manchen sozialen Teilsystemen gar nicht auffindbar bzw. rekonstruierbar ist?

Mögliche Schlussfolgerungen aus verhandelten Fragestellungen, welche während der Tagung in unterschiedlichen Sessions aufgeworfen worden waren, führten bereits zu möglichen Themenstellungen für künftige Jahrestagungen.

Im Rahmen der Generalversammlung der DeGEval wurde auch – nach mehrjähriger Auf- und Umgestaltung durch eine Arbeitsgruppe – die Revision der Evaluationsstandards beschlossen, was zeigt, dass auch die DeGEval ihre Arbeitsgrundlagen und -prinzipien immer wieder einer formativen Analyse unterzieht und dann evidenzbasiert Nachbesserungen und Updates vornimmt. Die neuen Evaluationsstandards werden schon bald auf der Homepage abrufbar sein: <http://www.degeval.de/degeval-standards/>

Die Momente des sinnlichen, intellektuellen und künstlerischen Genusses in der wunderschönen Universitätsstadt Salzburg waren laut der überwiegenden Mehrheit der Teilnehmerinnen und Teilnehmer durchaus gegeben.

In Erinnerung blieben gemeinsame opulente Abendessen im Stieglkeller in der Festungsgasse, die Stadtführung, die wunderbaren Kaffees und Buffets im Bibliotheksfoyer des Uniparks, der frühherbstlich-milde Gesamteindruck auf der prächtigen Dachterrasse des Uniparks mit Blick auf die Festung Hohensalzburg und die umliegenden Berge.

Die Teilnehmerinnen und Teilnehmer scheinen es genossen und sich auch wirklich gut verstanden, ja gemocht zu haben, kein unwesentliches Ziel einer Jahrestagung einer Gesellschaft, die der nachhaltigen Regenerationsfähigkeit von sozialen Systemen verpflichtet ist. Die Jahrestagung 2016 der DeGEval war ein Erfolg und – aus der Sicht der Gastgeber – für uns eine schöne Erfahrung des kollegialen Austausches, des Lernens und der fruchtbaren Debatte über das zentrale Thema der Wirksamkeit von Evaluation.

Das Salzburger Team

Nachwuchspreis der DeGEval 2016

*Edith Halves, Christiane Kerlen, Michael Köhler, Dörte Schott,
Sonja Sheikh*

Der DeGEval-Nachwuchspreis wurde auf der diesjährigen 19. Jahrestagung der DeGEval in Salzburg zum zehnten Mal vergeben. Ziel des Preises ist es, jährlich eine Nachwuchsevaluatorin/einen Nachwuchsevaluator oder auch eine Gruppe von Nachwuchsevaluator(inn)en auszuzeichnen und damit eine herausragende Arbeit im Bereich Evaluation im deutschsprachigen Raum zu würdigen. Die prämierte Arbeit soll einen fundierten Beitrag zur fachlichen und professionellen Weiterentwicklung von Evaluation leisten. Aus Sicht der DeGEval soll dieser Preis die Bedeutung des wissenschaftlichen Nachwuchses auf dem Gebiet der Evaluation sowie der Nachwuchsförderung insgesamt hervorheben, der sich die DeGEval in besonderem Maße verpflichtet fühlt.

Da der Preis in Salzburg zum zehnten Mal vergeben wurde und es sich somit um ein kleines Jubiläum handelte, nahmen wir dies zum Anlass, kurz zurückzublicken. Insgesamt hatten wir in den vergangenen zehn Jahren knapp 30 Einreichungen für den DeGEval-Nachwuchspreis. Bei dem Großteil davon (80%) handelte es sich um Dissertationen zum Thema Evaluation, der Rest waren Masterarbeiten und Evaluationsberichte aus der Praxis. Die eine Hälfte der Arbeiten wurde von Frauen, die andere Hälfte von Männern eingereicht. Zehn dieser Einreichungen wurden bisher ausgezeichnet (zehn deshalb, da es in einem Jahr zwei Preisträger(innen) gab). Darunter waren acht Dissertationen, eine Praxisarbeit und eine Masterarbeit. Fünf der bisherigen DeGEval-Nachwuchspreise wurden an Frauen, fünf an Männer vergeben.

Soweit der Rückblick. In diesem Jahr wurden insgesamt fünf Beiträge für den Nachwuchspreis 2016 eingereicht: zwei Dissertationen, zwei Masterarbeiten sowie ein Fachartikel. Und es muss dazu gesagt werden, dass alle Arbeiten im Vergleich zu den Einreichungen noch vor einigen Jahren ein enorm hohes Niveau aufweisen. Fast alle Arbeiten hätten, wären sie in einem anderen Jahr eingereicht worden, einen Preis erhalten können. Daher ist es der Jury in diesem Jahr ausgesprochen schwergefallen, eine Entscheidung für einen Preistragenden herbeizuführen. Aber wir mussten uns entscheiden: Der Preis ging diesmal – erstmals in den vergangenen fünf Jahren – an einen Fachartikel.

Der Artikel mit dem Titel „What Do We Know About the Demand for Evaluation? Insights From the Parliamentary Arena“ fragt nach den Gründen für die Nachfrage nach Evaluationen in Parlamenten und zwar am Beispiel des Schweizer

Parlaments. Anhand einer Befragung sämtlicher Parlamentsmitglieder der schweizerischen Kantone und des Bundes werden in dem Artikel u.a. Einstellungen gegenüber Evaluation untersucht und wie Evaluationen ausgelöst und später genutzt werden. Das Ganze wird untermauert durch aus der Prinzipal-Agent-Theorie hergeleiteten Hypothesen, die mittels einer anspruchsvollen Mehrebenenanalyse geprüft werden.

Aus Sicht der Jury werden mit diesem Thema gleich zwei Forschungslücken geschlossen. Zum einen wissen wir sehr wenig über die Entstehung von Evaluationen (vergleichsweise etwa zur Nutzung) und zum anderen war die Rolle der Evaluation in Parlamenten bisher kaum Gegenstand einschlägiger Forschung. Der Autor diskutiert dabei nicht nur die Relevanz der Befunde (über den schweizerischen Einzelfall hinaus), sondern reflektiert auch in souveräner Weise kritisch die Aussagekraft der Analyse. Die Jury war begeistert von dem Artikel, der auf Arbeiten des Autors an seiner Doktorarbeit beruht, die er an der Universität Zürich im Rahmen des Forschungsverbundes „Policy Evaluation in the Swiss Political System – Roots and Fruits“ verfasst.

Und damit ging der diesjährige zehnte DeGEval Nachwuchspreis 2016 an Herrn Pirmin Bundi. Der Artikel von Herrn Bundi ist vor kurzem in der renommierten internationalen Evaluationszeitschrift „American Journal of Evaluation“ erschienen. Wir gratulieren Herrn Bundi ganz herzlich zu dieser hervorragenden, theoretisch wie empirisch fundierten Forschungsarbeit und hoffen auf zahlreiche weitere Beiträge von ihm zur Praxis und Theorie der Evaluation!

Durch die Verleihung des Preises signalisiert die DeGEval ihr Interesse an akademischer und praktischer Forschung im Bereich Evaluation in unterschiedlichen thematischen Feldern und insbesondere auch an der Förderung des Nachwuchses im Bereich Evaluation. Wir freuen uns sehr, dass wir den Preis in diesem Jahr zum ersten Mal in die Schweiz verleihen konnten.

Herzlichen Glückwunsch!!!

Die Jury des Nachwuchspreises 2016

What Do We Know about the Demand for Evaluation? Insights from the Parliamentary Arena

Kurzfassung des mit dem DeGEval-Nachwuchspreis ausgezeichneten Artikels¹

*Pirmin Bundi*²

Im modernen Staat des 21. Jahrhunderts hat die Evaluation einen festen Platz eingenommen. Öffentliche Ausgaben stehen in Zeiten von hoher Staatsverschuldung immer stärker unter Beobachtung. Evaluationen dienen deshalb immer häufiger als wichtige Informationsquelle für politische Entscheidungsträger, um die Wirksamkeit und Wirtschaftlichkeit einer öffentlichen Maßnahme zu beurteilen. Während der letzten zwanzig Jahre hat die Evaluation im Zuge der internationalen Verbreitung der Evaluationstätigkeit stark an Bedeutung gewonnen, weshalb es nicht erstaunt, dass viele vom Zeitalter der Evaluation reden.

Obwohl die Regierungen jährlich eine beträchtliche Menge an finanziellen Mitteln für Evaluationen ausgeben, ist die Frage nach deren Ursprung bislang kaum untersucht worden, da sich die Evaluationsforschung vor allem mit der Nutzung von Evaluationen beschäftigt hat. Innerhalb des politischen Systems ist das Parlament ein wichtiger Stakeholder von Evaluationen, da diese für die Parlamentsmitglieder besonders nützlich sein können. Evaluationen versorgen Parlamentsmitglieder nicht nur mit wichtigen Informationen für die Gesetzgebung, sondern unterstützen diese auch in ihrer Aufsichtstätigkeit. Es kommt deshalb immer häufiger vor, dass Parlamente von sich aus Evaluationen anstoßen. Parlamentsmitglieder können Evaluationen vor allem mithilfe von parlamentarischen Vorstößen anregen, welche die Regierung damit beauftragen, eine bestimmte Maßnahme, ein Projekt oder gar ein ganzes Programm zu evaluieren. Trotz der hohen Relevanz der Evaluationen in den Parlamenten gibt es allerdings kaum Studien, welche die Nachfrage nach Evaluationen durch die Parlamente in den Mittelpunkt ihrer Analyse gestellt haben.

Vor diesem Hintergrund untersucht dieser Artikel, wodurch sich die Evaluationsnachfrage durch Parlamentsmitglieder erklären lässt. Dabei geht die Studie vor allem der Frage nach, weshalb Evaluationen angestoßen werden und welche Fakto-

1 American Journal of Evaluation, 37 (4), S. 522-541

2 Institut für Politikwissenschaft, Universität Zürich

ren diesen Vorgang unterstützen. Die Studie fokussiert sich auf die Schweizer Parlamente der kantonalen und nationalen Ebene. Zunächst galt die Schweiz als Nachzügler bezüglich der Institutionalisierung der Evaluation. Dies hat sich allerdings in den letzten Jahren stark verändert, besonders in den Parlamenten. In den frühen 1990er Jahren wurde die Parlamentarische Verwaltungskontrolle geschaffen, die im Auftrag des Bundesparlamentes Evaluationen durchführt. Weiter wurde mit der neuen Bundesverfassung von 1999 der Artikel 170 eingeführt, der die Bundesversammlung damit beauftragt, die Maßnahmen des Bundes auf ihre Wirksamkeit überprüfen zu lassen. Auch viele Kantone kennen mittlerweile eine solche Evaluationsklausel in ihrer Verfassung. So zählt die Schweiz heute zu denjenigen Ländern, in welchen die Evaluation den höchsten Stellenwert in den Parlamenten genießt.

Die Studie argumentiert, dass die Parlamentsmitglieder vor allem Evaluationen anregen, um ihre Oberaufsichtsfunktion gegenüber der Verwaltung wahrzunehmen. Grundsätzlich zeichnet sich der Policy Prozess der westlichen Demokratien durch eine Delegationskette aus, das heißt diejenigen, die autorisiert sind, eine Entscheidung zu treffen, beauftragen andere, diese für sie zu fällen. Die Parlamentsmitglieder delegieren die Umsetzung der Gesetzgebung der Regierung bzw. den Verwaltungseinheiten, die im Sinne des Parlamentes handeln sollen. Dabei entsteht eine typische Prinzipal-Agent-Situation, die jedoch zwei Probleme mit sich führt. Einerseits besitzen die Parlamentsmitglieder oftmals nicht genügend Informationen, um die Umsetzung der Gesetzgebung durch die Verwaltungen einschätzen zu können (Asymmetric Information). Andererseits sind die Verwaltungseinheiten heutzutage vermehrt auch politische Akteure und haben ein Interesse, die Politik in ihrem Sinne zu gestalten (Bureaucratic Drift). Aus diesen Gründen können die Parlamentsmitglieder nicht sicher sein, ob die von ihnen formulierte Politik auch in ihrem Sinne umgesetzt wird. Deshalb haben die Parlamente Anreize, Informationen über die Politikumsetzung zu sammeln und die Regierung und deren Verwaltungseinheiten zu beaufsichtigen.

Allerdings haben die Parlamentsmitglieder häufig Schwierigkeiten, an die nötigen Informationen zur Kontrolle des Gesetzesvollzugs zu gelangen. Aufgrund ihrer begrenzten Ressourcen ist eine permanente und flächendeckende direkte Überwachung durch das Parlament nicht möglich. In der Schweiz trifft dies besonders stark zu. Einerseits werden die Parlamente durch die direkt-demokratischen Instrumente in ihrer Macht eingeschränkt, besonders auf der kantonalen Ebene, die noch eine stärkere Volksbeteiligung als die Bundesebene kennt. Zudem sind die Aufsichtskapazitäten des Parlamentes limitiert. Die institutionell formalisierte Aufsicht beschränkt sich praktisch auf die Aufsichtskommissionen, die regelmäßig Inspektionen und Berichte publizieren. Die Empfehlungen der Berichte sind jedoch nicht bindend, weshalb die Regierungen oftmals keine unmittelbaren Maßnahmen beschließen. Zudem sind die Schweizer Parlamente – auch auf Bundesebene – noch immer Milizparlamente, weshalb die meisten Parlamentsmitglieder neben ihrem Mandat einer anderen Beschäftigung nachgehen und sich bei der Parlamentsarbeit eher auf die Gesetzgebung fokussieren.

Evaluationen nehmen deshalb für die Aufsichtsfunktion der Parlamente eine wichtige Rolle ein, da die Verwaltung während des Evaluationsprozesses Informa-

tionen zur Umsetzung einer Politik liefert und dem Parlament Rechenschaft ablegt. Parlamentsmitglieder können selektiv und ressourcensparend ihre Aufsichtsfunktion ausüben, indem sie mit einem parlamentarischen Vorstoß eine Evaluation zu einem spezifischen Thema anregen. So benötigen sie nicht zwingend institutionell starke Aufsichtsorgane, sondern können Evaluationen fordern, um die Regierung, bzw. die Verwaltung, zur Rechenschaftslegung anzuhalten. Evaluationen ermöglichen den Parlamentsmitgliedern somit ihre fehlenden Aufsichtskapazitäten zu einem gewissen Grade zu kompensieren und trotzdem die Exekutive zu kontrollieren.

Um zu überprüfen, ob Parlamente Evaluationen für die Erfüllung ihrer Aufsichtstätigkeit anregen, wurde im Frühjahr 2014 eine Online-Befragung unter den Parlamentsmitgliedern des Bundes und der Kantone durchgeführt. Dabei wurden die Parlamentsmitglieder nach ihrem Umgang sowie ihren Erfahrungen mit Evaluationen gefragt. Insgesamt beteiligten sich 1.570 Parlamentsmitglieder an der Befragung, was einer Teilnahmequote von 55,3% entspricht und für eine Parlamentsbefragung einen außerordentlich hohen Wert darstellt.

Die Befragung zeigt, dass viele Parlamentsmitglieder schon einmal nach einer Evaluation gefragt haben. Beinahe die Hälfte der Befragten haben angegeben, dass sie mindestens einmal eine Evaluation mit einem parlamentarischen Vorstoß angeregt haben. Ein Viertel der Befragten hat demnach bereits mehrere Evaluationen angestoßen. Gleichzeitig haben unter derjenigen Hälfte der Parlamentsmitglieder, die nach keiner Evaluation gefragt haben, am meisten Befragte angegeben, dass sie bisher keine passende Möglichkeit gefunden hätten oder dass die Verwaltung bereits genügend Evaluationen zur Verfügung stellen würde. Insgesamt zeigt sich damit ein äußerst evaluationsfreundliches Bild der Schweizer Parlamente.

Die Mehrebenenanalyse zeigt, dass individuelle Persönlichkeitsfaktoren wichtiger für die Evaluationsnachfrage sind als strukturelle Merkmale auf der Parlamentsebene. Die quantitative Analyse unterstützt dabei die These, dass Parlamentsmitglieder Evaluationen für die Ausübung ihrer Aufsichtsfunktion anregen. Diejenigen Parlamentsmitglieder, die den Eindruck haben, dass die Verwaltungseinheiten die Gesetzgebung nicht im Sinne des Parlamentes umsetzt, haben eine höhere Wahrscheinlichkeit eine Evaluation zu fordern. Im Gegensatz dazu spielt es jedoch keine Rolle, ob Parlamentsmitglieder genügend Informationen haben, um die Politikumsetzung beurteilen zu können. Des Weiteren haben die Mitglieder einer Aufsichtskommission eine signifikant höhere Wahrscheinlichkeit, einen Vorstoß zur Anregung einer Evaluation einzureichen. Das gleiche gilt auch für die Mitglieder des Parlamentsausschusses (die Vorsitzenden des Parlamentes und der parlamentarischen Gruppen) und die erfahrenen Parlamentsmitglieder, die oftmals eine höhere Sensibilisierung gegenüber der parlamentarischen Oberaufsicht besitzen. Schließlich spielt auch die Einstellung gegenüber Evaluationen eine wichtige Rolle, auch wenn jüngste Studien den Effekt vor allem der sozialen Erwünschtheit zuschreiben. Soziostrukturelle Faktoren sowie parlamentarische Kontextvariablen spielen hingegen kaum eine Rolle. Einzig eine Evaluationsklausel in der Verfassung hat einen positiven Effekt auf die Evaluationsnachfrage des jeweiligen Parlamentes.

Die Resultate bieten einen wichtigen Einblick in ein bisher vernachlässigtes Forschungsfeld. In den vergangenen Jahren wurde oftmals in der Literatur argumen-

tiert, dass die Evaluationsforschung sich von der Evaluationsnutzung zum Evaluationseinfluss verlagern sollte, um fortgeschrittene Auswirkungen und Konsequenzen von Evaluationen zu erfassen. Allerdings haben sich die bisherigen Studien überwiegend auf Informationen aus der Sicht der Evaluierenden beschränkt, die oftmals nicht besonders objektiv sind. Falls die Perspektive von Stakeholdern in der Studie berücksichtigt wurden, basierten die Informationen oftmals auf internen Berichten. Da Stakeholder häufig ein Interesse daran haben, als eine evidenzbasierte Organisation wahrgenommen zu werden, sind interne Berichte häufig geschönt. Diese Gefahr ist bei Parlamentsmitgliedern weniger gegeben, da empirische Evidenz nur eine mögliche Form der Entscheidungsfindung darstellt und andere Entscheidungsgrundlagen (Interessen, Ideologien, Institutionen) oftmals sogar über eine höhere Legitimität verfügen. Die Ergebnisse dieser Studie bieten deshalb einen Einblick in die parlamentarische Arena, die ein wichtiger Stakeholder von Evaluationen ist.

Die Ergebnisse beinhalten trotz der Eigenheiten des politischen Systems der Schweiz wichtige Implikationen für andere Parlamente. Die Studie zeigt, dass die Schweizer Parlamente besonders Evaluationen anregen, um ihre fehlenden Kontrollressourcen zu kompensieren. Auch wenn andere Parlamente möglicherweise über mehr Aufsichtskapazitäten verfügen, können Evaluationen eine wichtige Rolle für Parlamente einnehmen. Wenn Evaluationen besonders zur Kontrolle von Regierungen verwendet werden können, ist ein Einsatz der Evaluation als Oppositionsinstrument durchaus vorstellbar. Da in der schweizerischen Konsensdemokratie die meisten Parteien in der Regierung eingebunden sind, können keine Unterschiede zwischen der Opposition und den Regierungsparteien in ihrer Evaluierungsforderung beobachtet werden. In einer parlamentarischen Demokratie, die häufig über eine substantielle Opposition im Parlament verfügt, haben Parlamentsmitglieder aber sicherlich vermehrt Interessen, eine Evaluation als Kontrollinstrument einzusetzen.

Diese Schlussfolgerung ist nicht nur für die Forschung über Evaluationen besonders relevant, sondern auch für Evaluierende und Auftraggebende von Evaluationen. Carol Weiss hat vor 17 Jahren geschrieben, dass viele Evaluator(inn)en darüber erstaunt sind, dass die politischen Entscheidungsträger Evaluationsresultate nicht konsequenter nutzen. Dieser Artikel bietet für die parlamentarische Arena eine Erklärung für dieses Verhalten an. Parlamente scheinen sich mehr um die Funktion der Evaluation zu interessieren als um ihren Inhalt. Evaluationen bieten den Parlamentsmitgliedern eine Möglichkeit, die Arbeit der Verwaltung zu überprüfen, ohne selbst viel zu investieren. Die Ergebnisse implizieren deshalb eine wichtige Botschaft für Personen, die Evaluationen durchführen: Wenn eine Evaluation für ein Parlament relevant und einflussreich sein will, sollte sie eher auf Rechenschaftspflicht als auf das Lernen fokussieren. Evaluierende können so die Nutzung von Evaluationen in den Parlamenten verbessern, wenn sie die Bedürfnisse der Parlamentsmitglieder vermehrt in die Konzeption der Evaluation miteinbeziehen. Auch wenn Parlamente Evaluationsresultate vielleicht nur selten nutzen, nehmen Evaluationen trotzdem einen wichtigen Stellenwert für die politische Arbeit der Parlamentsmitglieder ein, denn Parlamente nutzen Evaluationen, nicht Evaluationsbefunde.

**„Evaluation (in) der Zukunft“
20. Jahrestagung der DeGEval –
Gesellschaft für Evaluation e.V.
vom 20. bis 22. September 2017
in Mainz**

Zum Anlass des 20. Jubiläums wird die DeGEval-Jahrestagung **den Blick zurück mit einem Blick in die Zukunft verknüpfen**. Wie hat sich seit Gründung der Gesellschaft das Handlungsfeld Evaluation in Praxis, Theorie und Forschung verändert? Welche Entwicklungen sind in den kommenden Jahren absehbar, zu befürchten oder zu erhoffen? Welche Rolle hat die DeGEval bei diesen Entwicklungen gespielt und welche wird sie in Zukunft einnehmen?

Unter Bezug auf einige pointierte Vorhersagen des gegenwärtigen Vorsitzenden unserer Schwestergesellschaft American Evaluation Society,¹ John Gargani, wollen wir in diesem Zusammenhang unter anderem folgende Fragen diskutieren:

- Wie verändert sich das **Verhältnis von interner zu externer Evaluation**? Werden Organisationen und Verwaltungen die Evaluation zunehmend internalisieren (Stichwort ‚Evaluation Capacity Development‘) oder wird der Bedarf nach Objektivität und Unabhängigkeit die Nachfrage nach externen Evaluationen hochhalten?
- Wie verändern sich die **Rolle von Evaluierenden** und die an sie herangetragenen Rollenerwartungen? Wie stark wird etwa von Evaluierenden verlangt werden, zunehmend gestaltend auf der Gegenstandsebene mitzuwirken, wie es etwa Pattons Ansatz der Developmental Evaluation vorsieht?
- Wie verändern die **sozialen Medien** die Evaluation und den sozialen Kontext, in dem sie agiert? Welche Möglichkeiten ergeben sich etwa für die Kommunikation in Evaluationen z.B. bei partizipativ angelegten Verfahren? Welche neuen Risiken ergeben sich in der digitalen Öffentlichkeit?
- Wie verändert sich unter den Bedingungen einer zunehmenden Digitalisierung das **Berichtswesen in Evaluation** hinsichtlich der medialen Aufbereitung (z.B. interaktive Online-Berichte, Data Visualisation), aber auch hinsichtlich der zeitlichen Taktung (‚just-in-time reporting‘)?
- Wie werden sich **Datenerhebung und Datenanalyse in Evaluationen** verändern? Welche Rolle werden Big Data, Data Mining oder zunehmend intelligente Auswertungsverfahren v.a. für qualitative Daten spielen?
- Welche weiteren Schritte zur **Professionalisierung der Evaluation** werden wir erleben? Wird es zu einer Vereinheitlichung der Fachsprache kommen? Wird es

1 Vgl. <https://evalblog.com/2012/01/30/the-future-of-evaluation-10-predictions/>

Schließungsanstrengungen geben, um Evaluation vor unqualifizierten Anbietern zu schützen? Welche Folgen haben internationale Ansätze zu standardisierten Anerkennungsverfahren für Evaluierende?

- Und schließlich: Wie verändert sich insgesamt die **Nachfrage nach Evaluationen**? Erleben wir nach 20 Jahren einer vielfach konstatierten steigenden Nachfrage (die sich auch in einem kontinuierlichen Wachstum der DeGEval spiegelt) derzeit eine Sättigung oder gar einen Gipfel vor dem Abstieg?

Neben diesen Fragen zur Zukunft der Evaluation wird ein zweiter Schwerpunkt der Tagung die **Zukunft als Gegenstand in der Evaluation** sein. Im Mittelpunkt steht dabei die Frage: Inwiefern kann und soll Evaluation valide Aussagen über die Zukunft treffen, obwohl sie sich als empirisch gestütztes Verfahren immer nur auf Vergangenes beziehen kann? Diese Frage berührt vielfache Aspekte der Evaluationstätigkeit:

- Aussagen über die Zukunft treffen Evaluationen immer dann, wenn sie die **Generalisierbarkeit von Befunden** feststellen, da damit eine Prognose hinsichtlich zukünftiger Wirkungen einer evaluierten Maßnahme getätigt wird. Unter welchen Bedingungen sind solche Prognosen tatsächlich tragbar? Warum scheitern positiv evaluierte Innovationen dennoch oft nach Einführung im Feld?
- Auch **Empfehlungen**, die in Evaluationsberichten ausgesprochen werden, enthalten immer eine Vorhersage in Form von prognostizierten Verbesserungen für die Zukunft. Auf welcher Basis macht Evaluation im Einzelfall solche Vorhersagen, vor allem, wenn sich Empfehlungen nicht zu 100% aus den empirischen Befunden ergeben? Wie geht Evaluation mit den dabei inhärenten Unwägbarkeiten um?
- Bei verschiedenen Evaluationsansätzen ist die Zukunft darüber hinaus genuine Perspektive der Betrachtung, indem sie Pläne, Anträge und sonstige in die Zukunft gerichtete Konzepte bewerten. Zu ihnen gehören die **Ex-ante-Evaluation**, die **präformative Evaluation** oder die **Evaluierbarkeitsanalyse**. Sie befassen sich genuin mit Fragen, die die Zukunft betreffen, etwa nach Bedarfslagen, Realisierbarkeit oder Erfolgchancen. Wie und in welchen Kontexten haben sich entsprechende Ansätze bewährt und wo besteht brachliegendes Potenzial? Wie verhalten sich die Ansätze untereinander und zu ähnlichen Verfahren wie der Zukunftsforschung, Foresight-Prozessen oder der Gesetzesfolgenabschätzung?

Nähere Informationen zur Tagung finden Sie unter: <http://www.degeval.de/veranstaltung/en/jahrestagungen/mainz-2017>.

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>

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

Positionspapier des Vorstandes der DeGEval –
Gesellschaft für Evaluation

Auf die zunehmende Bedeutung von Evaluation wird immer wieder hingewiesen. Evaluation will dabei häufig Wirkungen untersuchen, aber hat sie selbst in ausreichendem Maße Wirkung? Können Evaluationsergebnisse Gesellschaft oder Teile der Gesellschaft in intendierter Richtung verändern? Mit der Frage nach dem Nutzen von Evaluation ist die erste Kriteriengruppe der *Standards für Evaluation* angesprochen, die die *DeGEval – Gesellschaft für Evaluation* soeben in einer Überarbeitung vorgelegt hat (vgl. www.degeval.de). Dieser Standard der Nützlichkeit von Evaluation bezieht sich aber nur auf die Konzeption des Evaluationsprojektes, die Frage, inwieweit alle Beteiligten und Betroffenen einbezogen wurden und die Zwecke der Evaluation bestimmt wurden. Ob diese aber dann wirklich umgesetzt werden, das betrifft die hier angesprochene Frage des Nutzens und der Nachhaltigkeit von Evaluation.

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 vom 21. bis 23. September 2016 aufgenommen hatten. Die Bandbreite erstreckte 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 hat 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 beleuchtet.

Es haben sich dabei große bereichsspezifische Unterschiede gezeigt, je nachdem wie stark Evaluation im Praxisbereich verankert ist. Lassen Sie uns in exemplarische Bereiche genauer hineinschauen:

- Im Bildungsbereich, der aufgrund des Gastgebers der Jahrestagung, der *School of Education* der Universität Salzburg, diesmal im Vordergrund stand, ist in vielen Bereichen Evaluation und die Umsetzung der Evaluationsergebnisse fest verankert. An Schulen wurden zunächst Selbstevaluationsmodelle eingeführt. Die Teilnahme war vorwiegend freiwillig, so dass die Durchführung zum großen Teil auf dem Engagement, dem Willen zur Selbstvergewisserung und Kompetenzerweiterung der beteiligten Lehrerinnen und Lehrer aufbauen konnte und dadurch zu nachhaltigeren Wirkungen führen konnte. In den letzten Jahren wurde an Schulen verstärkt durch Kompetenzmessungen der Schülerinnen und Schüler evaluiert, was zwar zu aufrüttelnden Ergebnissen führte („Pisасhock“), nicht immer aber zu Veränderungen im Schulsystem beitragen konnte. Im Hochschulbereich sind in vielen Ländern Evaluationen gesetzmäßig verankert. In Österreich müssen beispielsweise Universitäten ihr Qualitätsmanagementsystem regelmäßig einer externen Überprüfung unterziehen, die zu Auflagen führen kann, die dann erfüllt werden müssen.
- Im Wirtschaftsbereich sind evaluative Maßnahmen, auch wenn das nicht immer als vollwertige Evaluation zu verstehen ist, vor allem im Bereich von Qualitätsmanagement zu finden. In aller Regel sind sie fest verankert in der Institution. Aufgrund ökonomischer Erwägungen hat das Unternehmen hier ein Interesse, die Ergebnisse direkt in Optimierungsaktivitäten münden zu lassen und damit Wirkung und Nachhaltigkeit sicherzustellen.
- Im Gesundheitsbereich wiederum sind feste Formen von Evaluationen (experimentell quantitative Studien in Form von ‚Randomized Controlled Studies‘) gesetzmäßig gefordert zur Zulassung von Medikamenten und Heilbehandlungen. Damit ist die Praxiswirksamkeit festgelegt. In anderen Bereichen wie beispielsweise der Evaluation von Gesundheitspräventionsmaßnahmen sind die Verfahren weniger verpflichtend und die Methodik ‚weicher‘ (z.B. Quasiexperimente). So bleibt hier der langfristige Nutzen der Maßnahmen oft im Unklaren.

Was lernen wir daraus? Wir denken, dass Nutzen, Einfluss und Nachhaltigkeit vor allem durch zwei Ansätze verstärkt werden können: Zum einen verstärken partizipative Elemente in der Evaluation deren langfristige Nützlichkeit. Die Beteiligten fühlen sich ernst genommen, formulieren selbst ein Interesse an den Evaluationsergebnissen, setzen die Evaluationsmethodik dort an, wo eigener Bedarf ist und können deshalb direkt an der Umsetzung der Ergebnisse arbeiten. Zum anderen wird langfristiger Nutzen von Evaluationen dann erleichtert, wenn die Evaluation im Praxisbereich fest verankert ist, zum Beispiel durch gesetzliche Verpflichtungen. Hier muss aber nicht nur die Durchführung der Evaluation, sondern auch die Umsetzung der Ergebnisse festgelegt werden.

Beide Ansätze sind durchaus gegensätzlich, oft nicht vereinbar, stellen damit unterschiedliche Wege dar, die in verschiedenen Praxisfeldern mehr oder weniger vielversprechend sein können. Beide Ansätze tragen aber zur stärkeren Verankerung von Evaluation in der Gesellschaft bei.

Zwischen Schema F und Innovation: Eine politikfeldübergreifende Diskussion zu methodischen Standards

*Thomas Altenburg*¹

Der Arbeitskreis (AK) Methoden in der Evaluation der DeGEval führte im Rahmen der DeGEval-Jahrestagung 2016 in Salzburg eine Session zum methodischen Erfahrungsaustausch durch. Mit dieser Auftaktveranstaltung wurde der Startschuss für eine politikfeldübergreifende Bestandsaufnahme der verschiedenen Evaluationsansätze und -standards innerhalb der DeGEval gegeben. Ziel ist es, Gemeinsamkeiten und Unterschiede in Evaluationsmethoden und methodische Trends in den jeweiligen Arbeitskreisen der DeGEval herauszuarbeiten, deren Ursachen festzustellen und mögliche Synergieeffekte oder Innovationen auszuloten.

Um einen intensiven Erfahrungsaustausch zu ermöglichen, wurde das Konzept der Fishbowl-Diskussion gewählt. Neben drei festen Diskussionsteilnehmerinnen

- *Vera Hennefeld* (Bereichsleitung Bildung und Kultur, CEval GmbH) – aktuell stellvertretende Sprecherin des AK „Kultur und Kulturpolitik“,
- *Sonja Sheikh* (stv. Direktorin der KMU Forschung Austria) – von 2005 bis 2012 Sprecherin des AK „Forschungs-, Technologie- und Innovationspolitik“ und
- *Susanne von Jan* (stv. Leiterin Evaluierung, Hanns-Seidel-Stiftung e.V.) – aktuelle Sprecherin des AK „Entwicklungspolitik“,

welche exemplarisch die Evaluationserfahrungen aus ihren Arbeitskreisen präsentierten, konnten hierdurch auch die Seminarteilnehmerinnen und Seminarteilnehmer (vorübergehend) in die Diskussionsrunde gelangen und ihre individuellen Sichtweisen und Erfahrungen schildern. Die Diskussion wurde von *Alexandra Caspari* und *Tobias Polak* (Sprecherteam des AK Methoden) moderiert, um sicherzustellen, dass möglichst viele Gäste am aktiven Gedankenaustausch teilnehmen konnten.

Bestandsaufnahme

Vera Hennefeld betonte in ihrem Eingangsstatement, dass im Arbeitskreis **Kultur und Kulturpolitik** zwei Handlungsfelder bestehen, welche getrennt voneinander

1 Berlin

betrachtet werden müssen: Evaluation von Kultur im Sinne von Projekten im Bereich Kunst und Kultur sowie Evaluation im Bereich Kulturpolitik im Sinne von (Auswärtiger) Kultur- und Bildungspolitik. Im Bereich Kultur wird im Vergleich zur Kulturpolitik weniger evaluiert. Gerade in der Kultur ist die Akzeptanz von Evaluation nach wie vor eine große Herausforderung: Notwendig ist hier oftmals die Aufklärung über die Ziele von Evaluation, welche ausdrücklich nicht eine Bewertung von Kultur und Kulturschaffenden selbst beabsichtigt, sondern die Analyse und Bewertung von einzelnen Projekten und deren Zielerreichung. Dieses von Gegnern der Evaluation gerne ins Feld geführte Missverständnis führt vermutlich zu den bestehenden Vorbehalten gegenüber Evaluationsprozessen. Im Bereich Kultur – so *Vera Hennefeld* – werden vornehmlich gängige qualitative Methoden (beispielsweise in Form von Fallstudien) angewandt, aber auch innovative quantitative und qualitative Ansätze werden verfolgt. Beispielsweise wurden kognitive Landkarten genannt, die auch auf andere Evaluationsfelder übertragbar sind. Gegenüber quantitativen Methoden existieren eher Vorbehalte. Eine Ausnahme bildet hier die weit verbreitete und gut etablierte Museumsevaluation und Besucherforschung.

In der Kulturpolitik sind hingegen wesentlich mehr Evaluationen zu beobachten. Ein Schwerpunkt sind klassische Querschnittsevaluationen zum Projektende, wobei großteils qualitative (z.B. Fokusgruppen, leitfadengestützte Interviews) oftmals in Kombination mit quantitativen Methoden, wie z.B. standardisierte Online-Befragungen, angewandt werden. Bislang werden jedoch kaum experimentelle bzw. quasi-experimentelle Designs realisiert. Auch Längsschnittanalysen sind eher eine Ausnahme.

Wünschenswert wäre in beiden Bereichen, wenn Evaluation verstärkt zum Start einer Maßnahme mitgedacht und so mehr Möglichkeiten zum Einsatz von anspruchsvolleren und rigoroseren Designs gegeben wären.

Im Handlungsfeld der **Forschungs-, Technologie- und Innovationspolitik** (FTI) gab *Sonja Sheikh* für den geografischen Bereich Österreich Auskunft – der sich stark von FTI-Evaluationen in Deutschland unterscheidet. Sie erläuterte zu Beginn den gleichwohl großen Evaluationsumfang, zum Teil wird gar für *ein* Programm eine Zwischen- und Endevaluation realisiert, welche dann bewusst von verschiedenen Evaluator(inn)en ausgeführt werden müssen. Im Fokus der Evaluation steht meist die Analyse des Gegenstandes, der Umsetzung sowie der Wirkungen, wobei letzteres das ‚Herzstück‘ ist. Der Fokus liegt hier auf einem Methodenmix aus Tiefeninterviews, Logik-Chart-Analysen, Standardbefragungen und/oder Fokusgruppen. Oft sind dabei sehr lange Zeiträume betroffen, wobei das Programmziel eher selten stringent dargestellt wird. Indikatoren für Programmwirkungen sind beispielsweise Publikationen, Veröffentlichungen oder Patente, die mittels einer Befragung erhoben werden. Experimentelle Designs kommen kaum zum Einsatz, obwohl diese unregelmäßig auch in Auftragsbeschreibungen benannt werden. Aus Ressourcengründen, aber auch mangels neutraler Kontrollgruppen, können solche Designs auch kaum realisiert werden. Ein aktueller Trend zeichnet sich derzeit dahingehend ab, dass häufiger dynamische Prozesse (beispielsweise Stakeholder-Workshops als World-Café) im Rahmen von Evaluierungen zum Einsatz kommen.

Im Politikfeld der **Entwicklungszusammenarbeit** (EZ) ist es – vor allem aufgrund der starken Heterogenität von Themen, Ländern und Akteuren – schwierig, von *einem* methodischen Standard zu sprechen, so *Susanne von Jan*. Dennoch können einige Evaluationsmethoden als typisch angesehen werden. Sie nennt hier das Dokumentenstudium (was sie von der Dokumenten*analyse* dezidiert abgrenzt) und qualitative Instrumente (Leitfadeninterviews, semistrukturierte Interviews, Gruppenbefragungen und -diskussionen). Aber auch standardisierte Interviews oder partizipative Methoden (PRA, Mapping) sowie Beobachtungen werden umgesetzt. Standardmäßig werden angestrebte Veränderungen gemessen, allerdings eher als Soll-Ist-Vergleich im Sinne von Zielerreichung, häufig auch über retrospektive Fragen. (Quasi-)Experimentelle Designs oder randomisierte, kontrollierte Studien zur Wirkungsmessung werden zwar auch in diesem Handlungsfeld realisiert, jedoch nicht als Standard. Als ein Zukunftstrend erscheint die vermehrte Erhebung und Nutzung von Baseline-Daten, so dass zukünftig sicherlich verstärkt Vorher-Nachher-Messungen zu erwarten sind. Als innovative Methodologien für die Zukunft gelten aktuell neue Technologien sowie die Nutzung von Big Data. So könnten Umfragen über mobile Endgeräte realisiert oder große Daten aus aktuellen Satellitenbildern (beispielsweise Nachtaufnahmen zur Auswertung der Helligkeit einzelner Regionen als Proxyindikator für Armut) genutzt werden. Im Zusammenhang unter anderem mit den SDG ist auch die stärkere Nutzung von Sekundärdaten in der Diskussion. Abschließend gab sie noch einen Hinweis zum Monitoring: Während das Aktivitäten-Monitoring durchaus stark ausgeprägt ist, stellt das wirkungsorientierte Monitoring viele Projekte noch vor große Herausforderungen.

Diskussion

Im Rahmen der anschließenden Diskussion wurden verschiedene Aspekte vertieft. Zunächst wurde betrachtet, **wer Evaluationen durchführt** und wie die **Kommunikation und Kooperation mit der Wissenschaft** aussieht. Die Referentinnen erläuterten beispielsweise, dass Evaluationen im Kulturbereich, in der Entwicklungszusammenarbeit und FTI-Politik meist durch freie Einzelgutachter aber auch außeruniversitäre Einrichtungen wie Consultings oder private Beratungsunternehmen und außeruniversitäre Forschungseinrichtungen durchgeführt werden. Eine große Ausnahme sind Universitäten, wodurch sich der wissenschaftliche Einfluss auf die Evaluationen schwer einschätzen lässt. Allerdings sind einige Consultingunternehmen an Universitäten angesiedelt oder die Akteure sind Professor(inn)en, die die Evaluationen allerdings nicht in ihrer universitären Rolle durchführen.

Ein zentraler Diskussionspunkt betraf die Frage der **Methodenqualität**, d.h. wie ‚sauber‘ wird gearbeitet und was bedeutet ‚methodisch sauberes‘ Arbeiten. Hieraus ergaben sich auch die Frage der **Darstellung und Transparenz** der bei Evaluation genutzten Methoden sowie die Frage, inwieweit **Plausibilitätsüberlegungen** akzeptabel sind.

Dorothee Mack (Bischöfliches Hilfswerk Misereor e.V.) berichtete, dass eine Metaevaluation aus 2016 gezeigt hat, dass die genutzten Methoden selten dargestellt werden und auch in den meisten Berichten nicht erkennbar ist, dass Daten ‚sauber‘ analysiert wurden – hier ist „sehr viel Luft nach oben“. Die zentrale Frage lautet daher, was muss und was kann ein Bericht enthalten? Aus Auftraggebendensicht zeigt sich hier – so *Dorothee Mack* – ein großes Dilemma, da in einem ständigen Balanceakt die knappen Ressourcen entweder in die Methodendarstellung oder Ergebnisdarstellung investiert werden müssen: Die wünschenswerte Qualität in der Methodenanwendung und Datenauswertung treibt jedoch die Kosten schnell in die Höhe und steht dann in keinem Verhältnis mehr zu den Projektkosten. Daher muss pragmatisch vorgegangen werden. So wird durchaus in Auswertungsgesprächen eine sehr saubere Methodenanwendung deutlich, was dann aber aus Kostengründen in den Berichten nicht ausführlich dokumentiert wird. D.h. ‚sauberes Arbeiten‘ zeigt sich insbesondere in einer klaren Verbindung von beschriebener Datenlage zu Schlussfolgerungen, d.h. wie deutlich können einzelne Schlüsse wirklich aus den Daten gezogen werden. Häufig fehlen jedoch in Berichten Zahlen zu Stichprobengrößen. Dann bleibt offen, ob es sich um eine überzeugende Einzelmeinung handelt oder ob diese Meinung trianguliert wurde. Diese Verbindung ist in Berichten teilweise vorhanden aber nicht immer eindeutig dargestellt, wird dann aber in Gesprächen geklärt, so *Dorothee Mack*.

Erwin Geuder-Jilg (DEval) führte aus der Perspektive des freien Gutachters, welcher auch für Misereor tätig war, an, dass z.B. die Bestimmung von Stichprobengröße und Stichprobenauswahl nicht in zwei bis drei Wochen erarbeitet werden kann, sondern ausreichend Vorlaufzeit benötigt. Als positives Beispiel nennt er eine Evaluation in Brasilien, die zwei Missionen umfasste. So konnte im Rahmen der ersten Mission die Gesamtmenge bestimmt werden und daraus die Stichprobe systematisch und transparent herausgearbeitet werden. Ausreichend Zeit für die Entwicklung einer systematischen Methodik ist allerdings seiner Meinung nach sehr selten gegeben – systematisches Vorgehen ist also eine Zeit- und somit Geldfrage.

Susanne von Jan führt an, dass Evaluatorinnen und Evaluatoren aus ihrer Erfahrung häufig die in Berichten aufgeführten Schlussfolgerungen auch auf Nachfrage hin nicht mit Daten belegen oder begründen könnten und auf ‚logische‘ Plausibilitätsüberlegungen verweisen. Plausibilitätsüberlegungen sind ihrer Meinung nach allerdings nicht immer angemessen – hier kommt es auf die Fragestellung an. Wenn Wirkungen im Vordergrund stehen und in der Ausschreibung ein quasi-experimentelles Design erwartet wird, dann ist eine Plausibilitätsüberlegung im Sinne von wie ‚realistisch‘ eine Wirkung ist, nicht ausreichend. Auch ihrer Meinung nach sind die finanziellen Mittel das große Problem: Gutachterinnen und Gutachter sollen mit einer kurzen Feldphase alle DAC-Kriterien beantworten – Geld für elaborierte qualitative Designs, wie z.B. eine Kontributionsanalyse oder auch elaborierte quantitative Designs, ist nicht vorhanden. Im Ergebnis greifen Gutachterinnen und Gutachter auf Plausibilitätsüberlegungen, auf ‚logische Überlegungen‘ zurück. Zwar werden bei ‚echten‘ Wirkungsanalysen durchaus mehr Tage, mehr Zeit und mehr Geld bereitgestellt, aber im Normalfall – auch unter Berücksichtigung der zur Verfügung gestellten Arbeitstage – „kann ich mich gar nicht *nicht* damit [Plausibilitätsüberlegungen]

zufrieden geben...“. Auftraggebende sollten daher ihre Anforderungen reduzieren und realistischer gestalten bzgl. dem, was von Auftragnehmerinnen und Auftragnehmern geleistet werden kann. Andererseits erwartet sie von diesen aber Offenheit. In Berichten muss offen dargelegt werden, was tatsächlich ableitbar ist – die Formulierung „das Projekt hat dazu beigetragen, dass...“ ist für *Susanne von Jan* nicht ausreichend, wenn dies ausschließlich auf Plausibilitätsüberlegungen basiert. Dann erwartet sie eine entsprechende Formulierung, im Sinne von „aus den und den Gründen ist es plausibel, dass das Projekt da und da hinführt“. *Dorothee Mack* unterstreicht diesen Hinweis – für sie sind Plausibilitäten im Bereich Wirkungen durchaus akzeptabel, aber „es muss einfach gut nachvollziehbar und belegt sein“. *Susanne von Jan* schlägt als Ausweg aus dem Dilemma Joint Evaluations oder Clusterevaluationen vor: Mehr Ressourcen ermöglichen die methodisch saubere Beantwortung bestimmter Fragestellungen.

Für den Bereich der Kultur teilte *Vera Hennefeld* mit, dass Geld weniger das Problem ist – hier fehlt es eher an Zeit. Verzögerungen in der Planung und Ausschreibung sowie Vergabe von Evaluationen und dann schnell zu treffende Entscheidungen, z.B. über die Weiterführung eines Projektes, gehen oft zu Lasten der Evaluationszeit bzw. ‚sauberer‘ Methodenanwendung. Der Zwang, zeitnah ein Ergebnis vorzulegen, führt dann häufig zu dem Umstand, dass auf Plausibilitätsüberlegungen zurückgegriffen werden muss. Es gehört allerdings zum guten Standard, dies genau in dieser Form offen zu legen – und das hat aus ihrer Sicht nichts mit dem Budget zu tun. D.h. im Bericht ist aufzuführen, dass z.B. Outcomes nicht gemessen werden konnten, aber die Tatsache, dass anhand erbrachter Outputs zumindest die Voraussetzungen gegeben sind, dass Outcomes überhaupt erwartet werden können.

Im Rahmen der Diskussion um Methodenqualität, ‚sauberes‘ Arbeiten und Transparenz kam die Frage auf, inwieweit Evaluationsberichte ein **Methodenkapi-
tel** als Anhang enthalten sollten und inwieweit im Text **Quellenverweise** angeführt werden sollten – sowohl im Sinne von Verweisen auf Literatur und Dokumente aber auch im Sinne von Verweisen auf Datenquellen, wie Interviews, standardisierte Befragungen. *Dorothee Mack* verweist darauf, dass dies wiederum Arbeitstage benötigt. Gutachterinnen und Gutachter würden ihr sagen, dass sie Tabellen und Interviewtranskripte vorliegen hätten – bei Rückfragen könnte man darauf zurückgreifen. Aber in den Berichten erwarte sie das nicht – auch wolle das in der Politik ja keiner lesen. *Sonja Sheikh* bestätigt dies aus Auftragnehmerinnensicht: Literatur und Datenauswertungen sind in großem Umfang vorhanden. Natürlich enthalten Evaluationsberichte ein Literaturverzeichnis, aber es werden nicht an jeder Stelle einzelne Verweise eingefügt, das sei zu aufwändig. Stichprobenumfänge etc. würden natürlich im Bericht aufgeführt aber nicht detaillierte Zahlen zu Standardabweichungen oder t-Test-Ergebnissen – „da wird uns schon vertraut“. *Alexandra Caspari* argumentierte, dass durch diese Verweise die Nachvollziehbarkeit erhöht wird: Wenn hinter einer Aussage in Klammern aufgeführt wird, aus welchen Quellen diese Aussage stammt, z.B. Dokumente, Literatur, einzelnen Intensivinterviews sowie standardisierte Befragung, dann ist erkennbar, dass hier Aussagen trianguliert wurden und aus mehreren Quellen stammen – das genau erhöht doch die geforderte Nachvollziehbarkeit und die Plausibilität. *Susanne von Jan* betonte, dass sie als Auftraggebe-

rin sehr wohl Quellenverweise auch zu Literatur erwartet. Zudem begrüßt sie Angaben, die einzelne Aussagen konkreten Gruppen zuordnen, z.B. ZG-Vertretern oder Projektpartnern, so lange die Anonymität der betreffenden Personen gewährleistet bleibt.

Tülin Engin-Stock (uzbonn GmbH und Sprecherin des AK Kultur und Kulturpolitik) warf ein, dass die Diskussion den Eindruck erwecke, seitens der Auftraggebenden gebe es den Konsens, dass in Evaluationsberichten alles plausibel dargestellt werden müsse. Dies deckt sich allerdings nicht mit ihren Erfahrungen als Evaluatorin. Insbesondere wenn sie wissenschaftlich arbeitet und alles nach „bestem Wissen und Gewissen“ niederschreibt, sind die Rückmeldungen eher negativ – von „das kann kein Mensch lesen“ über Kritik am Umfang bis hin zu „das kann ich für meine Öffentlichkeitsarbeit nicht nutzen“. Sie weist darauf hin, dass meist eigentlich zwei Aufbereitungen der Ergebnisse nötig sind – eine interne Aufbereitung, die dem hohen Anspruch nach Legitimität und Transparenz gerecht wird, und eine weitere Aufbereitung der Ergebnisse für den eigentlichen Zweck der Evaluation, die Legitimation, die für die Öffentlichkeit und Verbreitung gedacht ist. Diese doppelte Anforderung macht auch doppelt Arbeit, ist aber in den Evaluationsanforderungen meist nur implizit enthalten. *Susanne von Jan* bestätigt dies teilweise: Es kommt sicherlich auch auf die Person auf Seiten der Auftraggebenden an. Sie selbst wolle diese Informationen gerne dargestellt haben, beispielsweise auch im Anhang. Aber natürlich gibt es verschiedenen Zielgruppen des Berichtes – die Geschäftsführung einer größeren Institution hat sicherlich weniger Interesse an diesen detaillierten Informationen, hier wird möglicherweise eher die Kurzzusammenfassung gelesen.

Jörg Faust (Leiter DEval) brachte einen neuen Aspekt in die Runde: Woher erhalten Organisationen **Informationen über existierende Evaluationsmethoden**, d.h. wie funktioniert „Diffusion von Methoden“, was sind die Treiber für **Imitationen** – in Abgrenzung zu ‚Innovationen‘ im Sinne von originärer Methodenforschung. Kommen Imitationen aus anderen Politikfeldern, aus dem internationalen Umfeld und/oder über direkte Peers, also Konkurrenten? Das DEval – so *Jörg Faust* – hat hier eine Sonderrolle inne, da es mit umfangreichen Personal- und Finanzmitteln unterschiedlichste Methodenexpertise zusammentragen kann. Hierdurch kann bei Evaluationen baukastenmäßig auf die unterschiedlichen Expertisen zurückgegriffen werden. In der EZ herrscht darüber hinaus eine große Diversifizierung, die Imitationen kommen somit aus unterschiedlichen Politikfeldern (z.B. Gesundheitsökonomie, Geografie), wobei man sich an internationalen Standards orientiert.

Sonja Sheikh berichtet – auch mit Blick auf ‚sauberes‘ Arbeiten – dass insbesondere Peers eine große Rolle spielen: Im Bereich FTI-Politik besteht zwischen den Evaluatorinnen und Evaluatoren ein großes Wettbewerbsverhältnis. Gerade in einem kleinen geografischen Bereich wie Österreich ist die Anzahl der Auftraggebenden und auch der Auftragnehmenden „überschaubar“, die Wettbewerber kennen sich untereinander. Standards und Qualität müssen einfach eingehalten werden, weil die Community die Qualität kontrolliert. D.h. man *muss* gut arbeiten, denn man hat einen Ruf zu verlieren. Das erzeugt einen viel größeren Druck als von Seiten der Auftraggebenden. Bei Ausschreibungen verzichten Auftraggebende teils gar auf eine detaillierte Methodendarstellung in den Angeboten, da sie wissen, dass methodisch

elaboriert gearbeitet werden kann und wird. Die indirekte Qualitätskontrolle über den Ruf von Evaluierungsanbietern ist im FTI-Bereich möglich, da Evaluationsberichte so gut wie immer publiziert werden. *Jörg Faust* wies darauf hin, dass in der EZ die Transparenz von Evaluationen im Sinne von Veröffentlichungspraxis eher heterogen ist. Hier wird nun allerdings verstärkt über Transparenzinitiativen und Transparenzindizes, die abbilden, inwieweit Organisationen ihre Evaluationsberichte vollumfänglich veröffentlichen, Wettbewerbsdruck zur Standardverbesserung ausgeübt. In der EZ ist der Markt international – somit ist hier die internationale Community die Referenzgröße.

Dorothee Mack berichtet, dass Imitation bei Misereor ebenfalls aus dem internationalen Umfeld kommt – und zwar über die Gutachterinnen und Gutachter, die oft in internationalen Teams arbeiten und somit Methoden, die in anderen Ländern angewendet werden, kennen lernen. Machen sie mit einer neuen Methode gute Erfahrung, bringen sie diese bei anderen Evaluationen zur Anwendung. Ein anderer Einflussbereich sind – so *Dorothee Mack* – Peers und nicht zuletzt die DeGEval und der AK Entwicklungspolitik.

Was ‚Imitation‘ betrifft, kocht man – so *Sonja Sheikh* – im Bereich FTI noch ‚im eigenen Süppchen‘. Da alle Evaluationsberichte auf der Plattform fteval² eingestellt sind, kann geschaut werden, was die Konkurrenz macht – aber diese kommen ebenfalls aus dem Bereich FTI. Politikfeldübergreifende Ideen finden sich eher in solchen Diskussionen im Rahmen der DeGEval. Hier kommen Ideen, die meist aufgrund fehlender Zeit trotz gutem Vorsatz nicht nochmals vertiefend betrachtet werden. Aber hier ist ‚mehr Potenzial als wir nutzen‘. *Susanne von Jan* bestätigt dies: Konferenzen wie die DeGEval-Tagung geben Hinweise zu bisher unbekanntem Methoden, die man näher anschauen will, was dann leider nicht immer passiert. Als weitere Quelle für Imitation nennt sie auch wissenschaftliche Artikel.

Im Politikfeld der Kulturevaluation dagegen – so *Vera Hennefeld* – ist Transparenz und somit die Möglichkeit der Kontrolle oder auch des Lernens im Sinne von Imitation annähernd nicht gegeben. So gibt es etwa keine Metaevaluationen und auch Evaluationsberichte werden so gut wie nie publiziert und sind damit nicht zugänglich. Somit ist weder Kontrolle noch Imitation über Peers möglich. Zwar wünscht sich z.B. das Goethe-Institut in seinen Evaluationsausschreibungen explizit ‚innovative Ansätze‘, aber es ist schwierig diese zu finden. Mangels Veröffentlichungen bleiben hier nur der Blick in Methodenbücher oder benachbarte Felder, also auch in Richtung Ökonomie oder auch Betriebswirtschaftslehre und Soziologie. Sinn machen hier theoretische Strömungen, aus denen methodisch etwas abgeleitet werden kann.

Alexandra Dolff (Deutsche Welle) berichtet von dem ‚Leidensdruck‘, dass im Bereich Kulturpolitik häufig Dinge erforscht werden wollen, für die es noch keine Methoden gibt – und somit nur ‚Innovation‘ möglich ist. Hier werden dann gemeinsam mit Peers Methodenprototypen entwickelt, z.B. eine Forschungs-App, um Videos im internationalen Standard vergleichbar bewerten zu können. Die Veröffentlichungspraxis in der Kulturpolitik empfindet sie auch als ‚Katastrophe‘. Allerdings

2 www.fteval.at

verweist sie darauf, dass gemäß Deutsche-Welle-Gesetz veröffentlicht wird: „Wir veröffentlichen alle vier Jahre – dann wenn wir müssen, dürfen wir auch“. Das Wissen ist aufgrund der vielen Studien, die im Forschungsverbund mit anderen Auslandssendern wie der BBC durchgeführt werden, eigentlich sehr umfangreich. Der Forschungsverbund hat aber strenge Veröffentlichungsregeln, daher ist das, was veröffentlicht werden darf, letztendlich so marginal, dass eine Veröffentlichung dann keinen Sinn mehr machen würde.

Ausblick

In der Kürze der Zeit konnten viele Fragen angesprochen, jedoch nicht abschließend diskutiert werden. Kennzeichnend für diese erste Analysesession war jedoch die Feststellung, dass in allen Politikfeldern der Balanceakt zwischen Methodenqualität und Methodentransparenz einerseits und zur Verfügung gestellter Zeit und finanzielle Mittel andererseits gleichermaßen zu bewältigen ist. Auch hat sich gezeigt, dass Evaluationsmethoden vornehmlich aus dem eigenen Politikfeld entnommen werden. Informationen über in anderen Politikfeldern angewandte Methoden werden eher selten eingeholt. Der Arbeitskreis Methoden bietet hier Unterstützung an: Im Schlusswort wies *Alexandra Caspari* darauf hin, dass sich der AK Methoden über Hinweise auf noch eher unbekannt in einem Politikfeld erfolgreich implementierte Methoden, die für andere Politikfelder innovativ sein könnten, sehr freut. Der AK Methoden kann dann evtl. im Rahmen einer Frühjahrstagung oder Session auf der DeGEval-Jahrestagung Arbeitskreise zusammenbringen oder spannende Methoden vertiefen. In einem Jahr wird der AK Methoden – so ist geplant – mit gleichem Format drei weitere Arbeitskreise näher betrachten.

Genderkompetenz und Genderexpertise – Anforderungen an Kompetenzprofile von Evaluator(inn)en

Maria Gutknecht-Gmeiner,¹ Kirstin Eckstein,² Angela Wroblewski³

Hintergrund

Die Berücksichtigung einer Geschlechterperspektive in Evaluationen ist sowohl Zeichen qualitätvoller Evaluation wie auch eine rechtliche Verpflichtung (vgl. Gutknecht-Gmeiner/Wotha/Wroblewski 2013). Der Arbeitskreis Gender Mainstreaming (AK GM) hat sich im Rahmen des Frühjahrsworkshops 2015 erstmals mit der Operationalisierung von Genderkompetenz und Genderexpertise auseinandergesetzt. Damit soll allen an Evaluationen beteiligten Personen fundiertes Hintergrundwissen zur Verfügung gestellt werden, um notwendige Kompetenzen erkennen und beurteilen zu können. Im Rahmen einer Arbeitssession des AK GM auf der DeGEval-Jahrestagung 2015 wurden konkretisierte Profile vorgestellt und diskutiert. Die aktuelle Fassung berücksichtigt die Ergebnisse der Arbeitssession auf der Jahrestagung 2015 sowie die Beschlüsse aus der Endredaktion im Frühjahrsworkshop 2016.

Gender Mainstreaming von Evaluationen bedeutet, die Genderdimension von Evaluationsgegenständen zu erkennen und in integrierter Weise in der Evaluation zu behandeln. Es geht darum, die geschlechtsspezifischen Unterschiede in Zugang zum und Teilhabe am Evaluationsgegenstand sowie die geschlechterdifferenzierte Wirkung des Evaluationsgegenstandes, die durch vorherrschende gesellschaftliche Machtverhältnisse, Strukturen, Werte, Normen und Rollenbilder bedingt sind, zu verstehen und zu berücksichtigen. Dazu sind Daten geschlechtsdifferenziert zu erheben und auszuwerten, Kontext und Ergebnisse systematisch auf mögliche geschlechtsspezifische Unterschiede hin zu untersuchen und die Bewertung der Programmwirkungen in gleichstellungsorientierter Perspektive vorzunehmen.

Die Berücksichtigung der Geschlechterperspektive schließt an ein Verständnis von Gender als soziales Konstrukt an, das als differenziertes Konzept die Vielfalt von Männern und Frauen berücksichtigt. Gender bezieht sich dabei nicht auf vermeintlich homogene Genusgruppen, sondern wird im Sinne der Intersektionalität als

1 IMPULSE – Evaluation und Organisationsberatung, Wien

2 Universität Graz

3 Institut für höhere Studien, Wien

mit weiteren Strukturkategorien (wie z.B. Alter, sozioökonomischer Status oder ethnischer Hintergrund) verwoben betrachtet.

Genderkompetenz als Voraussetzung für gendergerechte Evaluationen

Der Genderdimension in Evaluation gerecht zu werden erfordert nicht nur entsprechende Rahmenbedingungen und Ressourcen, sondern insbesondere auch ausreichend Genderkompetenz bei den an der Evaluation Beteiligten. Dies betrifft insbesondere Evaluierende, aber auch die Auftraggebenden von Evaluationen, die die zentralen Zielsetzungen und Ausrichtungen des Programms bzw. der Evaluation definieren.

Genderkompetenz wird verstanden als die Fähigkeit von Personen oder Teams⁴, die Genderdimension des Evaluationsgegenstandes zu erkennen und in der Konzeption und Durchführung der Evaluation zu berücksichtigen. Diese wird als grundlegende Kompetenz (Level 1) verstanden, über die alle Evaluierenden verfügen müssen, es handelt sich dabei um eine Mindestanforderung. Davon unterschieden wird eine vertiefte Genderexpertise (Level 2), für die es spezialisiertes Wissen, einschlägige Erfahrung sowie entsprechende Kompetenzen braucht, die nur durch langjährige intensive Beschäftigung mit dem Thema Gender erlangt werden können. Je nach Evaluationsgegenstand und Rahmenbedingungen können unterschiedliche Bereiche von Genderexpertise gefragt sein. Die für ‚Genderkompetenz‘ genannten Kompetenzen gelten kumulativ auch für ‚Genderexpertise‘.

Operationalisierung von Genderkompetenz und Genderexpertise

Doch was bedeutet Genderkompetenz konkret? Wie kann sie vermittelt, wie festgestellt werden? Die vorliegenden Kompetenzprofile sind ein erster Schritt zur Operationalisierung von Genderkompetenz und Genderexpertise für die Selbstreflexion, die Aus- und Weiterbildung von Evaluator(inn)en sowie für die Zusammenstellung oder Auswahl von Evaluationsteams.

Genderkompetenz (Level 1) bedeutet, dass Evaluator(inn)en über Sensibilität für Genderfragen verfügen und die Fähigkeit und den Willen, die eigene Arbeitsweise aus einer Genderperspektive kritisch zu reflektieren. Ebenso wird grundlegendes allgemeines Wissen zu geschlechtsspezifischen Disparitäten und gleichstellungspolitischen Zielen vorausgesetzt, um die Genderrelevanz des Evaluationsgegenstandes feststellen zu können. Teil dieser Genderkompetenz ist es auch, zu erkennen, dass

4 Genderkompetenz ist sowohl Qualifikationsmerkmal einer Person als auch eines Teams. Aufgeschlossenheit für Genderfragen und Reflexionsfähigkeit aller Teammitglieder ist Voraussetzung dafür, dass Gender durchgängig auch auf Teamebene berücksichtigt wird. Eine einzelne Person im Team mit dem Thema Gender zu betrauen, wird diesem Anspruch nicht gerecht. Auch Genderexpertise sollte im Team vertreten sein, v.a. wenn der Evaluationsgegenstand eine hohe Genderrelevanz aufweist und daher vertiefte Kenntnisse und Erfahrung im Bereich Gender notwendig sind. Genderexpertise wird von Individuen in die Evaluationsteams eingebracht.

der Evaluator/die Evaluatorin selbst die Genderdimension bzw. Genderrelevanz des Evaluationsgegenstandes nicht ausreichend berücksichtigen kann und hier gezielt Genderexpertise beizuziehen ist.

Genderkompetenz bedeutet daher u.a.:

- Erkennen der Genderdimension im eigenen Evaluationsgegenstand: d.h. Fragen stellen können, um die Genderdimension bzw. Genderrelevanz des Evaluationsgegenstandes zu erkennen; Beispiele für die Genderdimension bzw. Genderrelevanz im eigenen Fachbereich kennen,
- Fähigkeit, gleichstellungspolitische Ziele bewerten zu können bzw. übergeordnete gleichstellungspolitische Zielsetzungen im Evaluationsgegenstand verorten zu können,
- konzeptionelle und methodische Berücksichtigung von weiteren Strukturkategorien und deren Verwobenheit mit dem Merkmal Geschlecht (Intersektionalität),
- Reflexionsfähigkeit aus einer Genderperspektive in Bezug auf eigene Geschlechterrollen, Haltungen und Wertsysteme,
- positive Haltung gegenüber Gleichstellungszielen,
- Verwendung einer gendergerechten Sprache.

Genderexpertise (Level 2) verlangt ein spezialisiertes Wissen zu Geschlechterfragen und einschlägige Erfahrungen in der Umsetzung von Gender Mainstreaming. Eine intensive Auseinandersetzung mit feministischen und Gendertheorien sowie einschlägigen Forschungsergebnissen ist gefordert, die Genderdimension eines Evaluationsgegenstandes kann gleichstellungstheoretisch verortet werden. Methoden und Erkenntnisse der Genderforschung sowie der feministischen Theorie werden in die Evaluationstätigkeit einbezogen.

Tabelle 1: Einsatzbereiche von Genderkompetenz und Genderexpertise

	Genderkompetenz	Genderexpertise
Einsatzbereiche	<ul style="list-style-type: none"> – Erkennen und Integrieren einer allgemeinen Genderperspektive in Evaluationen – wenn möglich durch Genderexpertise (im Team) unterstützt 	<ul style="list-style-type: none"> – Einsatz als Genderexperte/ Genderexpertin in Evaluationen – Durchführung von Evaluationen von Gender Mainstreaming

Profile für Genderkompetenz und Genderexpertise für Evaluator(inn)en

Wissen – Fähigkeiten – Haltungen

Die weitere Operationalisierung der Kompetenzen bedient sich gängiger Taxonomien für die Differenzierung von Kompetenzen nach Wissen (kognitive Aspekte, ‚knowledge‘), Fähigkeiten/Können (Handlungsaspekt, ‚skills‘) und Haltungen (affektive und motivationale Aspekte, Wollen, ‚attitudes‘). Haltungen und entsprechend personale und methodische Fähigkeiten sind für Genderkompetenz besonders wichtig, da diese zwischen Geschlechtertheorie/-forschung und Gleichstellungspraxis/

-auftrag vermitteln soll: Genderkompetenz ist keine rein theoretische Qualifikation, sondern schließt das Hinarbeiten auf Veränderung in Richtung Gleichstellung in der eigenen Praxis ein.

Diese Kompetenzbereiche werden ergänzt durch ‚Reflexionsfähigkeit‘; diese Kompetenz, die auch zur Selbstkompetenz gezählt werden kann, wird in Operationalisierungen von Genderkompetenz immer wieder aufgrund ihrer Bedeutung als eigenständiger Kompetenzbereich genannt. Fähigkeiten werden in methodische Kompetenzen und Sozial- und Selbstkompetenzen weiter differenziert.

Kompetenzfelder der Empfehlungen für die Aus- und Weiterbildung von Evaluator(inn)en

Im Folgenden werden Genderkompetenz und Genderexpertise anhand der ersten vier Kompetenzfelder der Empfehlungen für die Aus- und Weiterbildung von Evaluator(inn)en dargestellt (vgl. DeGEval 2004). Um Genderkompetenz konkret in Curricula für die Aus- und Weiterbildung von Evaluator(inn)en verankern zu können, wäre eine weiterführende detaillierte Aufschlüsselung erforderlich. Auch für das Kompetenzfeld „Praxis der Evaluation“ mit Bezug auf Gender müssen weitergehende erste Überlegungen angestellt werden.

Tabelle 2: Zusammenfassung der Profile für Genderkompetenz und Genderexpertise von Evaluator(inn)en nach Kompetenzbereichen⁵

Kompetenzbereiche	Beispiele für Genderkompetenz	Beispiele für Genderexpertise
Wissen	- Basiswissen über Geschlecht als soziokulturelle Strukturkategorie in Wechselwirkung mit anderen Strukturkategorien (Intersektionalität)	- Wissen über Geschlechtertheorien (wie feministische Theorien, kritische Männer-/Maskulinitätsforschung; LGBTQ-Theorien (Lesbian/Gay/Bisexual/Queer/Transgender))
	- Basiswissen über geschlechtsbezogene Macht- und Herrschaftsverhältnisse	- vertieftes Wissen über Gender Mainstreaming als politische Strategie
	- Kenntnis der geltenden gleichstellungspolitischen Ziele	- vertieftes Wissen über Forschungsergebnisse zu Genderaspekten und intersektionalen Zusammenhängen
	- zentrale Aspekte für die Genderdimension bzw. Genderrelevanz im eigenen Fachbereich kennen	- Modelle der Genderanalyse sowie relevante Methoden und Indikatoren kennen
Wollen/Haltung	- Wissen über die Bedeutung einer gendergerechten Sprache als wichtiges Element zur Veränderung von Denken und Handeln	
	- Bereitschaft zur Auseinandersetzung mit Geschlechterfragen	- Motivation, Geschlechtergerechtigkeit als Ziel auch gegen Widerstände zu verfolgen
	- positive Haltung gegenüber Gleichstellungszielen	- Motivation, anderen die Bedeutung einer Geschlechterperspektive näherzubringen
	- gendergerechte Sprache auch verwenden, wenn andere dies nicht für notwendig halten	

5 Die vollständigen Kompetenzprofile sind auf der Homepage des AK GM verfügbar: http://www.degeval.de/fileadmin/users/Arbeitskreise/AK_Gender/Profile_GenderkompetenzGenderexpertise_final.pdf

Kompetenzbereiche	Beispiele für Genderkompetenz	Beispiele für Genderexpertise
Reflexionsfähigkeit	<ul style="list-style-type: none"> - Reflexionsfähigkeit in Bezug auf Geschlechterrollen, Geschlechterbilder, Geschlechterstereotypen allgemein und insbesondere in Bezug auf den Evaluationsgegenstand - Reflexionsfähigkeit in Bezug auf eigene Alltagserfahrungen, Wahrnehmungen, Einschätzungen und Annahmen sowie die eigenen Sozialisationserfahrungen aus einer Genderperspektive 	<ul style="list-style-type: none"> - Reflexionsfähigkeit in Bezug auf Re-Produktionsmechanismen von Ungleichheit und deren Überschneidungen (Intersektionalität) - Unterstützung und Förderung der Reflexionsfähigkeit anderer Personen - Reflexionsprozesse mit Betroffenen und Beteiligten konzipieren und moderieren können
Können/Handeln: Sozial- und Selbstkompetenz	<ul style="list-style-type: none"> - Fähigkeit eines gendergerechten Sprachgebrauchs ist Voraussetzung - nicht stereotypisierende Dialogbereitschaft 	<ul style="list-style-type: none"> - Fähigkeit, Auseinandersetzungen zu Genderfragen anzuregen und zu begleiten - Fähigkeit, mit Angriffen und Widerständen umgehen zu können - Fähigkeit, mit genderspezifischen Macht- und Herrschaftsstrukturen und -verhältnissen umgehen zu können
Können/Handeln: Methodenkompetenz	<ul style="list-style-type: none"> - Erkennen der Genderdimension im eigenen Evaluationsgegenstand: d.h. Fragen stellen können, um die Genderdimension bzw. Genderrelevanz des Evaluationsgegenstandes zu erkennen - die benötigten Informationen (Daten, Forschungsergebnisse, etc.) zu zentralen Genderaspekten finden und einholen können - Geschlecht in Grafiken und Bildern angemessen und in nicht stereotypisierender Weise sichtbar machen können - Fähigkeit, zu erkennen, wann zusätzliche Genderexpertise eingeholt werden muss 	<ul style="list-style-type: none"> - geschlechtsspezifische Disparitäten in allen Dimensionen des Evaluationsgegenstandes (Kontext, Zielsysteme, Maßnahmenplanung und -durchführung, Outputs sowie Wirkungen) erkennen und durchgehend in allen Evaluationsphasen berücksichtigen können - geeignete Genderkriterien und -indikatoren für die jeweiligen Evaluationsgegenstände/Fragestellungen anwenden oder neu entwickeln können - gendersensibles Vorgehen in der Durchführung der Evaluation, insbesondere in Bezug auf die Erhebungsmethoden und die Kommunikation mit Zielgruppen

Literatur

- DeGEval – Gesellschaft für Evaluation e.V. (Hg.) (2004): Empfehlungen für die Aus- und Weiterbildung in der Evaluation. Anforderungsprofile an Evaluatorinnen und Evaluatoren. Verfügbar unter: http://www.degeval.de/fileadmin/Publikationen/Publikationen/Homepage/DeGEval_-_Empfehlungen_Aus-_und_Weiterbildung.pdf [21.12.2016].
- Gutknecht-Gmeiner, Maria/Wotha, Brigitte/Wroblewski, Angela (2013): Considering Gender. Qualitätsvolle Evaluierung durch Einbeziehung der Geschlechterperspektive in Evaluationen. In: Zeitschrift für Evaluation, 12 (2), S. 349-352.

The Dutch Evaluatorenennetwerk

Frank van den Broek,¹ Dick Hanemaayer²

Evaluation Practice in the Netherlands

The Netherlands have a relatively robust evaluation infrastructure. Elements of this include – at the level of the national government – a legal obligation to evaluate policies periodically; there is an obligatory ‘policy review’ by which each policy item in the budgets should be reviewed every few years; a review clause is included in many laws; and parliament regularly calls for evaluations in order to discuss specific policy issues. At province and municipality level, local/regional audit offices are established by law; they usually focus on the results of local policies. Furthermore, in the Netherlands there are a large number of research institutes and private agencies which have the expertise to deliver independent evaluations of a decent quality.

On the other hand, a homogeneous professional group of evaluators does not exist. There is no specific training for evaluators; and evaluation skills are only a small part of various study programs such as public administration and sociology. Nor are there any professional standards for evaluators.

Also, for some years there have been concerns about the quality and use of policy evaluations. Thus, the Dutch board of audit found in a number of studies that, despite the legal obligation to do so, not all policies are evaluated, in most evaluations the effectiveness of policies remains unclear and also the use of evaluations not always provides reason for optimism.

Vide and the Evaluatorenennetwerk

Within this context, the Evaluatorenennetwerk was founded five years ago. Given the context – a well-developed assessment infrastructure and a heterogeneous group of evaluators – a light organizational structure has been chosen: a network which is

1 Board Vide, Den Haag, The Netherlands

2 Beleidsevaluatie.info, Oegstgeest, The Netherlands

freely accessible and which connects about two hundred evaluators and other interested parties. They participate in meetings on evaluation and/or show their interest in other form, for example by joining the LinkedIn group 'Leren van evalueren'. Especially staff of the national authorities are members of the Evaluatorenennetwerk: employees of departments, planning offices, the board of audit, and research institutes. To a lesser extent evaluators are connected who work for municipalities, counties, universities and regional boards of audit.

The network is an organizational part of Vide. Vide is the Dutch association of professionals in the field of monitoring, inspection, enforcement and evaluation. Vide thus has the role of the Dutch Evaluation Society. Vide is an organization of and for people who either as individuals or professionals get in touch with colleagues via the association. Vide aims to be the driving force behind the professional development of its members and to contribute to the continuous improvement of monitoring, inspection, enforcement actions and evaluations. That way, its intention is to contribute to a well-functioning government and society.

In line with this philosophy of Vide, and given the already highly institutionalized evaluation infrastructure, the focus of the work of the Evaluatorenennetwerk lies on exchange of knowledge and getting into contact within the profession.

The costs for the functioning of the Evaluatorenennetwerk are minimal. There is a web page set up within the Vide website (<https://www.videnet.nl/evaluatorenennetwerk>) and the Evaluatorenennetwerk is represented in the editing of 'Beleidsonderzoek Online', a Dutch public administration online journal which publishes articles on policy research and evaluation (<http://www.beleidsonderzoekonline.nl/>).

Activities and Methods of the Evaluatorenennetwerk

Annually, four to six meetings are organized concerning the three main subjects of the Evaluatorenennetwerk: use and usefulness of evaluations, evaluation methods and evaluation infrastructure. Meetings are organized by a working group of six people. Most members of this group are employed by governmental organizations. One of them is also a board member of Vide. Usually, during these meetings one or more invited speakers give presentations on a specific topic from the three main subjects named above and enter into a discussion with the participants.

Examples of events include:

- A meeting on a successful evaluation program in the healthcare sector (Programma Doelmatigheid ('Program Effectiveness') conducted by ZonMw). This program is almost unique in the Netherlands because it was implemented about fifteen years ago and is still working. This is a good reason to learn and to discuss why this program still exists after all these years: What are its success factors and what are important favorable conditions?
- A meeting on the central government mandatory instrument 'beleidsdoorlichtingen' (policy reviews). Given the criticisms of this evaluation instrument, central questions with regard to the functioning of the policy reviews were raised.

What is working well, what could be better? And how could it get better? Should we switch to another system or should we look for improvements within the current framework?

- A meeting on Qualitative Comparative Analysis (QCA). Central questions in this meeting were what QCA exactly means, what QCA precisely offers evaluators and clients, and which challenges for evaluators and clients might arise by the application of QCA.

These are just three examples of meetings, but together they offer a good insight into the topics addressed by the Evaluatorenennetwerk. On average, the meetings are attended by 20 to 50 people; the composition of the respective groups depends to a great extent on the specific issues.

European Evaluation Society

In September 2016, the biennial conference of the European Evaluation Society took place in Maastricht (Netherlands). This was a good opportunity for the Evaluatorenennetwerk to present itself and Vide to the European evaluation community. Therefore, the network had prepared some sessions on evaluation issues in the Netherlands to get into conversation with other national evaluation societies. In our opinion, there are many reasons to perpetuate and tighten the ties that already exist through EES and the Network of European Evaluation Societies (NESE). As evaluators in different countries, we all have to face similar challenges. Thus, the EES conference in September 2016 provided a good platform to learn from each other.

Autorinnen und Autoren

Juliane Achatz
 Institut für Arbeitsmarkt- und Berufsforschung
 Regensburger Straße 100
 D-90478 Nürnberg
 E-Mail: juliane.achatz@iab.de

Thomas Altenburg
 Korkedamm 22d
 D-12524 Berlin
 E-Mail: t.altenburg@berlin.de

Dr. Antonia Baumeister
 Technische Universität Chemnitz
 Institut für Psychologie
 Wilhelm-Raabe-Straße 43
 D-09107 Chemnitz
 E-Mail: antonia.baumeister@psychologie.tu-chemnitz.de

Prof. Dr. Edgar Baumgartner
 Fachhochschule Nordwestschweiz
 Hochschule für Soziale Arbeit
 Institut Professionsforschung und -entwicklung
 Rigenbachstr. 16
 CH-4600 Olten
 E-Mail: edgar.baumgartner@fhnw.ch

Pirmin Bundi
 Universität Zürich
 Institut für Politikwissenschaft
 Affolternstrasse 56
 CH-8050 Zürich
 E-Mail: bundi@ipz.uzh.ch

Lena Franziska Burchartz
 Technische Universität Berlin
 Fachgebiet Umweltprüfung und Umweltplanung
 Sekr. EB 5
 Straße des 17. Juni 145
 D-10623 Berlin
 E-Mail: lena.f.burchartz@campus.tu-berlin.de

Alain Dössegger
 Eidgenössische Hochschule für Sport
 Magglingen EHSM
 Hauptstrasse 247
 CH-2532 Magglingen
 E-Mail: alain.doessegger@baspo.admin.ch

Dr.in Kirstin Eckstein
 Universität Graz
 Beethovenstraße 19
 Koordinationsstelle für Geschlechterstudien und
 Gleichstellung
 A-8010 Graz
 E-Mail: kirstin.eckstein@uni-graz.at

Klaus M. Frei
 Eidgenössische Hochschule für Sport
 Magglingen EHSM
 Hauptstrasse 247
 CH-2532 Magglingen
 E-Mail: klaus.frei@baspo.admin.ch

Felix Gaisbauer
 Deutsches Evaluierungsinstitut der
 Entwicklungszusammenarbeit (DEval)
 Fritz-Schäffer-Str. 26
 D-53113 Bonn
 E-Mail: felix.gaisbauer@deval.org

Priv.-Doz. DDr. Ulrike Greiner
 Universität Salzburg
 Unipark
 Erzabt-Klotz Str. 1
 A-5020 Salzburg
 E-Mail: ulrike.greiner@sbg.ac.at

Dr.in Maria Gutknecht-Gmeiner
 Impulse – Evaluation und Organisationsberatung
 Dr.-Josef-Resch-Platz 14/3
 A-1170 Wien
 E-Mail: m.gutknecht-gmeiner@impulse.at

Edith Halves
 Hochschule Angewandte Wissenschaften (HAW)
 Hamburg
 Zentrum für Praxisentwicklung (ZEPRÄ)
 Arbeitsbereich Evaluation und Praxisentwicklung
 Alexanderstraße 1
 D-20099 Hamburg
 E-Mail: edith.halves@haw-hamburg.de

Drs. Dick E. Hanemaayer
 beleidsevaluatie.info
 Willem de Zwijgerlaan 33
 NL-2341 EH Oegstgeest
 E-Mail: dick@beleidsevaluatie.info

Christoph Hartmann
Deutsches Evaluierungsinstitut der
Entwicklungszusammenarbeit (DEval)
Fritz-Schäffer-Str. 26
D-53113 Bonn
E-Mail: christoph.hartmann@deval.org

Dr. Sigrid Haunberger
Fachhochschule Nordwestschweiz
Hochschule für Soziale Arbeit
Institut Professionsforschung und -entwicklung
Riggenbachstr. 16
CH-4600 Olten
E-Mail: sigrid.haunberger@fhnw.ch

Franziska Heinze
Deutsches Jugendinstitut e.V.
Außenstelle Halle
Franckeplatz 1, Haus 12/13
D-06110 Halle (Saale)
E-Mail: heinze@dji.de

Prof. Dr. Jan Hense
Justus-Liebig-Universität Gießen
Professur für Hochschuldidaktik und Evaluation
Fachbereich 06 Psychologie und
Sportwissenschaft
Otto-Behagel-Str. 10 F
D-35390 Gießen
E-Mail: jan.hense@psychol.uni-giessen.de

Dr. P. H. Heidemarie Kelleter
Diözesan-Caritasverband für das Erzbistum
Köln e.V.
D-50676 Köln
E-Mail: heidemarie.kelleter@caritasnet.de

Dr. Christiane Kerlen
Evaluation und Beratung
6 Hamilton's Folly Mews
Edinburgh
UK-EH8 9AW
E-Mail: info@kerlen.de

Michael Köhler
Steinbeis-Transferzentrum für Interventions- und
Evaluationsforschung
Karl-Peters-Straße 11
D-33605 Bielefeld
E-Mail: mkoehler262@gmail.com

Frank König
Deutsches Jugendinstitut e.V.
Außenstelle Halle
Franckeplatz 1, Haus 12/13
D-06110 Halle (Saale)
E-Mail: fkoenig@dji.de

Dr. Stefanie Krapp
Deutsches Evaluierungsinstitut der
Entwicklungszusammenarbeit (DEval)
Fritz-Schäffer-Str. 26
D-53113 Bonn
E-Mail: stefanie.krapp@deval.org

Dr. Wolfgang Meyer
Universität des Saarlandes
Centrum für Evaluation (CEval)
Campus, Geb. C5.3
D-66123 Saarbrücken
E-Mail: w.meyer@mx.uni-saarland.de

Prof. Dr. Helfried Moosbrugger
Goethe-Universität Frankfurt am Main
Psychologische Methodenlehre, Evaluation und
Forschungsmethodik
PEG-Gebäude, Hauspostfach 75
Theodor-W.-Adorno-Platz 6
D-60629 Frankfurt am Main
E-Mail: moosbrugger@psych.uni-frankfurt.de

Dr. Andreas Paschon
Universität Salzburg
Unipark
Erzabt-Klotz Str. 1
A-5020 Salzburg
E-Mail: andreas.paschon@sbg.ac.at

Anke Rehhausen
Technische Universität Berlin
Fachgebiet Umweltprüfung und Umweltplanung
Sekt. EB 5
Straße des 17. Juni 145
D-10623 Berlin
E-Mail: anke.rehhausen@tu-berlin.de

Stefanie Reiter
Deutsches Jugendinstitut e.V.
Außenstelle Halle
Franckeplatz 1, Haus 12/13
D-06110 Halle (Saale)
E-Mail: sreiter@dji.de

a.o. Univ. Prof. Dr. Franz Riffert
Universität Salzburg
Fachbereich Erziehungswissenschaft
Unipark
Erzabt-Klotz Str. 1
A-5020 Salzburg
E-Mail: franz.riffert@sbg.ac.at

Prof. Dr. Brigitte Schels
Institut für Arbeitsmarkt- und Berufsforschung
Regensburger Straße 100
D-90478 Nürnberg
E-Mail: brigitte.schels@iab.de

Dr. Dörte Schott
Freie Evaluatorin
Heckelsgasse 4
D-53227 Bonn
E-Mail: info@doerte-schott.de

Dr. Sonja Sheikh
KMU Forschung Austria
Gußhausstr.8
A-1040 Wien
E-Mail: s.sheikh@kmuforschung.ac.at

Benoit Simon
Planète Publique
30, rue de Fleurus
F-75006 Paris 6ème
E-Mail: simon@planetepublique.fr

Drs. Frank P.G. van den Broek
Algemene Rekenkamer (Court of Audit)
Postbus 20015
NL-2500 EA Den Haag
E-Mail: f.vandenbroek@rekenkamer.nl

Kirsten Vorwerk
Deutsches Evaluierungsinstitut der
Entwicklungszusammenarbeit (DEval)
Fritz-Schäffer-Str. 26
D-53113 Bonn
E-Mail: kirsten.vorwerk@deval.org

Prof. (FH) Dr. Ingrid Wahl
Ferdinand Porsche FernFH-Studiengänge
Studiengang Betriebswirtschaft &
Wirtschaftspsychologie Bachelor
Zulingerstraße 4
A-2700 Wiener Neustadt
E-Mail: ingrid.wahl@fernfh.ac.at

Prof. (FH) Dr. Christa Walenta
Ferdinand Porsche FernFH-Studiengänge
Studiengang Betriebswirtschaft &
Wirtschaftspsychologie Bachelor
Zulingerstraße 4
A-2700 Wiener Neustadt
E-Mail: christa.walenta@fernfh.ac.at

David Weibel
w hoch 2 GmbH
Münstergasse 72
CH-3011 Bern
E-Mail: d.weibel@w-hoch2.ch

Bartholomäus Wissmath
w hoch 2 GmbH
Münstergasse 72
CH-3011 Bern
E-Mail: b.wissmath@w-hoch2.ch

PD Dr. Joachim Wolff
Institut für Arbeitsmarkt- und Berufsforschung
Regensburger Straße 100
D-90478 Nürnberg
E-Mail: joachim.wolff@iab.de

Prof. Dr. (em.) Hellmut Wollmann
Humboldt-Universität zu Berlin
Unter den Linden 6
D-10099 Berlin
E-Mail: hellmut.wollmann@rz.hu-berlin.de

Dr.in Angela Wroblewski
Institut für höhere Studien
Josefstädterstr. 39
A-1060 Wien
E-Mail: wroblews@ihs.ac.at

Hinweise für Autorinnen und Autoren

Allgemeine Hinweise: Die Zeitschrift für Evaluation (ZfEv) ist in zwei Kategorien unterteilt:

– **Originalbeiträge:** Wissenschaftliche, theoretische und konzeptionelle Beiträge, Ergebnisse empirischer Studien sowie Praxisbeiträge zu aktuellen Themen werden entweder unter der Rubrik ‚Theorien, Methoden und Praxis der Evaluation‘ oder zu einem ‚Themenschwerpunkt‘ zusammengefasst veröffentlicht. Alle als Originalbeitrag eingereichten Manuskripte werden einem doppelblinden Begutachtungsverfahren unterzogen. Die Bewertung der Manuskripte erfolgt (1) hinsichtlich der Einschlägigkeit bezogen auf die Schwerpunkte der Zeitschrift, (2) der Einhaltung wissenschaftlicher Standards und (3) des Erkenntnisgewinns. Der Anspruch der Originalität verlangt, dass eingereichte Manuskripte bisher nicht bereits an anderer Stelle veröffentlicht wurden bzw. bis zur endgültigen Annahme oder Ablehnung nicht anderweitig zur Publikation eingereicht werden.

– **Information & Service:** Rezensionen, Tagungsberichte, Vorstellung von Evaluationsinstituten sowie sonstige Informations- und Servicenachrichten.

Manuskripteinreichung: Manuskripte sind per E-Mail an die Redaktion als Datei in einem Standard-Textverarbeitungsprogramm (Microsoft Word, Word Perfect) oder rtf-Format einzureichen. Grafiken oder Abbildungen sind zusätzlich als PowerPoint- oder Excel-Datei beizufügen (andere Formate sind vorab mit der Redaktion zu klären).

Manuskriptvorgaben: Bei der Abfassung von Manuskripten für die ZfEv ist zu beachten:

– **Äußere Gestaltung:** Seitenrand oben, unten, links, rechts 2,5cm; Schriftart ‚Times New Roman‘; Schriftgröße 12pt; 1-zeilig; keine Kopf-/Fußzeilen; keine Trennungen oder Steuerzeichen; Hervorhebungen ausschließlich *kursiv* (kein Fettdruck, Unterstreichungen etc.); Eigenbegriffe können in einfache Hochkommata gesetzt werden, doppelte Anführungszeichen sind ausschließlich für direkte Zitate zu nutzen.

– **Länge:** Originalbeiträge sollten zwischen 15 und 30 Druckseiten umfassen (ca. 5.000 bis 10.000 Wörter). Beiträge für die Kategorie ‚Information & Service‘ sollten i.d.R. nicht mehr als 10 Druckseiten umfassen (ca. 3.000 Wörter), Praxisberichte max. 15 Seiten.

– **Deckblatt:** Enthält den deutschen Titel (kurz und bezeichnend), ggf. Untertitel sowie die englische Übersetzung des Titels, ggf. Untertitels; den Namen der Autorin/des Autors (ggf. akadem. Titel), die institutionelle Zugehörigkeit (Name und Anschrift der Institution) sowie die E-Mail-Adresse.

– **Zweite Seite:** Wie Deckblatt, jedoch ohne Namen, Institution und Anschrift, plus deutsche Zusammenfassung (max. 1.000 Zeichen inkl. Leerzeichen, ca. 120 Wörter) und 4 deutsche Schlagwörter sowie eine englischsprachige Zusammenfassung (max. 1.000 Zeichen inkl. Leerzeichen, ca. 120 Wörter) und 4 englische Keywords.

– **Text ab dritter Seite:** Kapitel sind numerisch zu gliedern; Fußnoten sind fortlaufend zu nummerieren, durch hochgestellte Ziffern ohne Klammern zu kennzeichnen und am Seitenende aufzuführen.

– **Lesbarkeit:** Manuskripte sind gendersensibel zu verfassen und vor der Einreichung sowohl auf Grammatik, Orthographie etc. als auch auf angemessene sprachliche Qualität zu prüfen; Grundlage ist die neue deutsche Rechtschreibung. Bitte geschlechtsneutrale Formulierungen oder sowohl die weibliche als auch die männliche Form verwenden. Zur verkürzten Schreibweise kann auf die Schrägstrich- oder Klammervariante zurückgegriffen werden (keine Binnenmajuskel). Beispiele: Mitarbeiter/-innen oder Mitarbeiter(innen), nicht: MitarbeiterInnen.

– **Tabellen/Abbildungen:** Tabellen/Abbildungen sind unter Verwendung der Überschrift ‚Tabelle‘ bzw. ‚Abbildung‘ chronologisch zu nummerieren und mit einem Titel zu versehen.

– **Literaturverweise:** Direkte und indirekte Zitate sind in den Text durch Klammerverweise auf das Literaturverzeichnis einzuarbeiten (‚amerikanische Zitierweise‘). Indirekte Zitate werden dabei mit ‚vgl.‘ eingeleitet. Zwischen verschiedenen Literaturverweisen steht ein Semikolon, bei mehrfacher Zitierung derselben Autorin/desselden Autors ein Komma. Bei mehr als drei Autorinnen/Autoren ist nach der Erstautorin/dem Erstautor ‚et al.‘ anzuführen. Beispiele: (vgl. Scriven 1967, 1991); (vgl. Wottawa/Thierau 1990; Mertens et al. 1994); (vgl. Brandtstädter 1990: 217ff.); (Levine 1984: 99).

– **Literaturverzeichnis:** Sämtliche Quellenangaben im Text sind vollständig und in alphabetischer Reihenfolge im Literaturverzeichnis am Ende des Manuskripts aufzulisten. Die bibliographischen Angaben sollen der folgenden Reihenfolge entsprechen: Familienname, Vorname (Erscheinungsjahr): Titel. Untertitel. Erscheinungsort: Verlag (Auflage). Beispiele für Literaturangaben:

Brandtstädter, Jochen (1990): Evaluationsforschung. Probleme der wissenschaftlichen Bewertung von Interventions- und Reformprojekten. In: Zeitschrift für Pädagogische Psychologie, 4 (2), S. 215-228.

Levine, Robert A. (1984): Programmevaluierung und Politikanalyse in Europa, USA und Kanada – Ein Überblick. In: Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Handbuch zur Evaluationsforschung, Band 1. Opladen: Westdeutscher Verlag, S. 94-133.

Rossi, Peter H./Freeman, Howard E./Lipsey, Mark W. (1999): Evaluation. A Systematic Approach. Thousand Oaks, CA u.a.: Sage (6. Aufl.).

Wottawa, Heinrich/Thierau, Heike (1990): Lehrbuch Evaluation. Bern u.a.: Verlag Hans Huber.