

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

WAXMANN

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



Waxmann 2014
Münster • New York

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

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

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

© Waxmann Verlag GmbH, 2014

www.waxmann.com

info@waxmann.com

Umschlaggestaltung: Inna Ponomareva, Münster

Umschlagabbildung: © karandaev – Fotolia.com

Satz: Stoddart Satz- und Layoutservice, Münster

Druck: Hubert & Co., Göttingen

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706



Printed in Germany

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.

Inhalt

<i>Wolfgang Böttcher, Christiane Kerlen, Peter Maats, Oliver Schwab, Sonja Sheikh</i> Arbeitsfelder und Herausforderungen der Evaluation	7
<i>Gottfried Konzendorf, Axel Piesker, Renate Reiter</i> Evaluation in der öffentlichen Verwaltung	15
<i>Manfred Rolfes, Jan Lorenz Wilhelm</i> Evaluationspraxis und Evaluationsforschung im Kontext der Stadt- und Regionalentwicklung.....	21
<i>Regina Grajewski, Stefan Meyer</i> Stand der Evaluation in der Strukturpolitik	37
<i>Susanne Neubert, Dorothee Mack, Helge Roxin</i> Vom Methodenstreit zum Methodenmix – Chronologie und Stand der Entwicklungsvaluierung in Deutschland	61
<i>Tülin Engin, Vera Hennefeld, Ute M. Metje, Tanja Nagel</i> Zur Evaluationskultur in Kultur und Kulturpolitik.....	73
<i>Michael Fraiss, Monika Renz</i> Evaluation im Schulbereich	89
<i>Jürgen Götzhaber, Peter Jablonka, Michael Kalman</i> Evaluation in der beruflichen Bildung: Ein weites Feld	99
<i>Susan Harris-Huemmert, Lukas Mitterauer, Philipp Pohlenz</i> Evaluation im Kontext der Hochschule.....	107
<i>Karin Waldherr, Marcus Capellaro, Monika Finsterwald, Ludwig Grillich, Wolfgang Haß, Joseph Kuhn, Mary Lindner, Boris Orth, Vera Popper, Ursula von Rüden, Georg Spiel, Jürgen Töppich, Annekatri Wetzstein</i> Evaluation im Bereich der Prävention und Gesundheitsförderung: Stand, Herausforderungen und Entwicklungsperspektiven in Deutschland und Österreich.....	117

Edith Halves, Marianne Lück-Filsinger, Stefan Schmidt
Evaluation in der Sozialen Arbeit: Entwicklungen und Herausforderungen.....133

*Michael Astor, Iris Fischl, Judith Hoffmann, Gesa Koglin,
Marianne Kulicke, Sonja Sheikh, Jan Wessels, Katy Whitelegg*
Evaluation von Forschungs-, Technologie und Innovationspolitik in
Deutschland und Österreich – ein Überblick.....139

Martin Rost, Julia Hapkemeyer, Sonja Kind
Evaluation im Feld der Wirtschaft – Entwicklungslinien und aktuelle
Herausforderungen151

Peter Kaufmann, Leo Urban Wangler
Evaluation in der Umweltpolitik am Beispiel Klimapolitik.....159

Christian Erzberger, Udo Kelle
Methodisch handeln im Evaluationsprozess171

Maria Gutknecht-Gmeiner, Angela Wroblewski
Evaluation von Gender Mainstreaming und Gendern von Evaluation179

Jan Ulrich Hense, Heike Steckhan
Aus- und Weiterbildung in Evaluation: Vergangenheit – Gegenwart – Zukunft?.....191

Die DeGEval und ihre Arbeitskreise207

Autorinnen und Autoren.....215

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, Forschungsverbünde 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 Evaluationeninstrumente 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 Auftraggebern 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.

Evaluation in der öffentlichen Verwaltung

1. Öffentliche Verwaltung als Gegenstand von Evaluation

Die öffentliche Verwaltung unterschiedlicher Ebenen (Bund, Länder, Kommunen) und das Zusammenspiel mit gesellschaftlichen Partnern (z.B. Sozialversicherung, Wissenschaft) sind 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 selbst übernehmen in Evaluationsprozessen unterschiedliche Rollen (z.B. als Auftraggeber, als Nutzer von Evaluationsergebnissen oder als Evaluationsgegenstand) – oftmals im Zusammenspiel mit gesellschaftlichen Partnern. Die Schnittstellen zwischen Politik, Verwaltung, Wissenschaft und Zivilgesellschaft haben daher eine erhebliche Bedeutung bei der Durchführung von Evaluationen. Schlüsselfragen dabei lauten:

- Welche Funktionen und Aufgaben übernehmen Politik und Verwaltung bei der Initiierung, Vergabe und Durchführung von Evaluationen?
- Wie können diese Funktionen und Aufgaben qualifiziert wahrgenommen werden?
- Wie kann ein optimales Zusammenspiel der an Evaluationen beteiligten Akteure gewährleistet werden?

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).

Ein gewichtiges Augenmerk ist auf die Nutzung der Evaluationsergebnisse im politischen Entscheidungsprozess zu richten. Im Idealfall dienen Evaluationen der Verbesserung von öffentlichem Politik- und Verwaltungshandeln i.S. einer erhöhten Wirksamkeit (Effektivität) von politischen Maßnahmen, einer Steigerung der Wirtschaftlichkeit (Effizienz) oder auch einer Erhöhung der Legitimität öffentlichen Handelns. In diesem Sinne interessiert sich die auf die öffentliche Verwaltung gerichtete Evaluationsforschung insbesondere für folgende Fragen:

- Wie kann der Einfluss von Evaluationen im politischen Entscheidungsprozess gewährleistet werden?
- Wie geht die öffentliche Verwaltung mit den Ergebnissen von Evaluationen um?

Es ist leicht, diese Fragen zu stellen. Ihre Beantwortung ist jedoch ungleich schwieriger, und es besteht nach wie vor erheblicher, vor allem empirischer Forschungsbedarf.

2. Historische Entwicklung

Das Interesse von Evaluationspraktikern und Evaluationsforscher/innen/n an der öffentlichen Verwaltung geht eng einher mit der Entwicklung der Evaluationsforschung selbst. Verwaltung und Verwaltungshandeln (in seinen höchst unterschiedlichen Facetten und Interventionsfeldern) sind Gegenstände von Evaluationen und ihrer sozialwissenschaftlichen Erforschung, seitdem sich Politiker/innen, Verwaltungsakteure und (Sozial-)Wissenschaftler/innen ab den 1970er Jahren verstärkt für eine systematische Überprüfung der Wirkungen von öffentlichen Politikprogrammen, -instrumenten und Verwaltungsakten zu interessieren begannen. Dieses Interesse an der Evaluation öffentlicher Verwaltung wiederum hatte verschiedene Gründe:

Seit Ende der 1960er Jahre entwickelte sich innerhalb der deutschen Staats- und Verwaltungswissenschaft ein Diskurs über die politische Rolle der öffentlichen Verwaltung und ihre Funktionen. In Anknüpfung an diesen Diskurs fand ein rationales Politikmodell Verbreitung, das von der Idee einer objektiven Plan- und rationalen Steuerbarkeit öffentlicher Politik und administrativen Handelns getragen wurde (Wollmann, 2009, S. 384). Demnach vollzieht sich Politik idealtypisch als geordneter Zyklus der politisch-administrativen Themenfindung, Programmformulierung, Implementation sowie Ergebnis- bzw. Wirkungsanalyse (Evaluation) mit der Folge einer eventuellen Programmkorrektur (Janning & Schneider, 2006, S. 48–63).

In diesem Kontext entstand die auf öffentliches Verwaltungshandeln gerichtete Implementations- und Evaluationsforschung in Deutschland (Mayntz, 1980). Ziel war es einerseits, die öffentliche Verwaltung dabei zu unterstützen, geeignete Organisationsstrukturen und Prozesse zur Erfüllung der ihr politisch übertragenen Aufgaben zu finden. Andererseits ging es um eine systematische Überprüfung des öffentlichen Verwaltungshandelns (z.B. im Hinblick auf seine Zielerreichung und Wirkungen), die über die reine Kontrolle der Recht- und Ordnungsmäßigkeit sowie der Angemessenheit administrativer Akte hinausreichen sollte. *Zudem* bildete der kontinuierliche Wandel der wirtschaftlichen, technologischen und soziodemographischen Rahmenbedingungen staatlichen Handelns eine Triebfeder für die Analyse der öffentlichen Verwaltung als Gegenstand von Evaluation. Im Zuge der Wirtschaftskrise Mitte der 1970er Jahre setzte in Deutschland (wie in anderen westlichen Wohlfahrtsstaaten) eine öffentliche Diskussion über die Grenzen staatlicher Interventionsfähigkeit ein. Mit der Diskussion um die Steuerungsfähigkeit des Staates wurde zunehmend das Verwaltungshandeln zum Gegenstand wissenschaftlicher Analysen. Verwaltungsreformen sind seither zentraler Gegenstand der Verwaltungswissenschaft. Gefragt wird nach der Wirksamkeit, Wirtschaftlichkeit und Legitimation öffentlichen Verwaltungshandelns sowie nach Optimierungsmöglichkeiten (Holtkamp, 2012). Die Ausdifferenzierung unterschiedlicher Varianten oder Typen der Verwaltungsevaluation wurde nicht zuletzt durch diese Fragestellungen angeregt (Ex-post- oder summative Evaluation, Ex-ante-Evaluation/Kosten-Nutzen-Analyse/[Gesetzes-]Folgenabschätzung, Evaluier-

barkeitsabschätzung, formative oder On-going-Evaluation, Begleitforschung, Monitoring, Erfolgskontrolle, Effizienzanalyse) (vgl. Wollmann, 2009, S. 382–383).

3. Evaluationsformen, Akteure und Standards

Evaluationen im Bereich der öffentlichen Verwaltung können auf zwei unterschiedlichen Ebenen ansetzen.

Erstens können Evaluationen auf Ebene der Verwaltungspolitik ansetzen, die darauf abzielt, organisatorische, personelle, instrumentelle und prozedurale Strukturen innerhalb der Verwaltung zu verändern. In diesem Zusammenhang wird dann untersucht, inwieweit diese Änderungen zu einer Steigerung der Effektivität, Effizienz etc. des Verwaltungshandelns beigetragen haben und welche weiterreichenden Effekte durch sie bewirkt wurden (Jann, 2001, S. 333; Wollmann, 2002, S. 75–77; Reiter et al., 2011). Als Beispiel kann hier auf das Konzept *Better Regulation* verwiesen werden, das zur Einführung bzw. verstärkten Nutzung von Folgenabschätzungen, Ex-post-Evaluationen, Konsultationen etc. beigetragen hat (vgl. hierzu ausführlich: Wegrich, 2012).

Zweitens geht es um die Evaluation konkreter Maßnahmen (Gesetze, Programme, Instrumente etc.) in bestimmten Politikfeldern (z.B. Wirtschaftspolitik, Innere Sicherheit). Zur Bewertung der jeweiligen Maßnahme werden Kriterien wie Effektivität, Effizienz und die nicht-intendierten Wirkungen herangezogen.

Bei der Planung und Durchführung von Evaluationen spielt die öffentliche Verwaltung eine zentrale Rolle. Je nach Konzeption eines Evaluationsvorhabens tritt sie als Auftraggeberin, als Impuls- und Konzeptgeberin, als Informationsquelle und/oder als Nutzerin der Evaluationsergebnisse auf. Grundsätzlich lassen sich zwei Evaluationsvarianten unterscheiden. Zum einen kann die öffentliche Verwaltung Evaluationen selbst durchführen und ist damit für alle Phasen des Evaluationsprozesses alleine verantwortlich. Dabei kommt es vor, dass die öffentliche Verwaltung für die Bearbeitung bestimmter – insbesondere methodischer – Fragestellungen auf externen Sachverstand zurückgreift. In diesen Fällen wird auch von einer sogenannten Methodenberatung gesprochen (Ziekow, Debus & Piesker, 2013, S. 30). Zum anderen kann die Planung und Durchführung von Evaluationsvorhaben nach außen gegeben werden. Die öffentliche Verwaltung (Bundes-, Landes- und Kommunalverwaltung) tritt somit als Auftraggeberin von externen Evaluationen auf. Zudem werden z.B. Gesetzesfolgenabschätzungen mittels gesetzlicher Evaluationsklauseln von den Parlamenten (Bundestag, Landtage) veranlasst. Da externe Evaluationen in den letzten Jahren – aufgrund internationaler Einflüsse und einem erhöhten Legitimationsbedarf – an Bedeutung gewonnen haben, hat sich hier ein neues Betätigungsfeld für unterschiedliche Akteure herausgebildet. Zu nennen sind Universitäten sowie außeruniversitäre Forschungseinrichtungen, die Evaluationsprojekte vielfach als drittmittelfinanzierte Auftragsforschung bearbeiten sowie privatwirtschaftlich arbeitende Beratungsunternehmen.

Die Einhaltung von Standards bei der Durchführung von Evaluationen ist kein gesondert diskutiertes Thema im Evaluationsfeld. Vielmehr werden die Standards der DeGEval bei Evaluationsvorhaben – sofern möglich – implizit berücksichtigt. Eine aktive Auseinandersetzung mit den Standards findet jedoch bislang nicht statt.

4. Perspektiven der Evaluation öffentlicher Verwaltung sowie ihrer Erforschung

Evaluationen im Bereich der öffentlichen Verwaltung sehen sich mit spezifischen Anforderungen des politischen und administrativen Systems (z.B. mit dem Wahlzyklus, mit der Komplexität des Akteursfeldes und mit dem hierarchischen Aufbau der öffentlichen Verwaltung) konfrontiert. Daher ist es von besonderer Bedeutung, Erkenntnisse darüber zu sammeln, zu welchem Zeitpunkt und auf welche Weise Evaluationen am besten durchzuführen sind. In diesem Zusammenhang ist es erforderlich, sich stärker als bisher mit der Frage zu beschäftigen, wie Evaluationsergebnisse tatsächlich von politischen Entscheidungsträgern genutzt werden und wie ggf. die Nutzungsmöglichkeiten verbessert werden können. Dies ist sowohl aus Sicht der Evaluationspraxis als auch aus Sicht der Evaluationsforschung von großem Interesse (Konzendorf, 2013). Allerdings gibt es bislang praktisch keine empirischen Analysen zur Nutzung von Evaluationsergebnissen durch Politik und Verwaltung im Policy-Making-Prozess in Deutschland.

Handlungsbedarf besteht darüber hinaus bei der Weiterentwicklung von Evaluationsverfahren. Bislang gibt es keinen Gesamtüberblick über durchgeführte Evaluationen; so fehlt eine Datenbank, die Informationen z.B. zu methodischen Aspekten, zu Evaluationsergebnissen und zur Umsetzung von Vorschlägen enthält (Konzendorf, 2009, S. 29). Solche systematisch aufbereiteten Informationen wären jedoch hilfreich, um beispielsweise eine Erfolgs- und Problemanalyse durchführen zu können. Damit könnte ein Lernprozess sowohl auf Seiten der Auftraggeber als auch auf Seiten der Evaluatoren ermöglicht und insgesamt eine Verbesserung und Weiterentwicklung von Evaluationen erzielt werden.

Eine interessante Entwicklung, die seit einigen Jahren zu beobachten ist, zeigt sich zudem in der Diskussion um die Einführung und Anwendung der sogenannten Nachhaltigkeitsprüfung, wie sie z.B. bereits auf Bundesebene und im Land Baden-Württemberg als Erweiterung der Gesetzesfolgenabschätzung durchgeführt wird. Mit ihrer Hilfe sollen die langfristigen Wirkungen erfasst und dargestellt werden. Jedoch sind mit der Nachhaltigkeitsprüfung noch einige – v.a. methodische – Herausforderungen für die öffentliche Verwaltung verbunden, die diese Prüfung durchführt. Zu nennen sind hier insbesondere die Komplexität und die Mehrdimensionalität des Nachhaltigkeitskonzepts. Eine ernst gemeinte Nachhaltigkeitsprüfung bedarf eines komplexen Erhebungsinstrumentariums und Bewertungsansatzes. Gleichzeitig müssen die Prüfergebnisse den politischen Entscheidungsträgern kurz und verständlich zur Verfügung gestellt werden können, ohne dass sie zu inhaltsleeren Aussagen verkommen. Interessant wird dabei sein, wie es gelingen wird, die Nachhaltigkeitsprüfung tatsächlich mit den Zielen der jeweiligen Nachhaltigkeitsstrategien zu verknüpfen.

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 – weiter 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 nicht-intendierter Nebenwirkungen zu überprüfen.

Literatur

- Holtkamp, Lars (2012). *Verwaltungsreformen: problemorientierte Einführung in die Verwaltungswissenschaft*. Wiesbaden: VS Verlag.
- Jann, Werner (2001). Verwaltungsreform als Verwaltungspolitik: Verwaltungsmodernisierung und Policy-Forschung. In: Schröter, Eckhard (Hrsg.), *Empirische Policy- und Verwaltungsforschung* (S. 321–344). Opladen: Leske + Budrich.
- Janning, Frank, Schneider, Volker (2006). *Politikfeldanalyse*. Wiesbaden: VS Verlag.
- Konzendorf, Gottfried (2009). Institutionelle Einbettung der Evaluationsfunktion in Politik und Verwaltung in Deutschland. In: Widmer, Thomas, Beywl, Wolfgang, Fabian, Carlo (Hrsg.), *Evaluation: ein systematisches Handbuch* (S. 29–39). Wiesbaden: VS Verlag.
- Konzendorf, Gottfried (2013). Zum Einfluss von Evaluationen auf die politische Entscheidungsfindung. *Verwaltung & Management Heft 4/2013*, S. 171–178.
- Mayntz, Renate (Hrsg.) (1980). *Implementation politischer Programme – Empirische Forschungsergebnisse*. Königstein /Ts.: Athenaeum.
- Reiter, Renate, Ebinger, Falk, Grohs, Stephan, Kuhlmann, Sabine, Bogumil, Jörg (2011). Dezentralisierungsstrategien im Leistungsvergleich: Wirkungen von Dezentralisierungspolitik auf die Leistungsfähigkeit der Lokalsysteme in Deutschland, Frankreich und England. In: Europäisches Zentrum für Föderalismusforschung Tübingen (Hrsg.), *Jahrbuch des Föderalismus 2011* (S. 67–82). Baden-Baden: Nomos.
- Wegrich, Kai (2012). *Das Leitbild „Better Regulation“*. Berlin: Edition Sigma.
- Wollmann, Hellmut (2009). Kontrolle in Politik und Verwaltung: Evaluation, Controlling und Wissensnutzung. In: Schubert, Klaus, Bandelow, Nils C. (Hrsg.), *Lehrbuch der Politikfeldanalyse 2.0* (S. 379–400). München: Oldenbourg.
- Wollmann, Hellmut (2002). Verwaltungspolitik und Evaluierung. *Zeitschrift für Evaluation*, Heft 1/2002, S. 75–99.
- Ziekow, Jan, Debus, Alfred, Piesker, Axel (2013). *Die Planung und Durchführung von Gesetzesevaluationen*. Baden-Baden: Nomos.

Manfred Rolfes, Jan Lorenz Wilhelm

Evaluationspraxis und Evaluationsforschung im Kontext der Stadt- und Regionalentwicklung¹

1. Der Evaluationskontext: Stadt- und Regionalentwicklung

Zu Beginn des Beitrages wollen wir den inhaltlichen Kontext skizzieren, in dem hier über Evaluationspraxis und Evaluationsforschung nachgedacht werden soll. Die Bezeichnung Stadt- und Regionalentwicklung ist erläuterungsbedürftig, denn zum einen kursieren in der Alltagssprache recht oberflächliche und verkürzte Vorstellungen darüber, um was es hierbei geht. Wie weitreichend diese Politikfelder das Alltagsleben aber berühren, wird häufig nicht wahrgenommen. Zum anderen fasst auch die Fach- bzw. Wissenschaftssprache darunter sehr unterschiedliche Handlungsfelder, die darüber hinaus nationalstaatliche Spezifika aufweisen können. Auch in Beiträgen, die sich mit Evaluation oder Monitoring in der Stadt- und Regionalentwicklung befassen, fehlt in der Regel eine Präzisierung dieses Kontextes (vgl. Jacoby, 2009; Sedlacek, 2004).

Zum einen kann Stadt- und Regionalentwicklung als Bezeichnung für die *Beschreibung von beobachteten Phänomenen und Prozessen* verstanden werden, die sich in Städten und Regionen zeigen. Insbesondere soziale, ökonomische, infrastrukturelle, ökologische oder städtebauliche Entwicklungen sind dabei von Interesse. Zum anderen werden unter der Bezeichnung Stadt- und Regionalentwicklung *planerische, strategische und/oder politische Handlungsfelder* gefasst: Akteure in Institutionen, die auf unterschiedlichen administrativ-räumlichen Ebenen operieren, organisieren Prozesse und treffen Entscheidungen, mit denen die zukünftige Entwicklung in diesen Räumen (z.B. Stadtquartiere, Städte, Landkreise und Regionen) gesteuert oder zumindest beeinflusst werden soll. Diese Steuerungsprozesse betreffen sehr unterschiedliche Handlungsfelder. So könnten der Stadtentwicklung insbesondere die Handlungsfelder Städtebau, Stadtentwicklungsplanung bzw. Aufstellung und Durchführung städtebaulicher Entwicklungskonzepte, Stadt- und Kommunalplanung oder Stadtumbau zugeordnet werden. Noch vielfältiger erscheint das Feld der Regionalentwicklung. Hierzu ließen sich die Handlungsfelder Landes-, Raum- und Regionalplanung, Landschafts- und Umweltplanung, Raumentwicklung, Raumordnung(spolitik), Aufstellung und Durchführung regionaler Entwicklungskonzepte, regionale Strukturpolitik, Regionalmanage-

1 Wir danken unseren Reviewer/inne/n Michael Kalman und Ute M. Metje sowie darüber hinaus Richard Hummelbrunner, Oliver Schwab, Thomas Weith und Brigitte Wotha für die konstruktiven Feedbacks zu dem Beitrag.

ment oder -marketing, Verkehrsplanung sowie kommunale/regionale Wirtschaftsförderung rechnen (vgl. Akademie für Raumordnung und Landesplanung, 2005).

Sowohl bei der Stadtentwicklung als auch bei der Regionalentwicklung werden folglich neben fachplanerischen Aufgaben insbesondere querschnittsorientierte *politische Handlungsfelder* auf allen administrativ-räumlichen Maßstabsebenen bedient, und zwar von der Quartiersebene bis zur EU-Ebene. Und in dieser Eigenschaft sind sie für diesen Beitrag von Bedeutung. Es geht also um die politische Koordination und Planung von ökonomischen, soziodemographischen, ökologischen und infrastrukturellen Entwicklungen und Anforderungen, unter Berücksichtigung potenzieller oder beobachteter Wechselwirkungen zwischen diesen Feldern. Stadt- und Regionalentwicklung als politische Steuerungsaufgabe behandelt somit erstens die Erfordernisse der räumlichen Planung. Hierbei steht der Versuch einer koordinierten Arrondierung sozialer, ökonomischer sowie ökologischer Raumnutzungen und Raumentwicklungen im Vordergrund. Zweitens geht es, insbesondere auf der kommunalen Ebene, auch um die Sicherung der öffentlichen Daseinsvorsorge, also die Abfallbeseitigung, Versorgung mit Wasser, Gas und Strom sowie Angebote im öffentlichen Personennahverkehr. Eine weitere Systematisierung der Stadt- und Regionalentwicklung ist im Kontext von Evaluierungen von Bedeutung: die Unterscheidung in zeitlich begrenzte, ressortbezogene und/oder fördermittelfinanzierte Planungs- und Entwicklungsprogramme auf der einen Seite (z.B. EU-Förderprogramme oder die in Deutschland durchgeführten Bund-Länder-Programme wie Soziale Stadt oder Stadtumbau Ost/West oder die in der Schweiz im Rahmen der Neuen Regionalpolitik (NRP) zu implementierenden Programme und Projekte zur Stärkung der Innovationskraft, Wertschöpfung und Wettbewerbsfähigkeit der Regionen) und kontinuierliche, öffentliche Daueraufgaben der kommunalen und räumlichen Planung auf der anderen Seite (z.B. die Erstellung von Planwerken, Bauleitplanung, Regionalplanung, Maßnahmen der öffentlichen Daseinsvorsorge).

2. Entwicklung der Evaluationspraxis

Stadt- und Regionalentwicklung ist heute ohne Evaluation nicht mehr denkbar. Der Versuch, einen Überblick über die Entwicklung der **Evaluationspraxis** innerhalb der Stadt- und Regionalentwicklung zu geben, gestaltet sich jedoch nicht einfach. Er wird dadurch erschwert, dass in diesem Politikfeld eine Reihe von Aktivitäten und Instrumenten der Selbst- und Fremdbeobachtung existieren, die nicht mit dem Begriff Evaluation verknüpft werden, und zwar teilweise schlicht deswegen, weil der Evaluationsbegriff erst danach etabliert wurde (vgl. Hübler, 1984). Altrock weist beispielsweise (2007, 29f.) darauf hin, dass bereits ab Mitte der 1970er Jahre in den alten Bundesländern der BRD im Bereich der Stadtentwicklung wissenschaftliche Erkenntnisse zur Wirkung und den Effekten von Stadterneuerungs- und Stadtumbauprozessen generiert wurden. Ein anderes Beispiel liefert Jacoby (2009, 1): Seiner Auffassung nach hat sich für die Raumplanung die Laufende Raumbewertung² schon lange als wichti-

2 Die Laufende Raumbewertung ist in Deutschland wie in Österreich eine dauerhafte, Indikatoren gestützte Erfassung und Darstellung der räumlichen Entwicklung in Bereichen wie

ges Monitoring-Instrument zur Generierung raumrelevanter Informationen und zur Bewertung der Umweltauswirkungen von Plänen und Programmen bewährt. Diese Selbstbeobachtungen dokumentieren einen grundsätzlichen Willen und Wunsch der Akteure, Informationen über die Effizienz, die Wirkung oder den Erfolg ihrer Programme und Maßnahmen zu bekommen, ohne explizit von Evaluationen zu sprechen. Im Feld der Stadt- und Regionalentwicklung gibt es daher Versuche, die unter den Begriffen Controlling, Monitoring, Erfolgs- und Wirkungskontrolle und Evaluation subsummierten Aktivitäten zueinander in Beziehung zu setzen oder voneinander abzugrenzen (vgl. Diller, 2012, 2; Jacoby, 2009, 11 ff.). Es zeigt sich: Was dann unter den jeweiligen Begriffen konkret verstanden wird, hängt zum einen sehr stark von den spezifischen Organisations- und Politikkontexten ab, in denen diese evaluierenden Aktivitäten stattfinden. Und zum anderen wird die inhaltliche Ausrichtung dieser Aktivitäten erheblich davon beeinflusst, von wem welche Konsequenzen aus den erwarteten Ergebnissen gezogen werden. So kann leider eine begriffliche Diffusität nicht ganz ausgeräumt werden, auch wenn sich in der aktuellen Debatte die Bezeichnungen Evaluation oder Evaluierung als übergeordnete, zusammenfassende Begriffe weitgehend durchgesetzt haben. Dies berücksichtigend, sollen einige grundsätzliche Entwicklungslinien zur Evaluationspraxis in der Stadt- und Regionalentwicklung skizziert werden.

Evaluationen gewinnen an Bedeutung

Evaluationen haben in der Stadt- und Regionalentwicklung in den vergangenen drei Jahrzehnten erheblich an Bedeutung gewonnen. Nach unserer Ansicht lassen sich drei teilweise stark interagierende Prozesse bzw. Diskurse benennen, die diesen Bedeutungszuwachs kennzeichnen. *Erstens* ist die im Zuge des ökonomischen Strukturwandels der 1970/1980er Jahre einsetzende Reduzierung öffentlicher Mittel und die damit einhergehende Forderung nach einer nachweisbaren, effizienten Verwendung von Steuergeldern und Verringerung von Subventionen zu nennen (vgl. Becker, 2003, 209; Göddeke-Stellmann, 2007, 99; Schwab, 2009, 406). Die sachgerechte Nutzung der öffentlichen Gelder muss explizit legitimiert werden. *Zweitens* ist die Europäische Union sowohl durch ihre Rolle als Fördermittelgeber als auch durch die Etablierung von Prüflogiken innerhalb der Stadt- und Regionalentwicklung diskursbestimmender geworden. Auch der EU-Mittelfluss wird vor allem daran gebunden, inwieweit Zweckmäßigkeit und Wirksamkeit von Fördermaßnahmen nachgewiesen werden können. Dementsprechend hat eine vermehrte EU-Förderung ebenfalls dazu beigetragen, in der Stadt- und Regionalentwicklung eine Effizienz- und Bewertungslogik zu etablieren (vgl. Sedlacek, 2004, 21 f.; Thierstein, 2009, 425; Hummelbrunner/Maier, 2009, 415). *Drittens* wird seit Mitte der 1990er Jahre im deutschsprachigen Raum eine Privatisierung und Ökonomisierung innerhalb der öffentlichen Verwaltung diskutiert und beobachtet. Diese zeigt sich z.B. in Debatten über das New Public Management oder das Neue Steuerungsmodell, die vor allem an die Kommunalverwaltungen erhebliche Reform- und Modernisierungsanforderungen gestellt haben (vgl. Bogumil, 2007, 12 ff.;

Arbeitsmarkt, (Aus-)Bildung, Wirtschaft, Bevölkerung, Verkehr oder Umwelt. Der Bund und die meisten Bundesländern führen sie durch. Sie ist häufig Grundlage räumlicher Planung.