

Sonderdruck aus

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

Evaluation in Deutschland und Österreich

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

Print-ISBN 978-3-8309-3149-2

E-Book-ISBN 978-3-8309-8149-7

© Waxmann Verlag GmbH, 2014
Postfach 8603, 48046 Münster

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten. Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Bestellung per Fax: 0251 26504-26 oder
telefonisch: 0251 26504-0;
per Internet unter www.waxmann.com/buch3149 oder
per E-Mail: order@waxmann.com

Arbeitsfelder und Herausforderungen der Evaluation

Evaluation! Es ist noch nicht lange her, dass dies ein Fremdwort war, dessen fehlerfreie Aussprache manch einem Zeitgenossen bereits Probleme bereitete. Die Karriere der Evaluation ist steil, womöglich auf Kosten ihrer Konturierung. Allzu leicht werden heute auch triviale Formen der Rückmeldung zu Evaluationen. Das ist die eine Seite. Auf der anderen verleihen sich kontrollierende Messverfahren den Titel Evaluation.

Wie lässt sich Evaluation konzise beschreiben? Die „DeGEval – Gesellschaft für Evaluation“ formuliert es so: „Evaluation ist die systematische Untersuchung des Nutzens oder Wertes eines Gegenstandes. Solche Evaluationsgegenstände können z.B. Programme, Projekte, Produkte, Maßnahmen, Leistungen, Organisationen, Politik, Technologien oder Forschung sein. Die erzielten Ergebnisse, Schlussfolgerungen oder Empfehlungen müssen nachvollziehbar auf empirisch gewonnenen qualitativen und/oder quantitativen Daten beruhen“ (DeGEval 2002).

Bereits dieser knappe Versuch einer Bestimmung wirft ein Licht auf die Komplexität von Evaluation im Hinblick auf Gegenstände, Methoden und Logik von Schlussfolgerungen. Diese wird erhöht, wenn man bedenkt, dass Evaluationen unterschiedliche Leistungen erbringen bzw. Zielsetzungen verfolgen können. Bereits die schlichte Unterscheidung zwischen „formativer Evaluation“, die – grob gesprochen – den jeweiligen Gegenstand „begleitet“ und auf Verbesserung zielt, und der „summativen Evaluation“, die Ergebnisse bilanziert, kann das verdeutlichen.

Auch die von der DeGEval vertretenen Standards der Evaluation zeigen, wie schwierig die Aufgabe ist, eine „gute“ Evaluation durchzuführen. Eine qualitativ hochwertige, also in diesem Sinne „professionelle“ Evaluation ist anspruchsvoll. In Anlehnung an die Program Evaluation Standards des Joint Committee for Standards for Educational Evaluation (JCSEE 1994) und die Standards der schweizerischen Schwes-tergesellschaft SEVAL wurden insgesamt 25 Einzelstandards definiert, die in die vier Standardgruppen Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit gegliedert sind. Die *Nützlichkeitsstandards* thematisieren die Notwendigkeit, dass Evaluation sich an abgestimmten und geklärten Zwecken orientiert, damit sie am Informationsbedarf der vorgesehenen Nutzer und Nutzerinnen ausgerichtet ist. Evaluationen sollen die Basis fürs Lernen, also die Entwicklung und Verbesserung des jeweiligen Evaluationsgegenstands erzeugen. Die *Durchführbarkeitsstandards* sollen sicherstellen, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst geplant und ausgeführt wird. Die *Fairnessstandards* sollen gewährleisten, dass in einer Evaluation respektvoll und fair mit den beteiligten und betroffenen Personen und Gruppen umgegangen wird. Die *Genauigkeitsstandards* sollen dafür Sorge tragen, dass eine Evaluation gültige und überprüfbare Informationen und Ergebnisse zu dem jeweiligen Evaluationsgegenstand und den Evaluationsfragestellungen hervorbringt und vermittelt. Im Wesentlichen geht es hier um die Wissenschaftlichkeit der Evaluation, mithin ihre methodische Seriosität, um Validität, Reliabilität und Intersubjektivität.

„Die Standards sollen die Qualität von Evaluationen sichern und entwickeln helfen. Sie sollen als Dialoginstrument und fachlicher Bezugspunkt für einen Austausch über die Qualität von professionellen Evaluationen dienen“ (DeGEval 2002, S. 2). In der Kommunikation mit Auftraggebern, Adressaten und Adressatinnen sowie einem weiten Kreis von Beteiligten und Betroffenen, also auch einer interessierten Öffentlichkeit, können Standards als Referenz für „Evaluation als professionelle Praxis“ dienen (vgl. ebd.).

Es dürfte einleuchtend sein, dass hiermit hohe Standards der „Professionalität“, mithin der Güte von Evaluation gesetzt sind. Und ebenso dürfte deutlich sein, wie schwierig es sein kann, in der Praxis die Standards einzuhalten. Dies nicht nur deshalb, weil es vorkommt, dass Interessensvertreter (insbesondere Auftraggeber/innen) Druck auf die Evaluierenden ausüben, weil die zur Verfügung stehenden Ressourcen (Zeit, Geld) zu knapp bemessen oder Erwartungen an Evaluationsergebnisse zu hoch gesteckt sind. Dies gilt auch deshalb, weil in der Praxis nicht ohne weiteres alle Standards gleichgewichtig und gleichzeitig erfüllt werden können.

Anspruch und Komplexität von Evaluation sind damit aber noch nicht annähernd vollständig beschrieben. Ein wesentlicher Grund dafür, dass Evaluation kein generisches Konzept ist, liegt in der Tatsache begründet, dass sie in sehr unterschiedlichen Bereichen durchgeführt wird. Sie muss also immer feldspezifischen Anforderungen genügen. Und hiermit ist eine zentrale Absicht benannt, die mit der vorliegenden Publikation verfolgt wird. Bei aller Notwendigkeit von Standards – also allgemeinen Anforderungen an eine gute Evaluation – gilt es auch zu zeigen, wie breit das Feld ist, in dem evaluiert wird; und auch, wo zukünftig mehr und besser evaluiert werden muss.

Die DeGEval organisiert folgerichtig Evaluator/inn/en in sehr unterschiedlichen Feldern. Sie versammelt – vorwiegend in Deutschland und Österreich – aktuell gut 600 Personen, die Evaluationen durchführen, in Auftrag geben oder ein grundsätzliches Interesse an ihnen haben, sowie mehr als 160 institutionelle Mitglieder. Die Mitglieder kommen aus renommierten wirtschafts- und sozialwissenschaftlichen Forschungseinrichtungen, Hochschulen, Unternehmensberatungen und Politikberatungsunternehmen sowie Ministerien, Verwaltungen und Ressortforschungseinrichtungen.

Evaluation in Deutschland und Österreich steht heute – nach Gründung der DeGEval im Jahr 1997 – vor der Aufgabe, die eigene Arbeit stärker zu „professionalisieren“. Es geht nicht nur darum, die Güte von Evaluation zu sichern, es geht auch um die Frage, ob Evaluation sich als „Beruf“ definieren kann, der erlernt und entsprechend zertifiziert werden muss. Oder kann sich, wer immer Maßnahmen, Programme, Organisationen bewertet, Evaluator/in nennen?

Wir wissen noch immer zu wenig darüber, was und wer sich tatsächlich unter dem Begriff Evaluation versammelt, über welche Kompetenzen diejenigen verfügen, die Evaluationen verkaufen und was mit Evaluationen in der Folge geschieht. Allein diese Fragen markieren einen immensen Bedarf an Forschung über Evaluation.

Der vorliegende Band geht einen kleinen Schritt in diese Richtung. Hier beschreiben Protagonisten, was Evaluation in verschiedenen Politikfeldern zur Entwicklung von Programmen, Projekten, Organisationen u.ä. beiträgt. Die Perspektive ist dadurch geprägt, dass hier aktive Mitglieder eines Verbandes berichten, der sich unter anderem zur Aufgabe gesetzt hat, Evaluation zu professionalisieren. Ein weiterer Schritt auf diesem Weg

könnte sein, Auftragnehmer/innen und Auftraggeber/innen auf Standards der Evaluation zu verpflichten oder Standards wenigstens derart zu etablieren, dass sie selbstverständlicher Bezugsrahmen einer Reflexion über die Qualität von Evaluationen werden.

Die Berichte zu den unterschiedlichen Politikfeldern, die in diesem Band versammelt sind, wurden von Vertreter/innen der Arbeitskreise der DeGEval (vgl.: www.degeval.de/arbeitskreise) erstellt. Sie machen deutlich, wie unterschiedlich der Stand der Evaluierung in den einzelnen Feldern ist.

Die öffentliche Verwaltung mit ihren unterschiedlichen Ebenen (Bund, Länder, Kommunen) und ihrem Zusammenspiel mit gesellschaftlichen Partnern (z.B. Sozialversicherung, Wissenschaft) ist seit den späten 1960er Jahren zum Gegenstand von Evaluation und Evaluationsforschung geworden. Zunehmend wurden und werden unterschiedliche Aufgaben und Funktionen von Politik und Verwaltung evaluiert. Politik und Verwaltung haben die Bedeutung von Evaluationen als wichtiges Instrument zur Sammlung und Bewertung von Informationen im politischen Entscheidungsprozess erkannt. Mit ihrer Hilfe werden die Wirkungen oder Folgen von Gesetzen, politischen Programmen etc. ex ante abgeschätzt bzw. ex post analysiert (z.B. Gesetzes-Folgenabschätzung). Es ist davon auszugehen, dass Evaluationen in der öffentlichen Verwaltung auch in Zukunft eine wichtige Rolle spielen werden. Der Legitimationsdruck auf Politik und Verwaltung bleibt – nicht zuletzt aufgrund der beschränkten finanziellen Ressourcen – weiterhin hoch. Aus diesem Grund wird es auch zukünftig notwendig sein, politische Maßnahmen (z.B. Gesetze, Programme) sowohl bei ihrer Planung als auch bei ihrer Durchführung auf ihre Wirksamkeit und Effizienz unter Berücksichtigung möglicher nichtintendierter Nebenwirkungen zu überprüfen.

Das Feld der *Stadt- und Regionalentwicklung* grenzt sich vorwiegend über den territorialen Bezug ab, es umfasst inhaltlich sehr unterschiedliche Bereiche (z.B. Städtebau, Stadt- und Regionalplanung, Umweltplanung, Raumordnung). Es handelt sich hierbei um querschnittsorientierte Handlungsfelder mit vielfältigen Verknüpfungen in andere Bereiche. Evaluation gewinnt hier seit den 1970er/1980er Jahren stetig an Bedeutung und an Qualität. Es kann aber auch festgestellt werden, dass Evaluationen hier häufig von Legitimations- und Kontrollfunktionen geprägt sind. Herausforderungen im Feld liegen zum einen im methodischen Bereich und zum anderen in der Akteursstruktur. Kritisch wird thematisiert, dass Evaluation auf eine Symbolfunktion reduziert werden könnte. Ein Ausweg könnte in stärker selbstreflexiven Evaluierungsansätzen liegen, die ihre eigene Rolle im Feld bedenken.

Das Feld der *Strukturpolitik* ist stark durch die einschlägige europäische Politik beeinflusst. Gegenstand sind in aller Regel relativ komplexe (aus einer Vielzahl von Einzelinstrumenten bestehende) Förderprogramme, die die Bereiche Regionalpolitik, regionale Wirtschaftspolitik, Arbeitsmarktpolitik und Agrarpolitik betreffen. Die letzten 20 Jahre sind von einem stetigen Aufbau der Evaluierungskapazitäten und einer Ausdifferenzierung des Evaluierungssystems geprägt. Die Evaluierungen werden zumeist auf Ebene der Bundesländer durchgeführt und sind relativ stark auf Umsetzungsaspekte konzentriert, demgegenüber werden Wirkungen eher selten umfassend evaluiert. Innerhalb des Feldes ist der systematische Diskurs über Evaluierung eher schwach ausgeprägt. Sowohl Nutzung von Evaluierungsergebnissen als auch Qualitätsentwicklung könnten durch einen Ausbau profitieren.

Kaum ein anderes Politikressort wird so oft evaluiert wie das der *Entwicklungspolitik* und ist gleichzeitig so umstritten. Spiegelbildlich zum Wandel entwicklungspolitischer Ansätze – etwa von kleineren Projekten hin zu flächendeckenden Programmen – wandelte sich innerhalb der letzten drei Dekaden auch die Evaluierungsmethodik und wurde zugleich internationaler und professioneller. Ein wichtiges Thema im EZ-Ressort ist die von der Öffentlichkeit geforderte Unabhängigkeit der Evaluierung. Institutionell stellt die Gründung des Deutschen Instituts für Entwicklungsevaluierung (DEval) im Jahr 2012 die Konsequenz dieser Diskussion dar. Durch die stetige Weiterentwicklung des Politikfeldes und die neuen Möglichkeiten, die ein eigenes Institut für Entwicklungsevaluierung insbesondere für die weitere Professionalisierung bietet, sind zukünftig möglicherweise weitere Impulse zu erwarten.

Eine Evaluationskultur, wie sie in anderen Bereichen wie z.B. in der Entwicklungszusammenarbeit, in der Bildungs- oder auch Arbeitsmarktpolitik schon lange existiert, gibt es im Feld der *Kultur und Kulturpolitik* in Deutschland und Österreich noch nicht. Für die Darstellung aktueller Entwicklungen werden für Deutschland die Bereiche Kunst, Kultur und Kulturpolitik und Auswärtige Kultur- und Bildungspolitik eingehender betrachtet. Im Anschluss folgt ein Exkurs zur Situation in Österreich. Als besondere Herausforderung wird die Entwicklung kultursensitiver Konzepte (z.B. partizipative Verfahren oder auch Methoden der ästhetischen Forschung) herausgestellt. Stärker als bisher sollten sowohl Anforderungen als auch Lernmöglichkeiten genutzt werden, die sich aus der Heterogenität des Feldes und der Akteure ergeben. Eine weitere große Herausforderung ergibt sich aus der Schwierigkeit, einzelne, wichtige Akteursgruppen (Vertreter aus Kunst und Kultur) zu erreichen.

„Evaluation“ ist im *Schulwesen* des deutschsprachigen Raums heute ein allgegenwärtiger Begriff, wenn auch mit schillernden Inhalten. Die Anfänge reichen in die 1970er Jahre zurück, wo es im Zuge des bildungspolitischen Aufbruchs und einer Zusammenarbeit von Bildungspolitik und Bildungsforschung zur wissenschaftlichen Begleitung von Schulversuchen und Begutachtung von Reformprojekten durch ministeriellen Auftrag kam. Das Methodenspektrum ist umfassend und reicht von schlichten mündlichen Formaten (Feedbacks, Auswertungsgespräche) in schulischer Selbstevaluation bis zu groß angelegten Studien mit hohen wissenschaftlichen Ansprüchen und Teilnehmerzahlen im sechsstelligen Bereich wie im Falle von Schulleistungsstudien. Die Herausforderung der Zukunft wird darin liegen, die Steuerung im Mehrebenensystem des Schulwesens zwischen „Top-down“ (zentrale Standardsetzung und Kontrollmechanismen) und „Bottom-up“ (Qualitätsentwicklung und Qualitätssicherung auf Schulebene) auszubalancieren und dabei den Beitrag von Evaluation neu zu justieren.

Zur *beruflichen Bildung* zählen die berufliche Erstausbildung, die in Deutschland und Österreich überwiegend in Form der dualen Ausbildung, also an den beiden Lernorten „Betrieb“ und „Berufsschule“, aber auch als vollqualifizierende schulische Berufsausbildung stattfindet sowie das weniger stark regulierte Feld der beruflichen Weiterbildung. Evaluation in der beruflichen Bildung ist seit dem Beginn der 1970er Jahre aufgrund entsprechender gesetzlicher Regelungen fest verankert. Evaluationen in den verschiedenen Feldern der beruflichen Bildung setzen sich momentan vorwiegend mit Fragen der Wirksamkeit und der Effizienz auseinandersetzen. Evaluationsgegenstände – auf Bundesebene vor allem Programme – werden immer komplexer (Vielzahl von Zielen/Zielebenen, Verzahnung mit anderen Programmen, Vielzahl von Akteuren,

Programme die aus vielen Einzelprojekten bestehen), was zunehmende Herausforderungen an eine komplexe Evaluation nach sich zieht. Damit einher gehen zunehmende Rollenanforderungen an Evaluierende, die zum Teil dem Gedanken der Unabhängigkeit widersprechen.

Im Feld der *Hochschule* wird Evaluation zum einen als interne (Selbst-)Evaluation von Lehrveranstaltungen und Studiengängen, zum anderen als externe Akkreditierung vorwiegend von Studiengängen durchgeführt. Evaluation ist hier eng mit dem Qualitätsmanagement innerhalb der jeweiligen Organisationen verknüpft. Durch die starke interne Verankerung stellt sich aber ständig das Problem der Rollendefinition der Evaluierenden: Einerseits müssen sie sich im Verhältnis zu den Wissenschaftlern, andererseits zur Hochschulverwaltung positionieren. Zur Evaluierung von Lehrveranstaltungen kommen meist Designs zum Einsatz, die auf der Ebene der einzelnen Teilnehmer ansetzen. Die Bewertung von Studiengängen erfolgt meist auf Basis der Konzepte, teils auch mit Befragungen von Studierenden, jedoch kaum ohne den Kontext eines Akkreditierungsverfahrens. Schließlich erheben Hochschulen auch Verbleibsdaten ihrer Absolventen. Forschung wird entweder über Peer-Review oder bibliometrische Verfahren analysiert. Herausforderungen bestehen in einer besseren Methodik zur Forschungsvaluierung sowie der Verknüpfung der Evaluierung von Forschung und Lehre.

Im Arbeitsfeld *Gesundheit* steht die Beschäftigung mit der Evaluation von Gesundheitsförderungs- und Präventionsmaßnahmen im Vordergrund. Diese werden häufig in Form von Kampagnen durchgeführt. Bundesweit bekannte Kampagnen der Bundeszentrale für gesundheitliche Aufklärung (BZgA) sind z.B. „Gib Aids keine Chance“ oder „Alkohol? Kenn dein Limit.“ Eine besondere Herausforderung der Evaluation in diesem Feld liegt in der Komplexität, da sich auf der einen Seite die Programme aus einer Vielzahl von Interventionen zusammensetzen, die sich wechselseitig beeinflussen, und auf der anderen Seite diese Maßnahmen in einem komplexen Kontext implementiert werden. Während zum Beispiel im Bereich der Arzneimittelforschung Produktentwicklung und Evaluation heutzutage Hand in Hand gehen, sind Voraussetzungen, um belastbare Aussagen treffen zu können, im Feld der Gesundheitsförderung und Prävention in der Regel nicht gegeben. Insgesamt wird im Feld das Erfordernis von guten Evaluationen zunehmend anerkannt. Die Haltung der Auftraggebenden, Evaluationen von Anfang an mitzudenken, ist dazu verstärkt erforderlich.

Das Feld der *Sozialen Dienstleistungen* war lange von Selbstevaluationen geprägt. Erst in den letzten Jahren gibt es eine zunehmende Auseinandersetzung mit dem Evaluationsthema, insbesondere auch der externen Evaluation. Eine Besonderheit des Feldes liegt darin, dass die sozialen Dienstleistungen stets in Ko-Produktion zwischen Dienstleistungserbringern und den Zielgruppen erbracht werden. Und die Praxisfelder sind sehr unterschiedlich und umfassen die Kinder- und Jugendhilfe, Altenhilfe, sozialpädagogische Angebote im Gesundheitssystem sowie erwachsenenbezogene soziale Hilfen. Der Schwerpunkt der Evaluationstätigkeit liegt in der Kinder- und Jugendhilfe. Die Herausforderung für eine breitere Verankerung der Evaluationsfunktion besteht darin, den Spagat zwischen der starken Abhängigkeit vom jeweiligen Kontext einerseits und den Anforderungen an gute Evaluierungen andererseits zu meistern.

Staatliche Interventionen in *Forschung, Technologie und Innovation (FTI)* haben eine lange Tradition und bilden heute ein eigenständiges Politikfeld. Seit Jahrzehnten

kommen hier auch Verfahren der Evaluation zur Anwendung, die nach der wissenschaftlichen oder technologischen Qualität und dem Erfolg von FTI-Vorhaben oder entsprechenden Institutionen fragen und unterschiedliche Effekte öffentlich finanzierter Maßnahmen untersuchen. Gegenstand der Evaluation im Feld der FTI-Politik sind in erster Linie Fördermaßnahmen und die durch sie geförderten Projekte und Fördernehmer. Ebenso evaluiert werden Institutionen des Innovationssystems, z.B. Netzwerke (Cluster, Forschungsverbände u.ä.) sowie Forschungseinrichtungen. Evaluationen sehen sich ähnlichen Herausforderungen gegenüber, wie Evaluationen in den meisten anderen Politikfeldern: Begrenzte zeitliche und finanzielle Ressourcen bei gleichzeitig umfangreichen Anforderungen, zum Teil unzureichende Datenverfügbarkeit und eingeschränkter Zugang zu vorhandenen Daten.

Das Evaluationsfeld *Wirtschaft* umfasst Evaluationsaktivitäten, die sich auf privatwirtschaftliche Unternehmen sowie Intermediäre im Bereich der Wirtschaft wie beispielsweise Verbände, Kammern, Wirtschaftsförderer oder Stiftungen beziehen. Noch ist Evaluation in der Wirtschaft ein vergleichsweise unbekanntes Terrain und findet im privatwirtschaftlichen Kontext als Steuerungs- und Qualitätssicherungsinstrument bisher wenig Anwendung. Grundsätzlich dient sie insbesondere zur Beantwortung der Frage, wie Ziele von Programmen bzw. Projekten im Wirtschaftskontext besser erreicht und nichtintendierte negative Auswirkungen vermieden werden können. Evaluationen in der Wirtschaft werden tendenziell für bestimmte Maßnahmen eingesetzt (z.B. Personalentwicklung, Messung von PR, Einführung Arbeitsschutz) in den meisten Fällen findet eine Evaluation jedoch eher sporadisch statt, teilweise auch unter anderer Bezeichnung (z.B. Strategieentwicklung, Technologieanalyse). Dabei wird Evaluation im wirtschaftlichen Umfeld häufig mit der Rechtfertigung von politischen Maßnahmen und insofern eher mit einem aufwendigen bürokratischen Prozess assoziiert. Eine der wesentlichen Herausforderungen in diesem Arbeitsfeld ist daher darin zu sehen, den bestehenden Informationsbedarf zu decken und Nutzen und Ziele von Evaluationen in der Sprache der Wirtschaft zu formulieren.

Evaluationen von *umweltpolitischen Maßnahmen* zeichnen sich durch ein vergleichsweise hohes Maß an Komplexität aus, da es sich häufig um Querschnittsmaßnahmen handelt, die zum Beispiel darauf ausgelegt sind, die negativen Auswirkungen von sektorspezifischen Aktivitäten zu minimieren. Sie sind häufig geprägt von hohen Interessenlagen, die in der Regel mit Zielkonflikten einhergehen, und sie orientieren sich zudem an sehr langen Zeiträumen und Auswirkungen, die global zu beobachten sind und nicht an den Grenzen einer Politikmaßnahme halt machen. Die Vielfältigkeit des Umweltthemas spiegelt sich damit auch in der Evaluationspraxis. Unterschiede ergeben sich aus unterschiedlichen Evaluationskulturen der verschiedenen Disziplinen und durch neuere Erkenntnisse über die Klimawirkungen und deren Zusammenhänge. Deshalb wächst die Notwendigkeit, fächerübergreifend zu kooperieren

Neben der Ausdifferenzierung der Politikbereiche hat sich im Verlauf der Entwicklung der DeGEval auch die Notwendigkeit gezeigt, bestimmte *Querschnittsbereiche* durch die Gründung von Arbeitskreisen abzudecken. Hier wird deutlich, dass es (mindestens) drei Themen gibt, die für alle Politikfelder und die Evaluation als „Profession“ bedeutend sind.

Aus dem Bericht zu dem Arbeitsfeld *Methoden* wird die Verknüpfung mit den methodischen und methodologischen Debatten der Sozialwissenschaften deutlich. Der

enge Bezug zur wissenschaftlichen Methodologie wird aber praktisch immer wieder zum kritischen Thema, denn die dort entwickelten Methoden, Strategien und Techniken müssen auf die spezifischen Bedürfnisse von Evaluation bezogen und daraufhin angepasst werden. Evaluation ist eben u.a. auch dadurch gekennzeichnet, dass sich Anforderungen und auch Beschränkungen dadurch ergeben, dass Evaluationen auftragsgebunden sind. Der Beitrag beschreibt die Herausforderung, Standards und Gütekriterien mit den Besonderheiten einer Evaluation und ihre Eingebundenheit in Stakeholder-Interessen zu vereinbaren. Insbesondere wird auch die Gefahr diskutiert, dass diese Spannung womöglich zu problematischer und unwissenschaftlicher Evaluationspraxis führen könnte. Das ständige Bemühen unterschiedliche, einerseits akademische, andererseits administrativ-politische Diskurswelten in eine konstruktive Beziehung zu bringen, „setzt sowohl fachwissenschaftliche und methodische Kompetenz als auch eine Offenheit bezüglich des evaluativen Vorgehens voraus – beides konstitutive Elemente des Handelns im Bereich der Evaluation“.

Gender Mainstreaming bezeichnet die Strategie, geschlechtsspezifische Benachteiligungen durch die durchgängige Berücksichtigung der Geschlechterperspektive in allen Politik- und Handlungsfeldern zu beseitigen. Seit 1999 ist Gender Mainstreaming um das Ziel der Gleichstellung der Geschlechter zu erreichen, für alle Mitgliedstaaten der Europäischen Union als vertragliche Verpflichtung festgeschrieben. Daraus ergibt sich auch die Pflicht, Gender Mainstreaming in der Evaluation zu verankern. Nach einer kurzen historischen Verortung werden aktuell zum Einsatz kommende Evaluationsinstrumente vorgestellt, wie Gender Impact Assessment und Gender Budgeting, mit denen sich die Evaluationsfragestellungen immer mehr von der Implementation hin zur Wirkung verschoben haben. Besondere Herausforderungen liegen darin, die Geschlechterperspektive durchgängig zu berücksichtigen sowie die Genderkompetenz von Auftraggebenden und Evaluierenden zu stärken.

Viele der Herausforderungen, die in den jeweiligen Arbeitsfeldern geschildert werden, sind auch als Herausforderungen für die *Aus- und Weiterbildung* von Evaluatoredinnen und Evaluatoren interpretierbar. Hier wird ein enormer Entwicklungsbedarf deutlich: „Wie wird man eigentlich Evaluator oder Evaluatorin?“ lautet die Ausgangsfrage aus dem Arbeitsfeld, und es wird schnell deutlich, wie schwierig eine Antwort ist. Weder in Deutschland und Österreich, noch in anderen Ländern, kann man Evaluation als „Profession“ adressieren. In der Evaluation verfügen wir weder über eine Regulierung des Marktzugangs noch über standardisierte Ausbildungsgänge. Ein Faktor für diesen Tatbestand liegt in der Transdisziplinarität und Heterogenität des Anwendungsfelds Evaluation.

Mit dem letzten Beitrag ist die womöglich wesentliche Frage adressiert, die sich die Mitglieder der DeGEval stellen müssen: Sind die im vorliegenden Band deutlich werdende Vielfalt der disziplinären und inhaltlichen Bezüge der Evaluation, die Komplexität der Prozesse und Ziele in politischen Kontexten oder die konkurrierenden methodischen Prioritäten Gründe, die dazu führen können, ein Ungleichgewicht zwischen Expansion von Evaluationen auf der einen Seite und unzureichender Professionalisierung auf der anderen Seite hervorzurufen?

Mit den Standards der Evaluation hat die DeGEval ein wichtiges Instrument für die Qualität von Evaluation vorgelegt. Die aktuell stattfindende Revision der Standards verdeutlicht deren Bedeutung. Basierend auf den Standards liegen auch „Empfehlun-

gen für die Aus- und Weiterbildung in der Evaluation“ vor (DeGEval 2004). Dieses Dokument konkretisiert die Standards und macht deutlich, dass sie „lernbar“ und „lehrbar“ sind. Hier werden die Kompetenzen von Evaluatorinnen und Evaluatoren erörtert und damit wird eine hilfreiche Orientierung für die Aus- und Weiterbildung geliefert.

Als einen wichtigen Schritt auf dem Weg der zukünftigen Sicherung von Güte von Evaluation betrachtet die DeGEval die Stärkung des Nachwuchses. Seit 2008 hat sich aus einer Initiative des AK Aus- und Weiterbildung ein Netzwerk für den Nachwuchs in Wissenschaft und Praxis gebildet, welches sich 2013 offiziell als „Nachwuchsnetzwerk“ unter dem Dach der DeGEval konstituiert hat. Das Nachwuchsnetzwerk richtet sich zum einen an Personen, die im Rahmen von Qualifizierungsarbeiten über Evaluation (z.B. ihre Formen, Methoden, Prozesse, Resultate etc.) forschen. Das Nachwuchsnetzwerk kooperiert darüber hinaus mit den Arbeitskreisen der DeGEval, um den Nachwuchs an die Diskussionslinien innerhalb der DeGEval heranzuführen sowie die organisatorische und inhaltliche Struktur der DeGEval transparent zu machen. Als zentrale Elemente der Vernetzung und des Austausches haben sich das jährliche Treffen auf der Jahrestagung der DeGEval und ein Forschungs- und Praxiskolloquium etabliert, bei welchem Nachwuchswissenschaftlerinnen und -wissenschaftler ihre Forschungsarbeiten und Nachwuchsevaluatorinnen und -evaluatorennen ihre Evaluationsprojekte vorstellen und miteinander diskutieren können.

Der Beitrag aus dem Arbeitskreis Aus- und Weiterbildung formuliert eine Aufgabe, die sich auf Basis des hier vorliegenden Sammelbandes eindrucksvoll bestätigen lässt: „Eine zentrale Frage für die weitere Entwicklung wird aber sein, wie sich die Gesellschaft und ihre Mitglieder zukünftig zur Notwendigkeit weitergehender Professionalisierungsschritte positionieren“. Die DeGEval nimmt sich der Aufgabe an, diese Frage mit ihren Mitgliedern zu diskutieren. Es wird sich zeigen, ob das in eine Richtung verläuft, wie sie sich international zumindest vorsichtig abzeichnet: Beglaubigung, Begutachtung oder Zertifizierung von Evaluatorinnen und Evaluatoren. Auch andere Ansatzpunkte sind denkbar: die Unterstützung hochwertiger Aus- und Weiterbildungsangebote, die Stärkung des Diskurses über Qualität von Evaluationen und generell die Erhöhung der Transparenz. Die DeGEval beginnt diese Debatte auf ihrer in Kooperation mit der SEVAL in Zürich durchgeführten Jahrestagung 2014.

Literatur

- Deutsche Gesellschaft für Evaluation (DeGEval). (2002). *Standards für Evaluation*. Köln: DeGEval – Gesellschaft für Evaluation e.V.
- Deutsche Gesellschaft für Evaluation (DeGEval). (2004). *Empfehlungen für die Aus- und Weiterbildung in der Evaluation. Anforderungsprofile an Evaluatorinnen und Evaluatoren*. Köln: DeGEval – Gesellschaft für Evaluation e.V.
- Joint Committee on Standards for Educational Evaluation (JCSEE). (1994). *The program evaluation standards. How to assess evaluations of educational programs*. Thousand Oaks, CA: Sage.