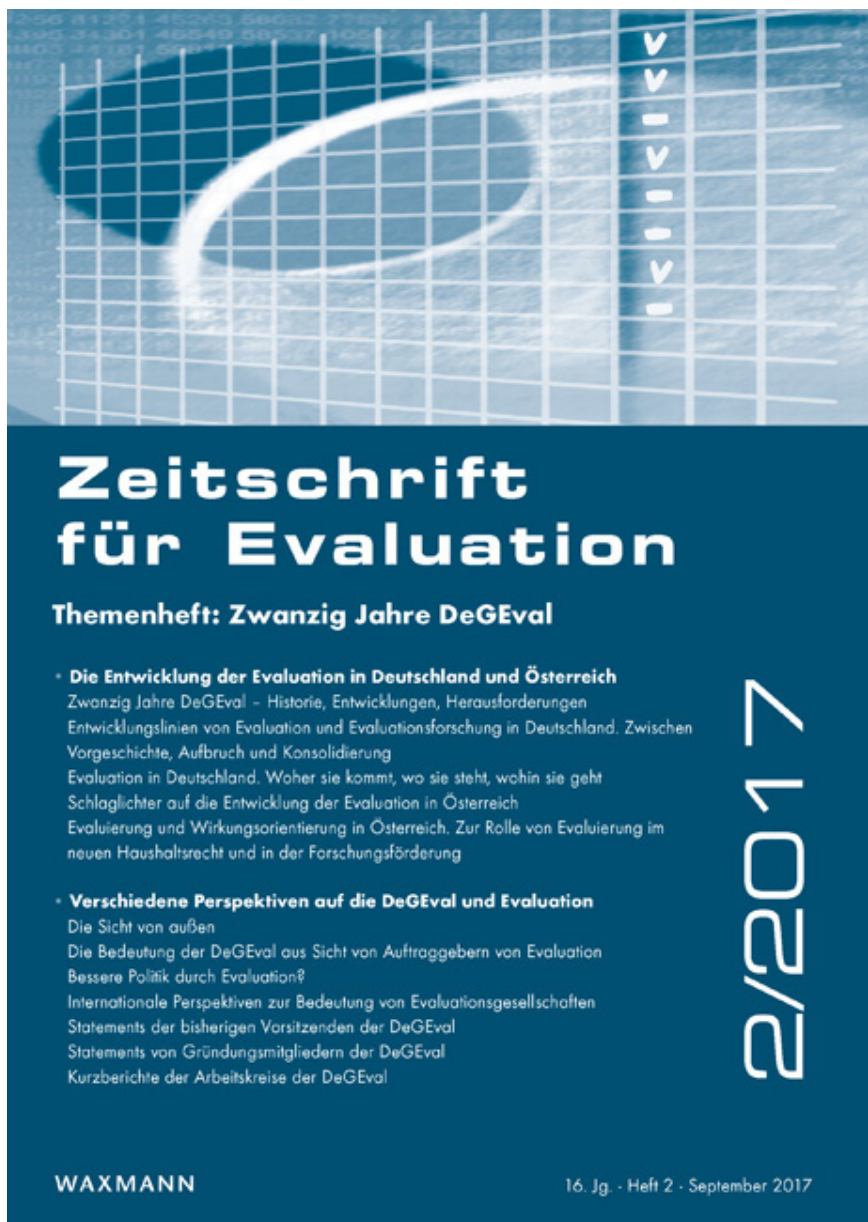


DIGITALER SONDERDRUCK

Zeitschrift für
Evaluation,
16. Jahrgang,
Heft 2/2017
ISSN 1619-5515

© Waxmann
Verlag GmbH,
2017



WAXMANN

Steinfurter Str. 555
48159 Münster

Fon 02 51 – 2 65 04-0
Fax 02 51 – 2 65 04-26

info@waxmann.com
order@waxmann.com

www.waxmann.com
Mehr zur ZfEval [hier](#).

Entwicklungslinien von Evaluation und Evaluationsforschung in Deutschland. Zwischen Vorgeschichte, Aufbruch und Konsolidierung

Hellmut Wollmann¹

Zusammenfassung: In dem Aufsatz werden in der Entwicklung von Evaluation und Evaluationsforschung in Deutschland eine Aufstiegs- und eine Konsolidierungsphase unterschieden. Die erste, die in den späten 1960er Jahren einsetzte, ist mit der Expansion von Reformpolitiken verknüpft, in der, unter dem konzeptionellen Einfluss der entsprechenden US-amerikanischen Entwicklung, Evaluation als Verfahren der Analyse und „Rückmeldung“ der Zielerreichung und Wirkungen politischer Programme und Maßnahmen, teilweise in der Absicht einer ‚verwissenschaftlichten‘ Politik („experimentelle Politik“), eingeführt wurde. Mit der Entstehung von Evaluationsforschung als (externer) Auftragsforschung wurden die traditionelle (von universitärer ‚Grundlagenforschung‘ geprägte) Wissenschaftslandschaft und die Beziehung zwischen Politik und der ‚scientific community‘ tiefgreifend verändert. In der ‚Konsolidierungsphase‘ seit den 1990er Jahren erwiesen sich die vom New Public Management inspirierte Politik- und Verwaltungsmodernisierung und die für EU-Programme vorgeschriebene systematische Evaluation als wirksame Impulse der weiteren Entwicklung. Die fortschreitende Institutionalisierung und Professionalisierung der Evaluationsforschung fanden in der Gründung der Gesellschaft für Evaluation (DeGEval) (1997) und der Gründung der Zeitschrift für Evaluation (2002) ihren markanten Ausdruck.

Schlagwörter: Reformpolitik, Wirkung politischer Programme, ‚Verwissenschaftlichte‘ Politik, Auftragsforschung

Development of Evaluation and Evaluation Research in Germany

Abstract: In discussing the development of evaluation and evaluation research in this article the distinction is made between an initial and a consolidation phase. The former setting off in the late 1960s is linked with the expansion of reform policies in which, under the conceptual influence of the related development in the USA, evaluation has been introduced as a procedure of analysis and “feedback” on goal attainment and on the effects of political programmes and measures, in part connected with the claim of a ‘scientification’ of policymaking (‘experimental policies’). Due to the emergence of evaluation research as (external) ‘contractual research’ the traditional research ‘landscape’ shaped by university-based ‘basic’ research, and the relation between the political world and the ‘scientific community’ have been profoundly changed. In the ‘consolidation phase’ the further development has been significantly impacted by New Public Management-inspired political and administrative modernisation and the systematic evaluation required by the European Union for its structural funds. The progressive institutionalisation and professionalisation of evaluation research is evidenced by the foundation of the (German) Evaluation Society (DeGEval) in 1997 and of the (German language) Journal of Evaluation (*Zeitschrift für Evaluation*) in 2002.

Keywords: Reform Policy, Effects of Political Programs, ‘Scientification’ of Policymaking, External Contractual Research

1 Em. Professor für Verwaltungswissenschaft, Humboldt-Universität zu Berlin

1. Vorbemerkung

Der Aufsatz soll einen Überblick über die Entwicklung der Evaluation und Evaluationsforschung in Deutschland geben. Ihm sollen einige Bemerkungen vorausgeschickt werden.

Zunächst einige Worte zur begrifflichen und terminologischen Verständigung. Ohne auf die in der einschlägigen Literatur zu dieser Frage geführten facettenreichen Diskussionen näher einzugehen (vgl. hierzu etwa Mertens 2006: 46-47; Wottawa/Thierau 2003; Stockmann 2006a: 15-25; Wollmann 2009, S. 382-385), sei – auf die Gefahr der Vereinfachung und Verkürzung – knapp das Folgende gesagt. Unter *Evaluation* (*Evaluierung*) werden die mit politisch-administrativen Entscheidungs- und Handlungsprozessen verbundenen Analyse- und Erkenntnisverfahren verstanden, die darauf gerichtet sind, die Zielerreichung, die Wirkungen, Kosten und Nutzen politischer Entscheidungen, Programme und Maßnahmen zu identifizieren und in die Entscheidungs- und Handlungsprozesse ‚zurückzumelden‘ (feedback). Im Unterschied hierzu werden als *Evaluationsforschung* Analyseverfahren und -tätigkeiten verstanden, in denen (sozial-)wissenschaftliche Methoden systematisch angewandt werden, um „valide“ Aussagen über die (auch ursächliche) Wirksamkeit politischen und administrativen Handelns zu gewinnen.² *Evaluationsforschung* wird typischerweise institutionell und personell von Wissenschaftlern und Forschungsinstituten – vielfach im Rahmen von *Auftragsforschung* – geleistet.

Auch wenn sich Evaluation/Evaluierung und Evaluationsforschung weitgehend parallel – gleichsam als „Siamesische Zwillinge“ – entwickelt haben, soll im Folgenden die Unterschiedlichkeit, wo dies der Zusammenhang nahe legt, begrifflich und sprachlich kenntlich gemacht werden.

Zwar haben sich Evaluation und Evaluationsforschung im heutigen Verständnis in der Bundesrepublik erst in den 1960er Jahren entfaltet, jedoch soll an dieser Stelle zunächst ein kurzer historischer, in das 19. Jahrhundert zurückreichender (in vorliegenden Darstellungen meist vernachlässigter) Rückblick gegeben werden, um auf frühe gewissermaßen ‚embryonale‘ Evaluationsansätze aufmerksam zu machen. Zugleich sollen hierdurch die tiefgreifenden Veränderungen deutlicher werden, die sich seit den 1960er Jahren im Verhältnis von Politik und Wissenschaft gerade auch durch das Aufkommen der Evaluationsforschung vollzogen haben.

In der Beschreibung der inzwischen über 50-jährigen Entwicklung der Evaluation und Evaluationsforschung sollen zwei Phasen unterschieden werden, zum ersten eine in den 1960er Jahren beginnende „Aufbruchsphase“ und zum andern eine „Konsolidierungsphase“, die, in den frühen und mittleren 1990er Jahren einsetzend, in der von New Public Management (NPM) inspirierten Politik- und Verwaltungsmodernisierung und der Evaluierungspflicht der EU-Förderprogramme wesentliche Triebfedern hatte. Beim Vergleich dieser Phasen sollen relevante Kontinuitäten oder Diskontinuitäten sichtbar gemacht werden.

2 Vgl. Peter Rossi/Howard Freeman (1993: 5): „Evaluation research is the systematic application of social research procedures for assessing the conceptualization, design, implementation, and utility of social intervention programs.“

Die nachstehende Diskussion beschränkt sich auf die Entwicklung in der Bundesrepublik und blendet jene im sonstigen deutschsprachigen Entstehungs- und Entfaltungsbereich, also in Österreich und in der (deutschsprachigen) Schweiz, aus.

Schließlich sei als Warnung eine persönliche Bemerkung vorausgeschickt. Da der Autor an der Entwicklung der Evaluationsforschung – insbesondere in deren *Aufbruchsphase* – engagiert mitwirkte, trägt die folgende Darstellung in möglicherweise überakzentuierter Weise „auto-biographische“ und auch „auto-bibliographische“ Züge.

2. Ein geschichtlicher Rückblick

In der deutschen Politik- und Verwaltungsgeschichte führt die Frage nach der Gewinnung und Anwendung politik- und verwaltungsnützlichen Wissens auf die *Polizeywissenschaften* zurück, die – im Kontext der spätabolutistischen Staaten des 18. und 19. Jahrhunderts – darauf gerichtet waren, das seinerzeit verfügbare politik- und verwaltungspraktische Wissen zu kumulieren und dem künftigen Verwaltungspersonal in universitärer Ausbildung und zur späteren Anwendung zu vermitteln (vgl. Maier 1980; Wollmann 1989: 235; Wittrock/Wagner/Wollmann 1991: 29). Mit Blick auf ihre praxeologische Wissenskumulierung, ihren Ausbildungs- und Anwendungsbezug sind diese „älteren Staatswissenschaften“ – in Anspielung auf das von *Harold Lasswell* in den 1950er Jahren geprägte Konzept der *policy sciences* (vgl. Lasswell 1951) – als „policy sciences des 18. und 19. Jahrhunderts“ bezeichnet worden (vgl. Beyme 1985: 7ff.; Wittrock/Wagner/Wollmann 1991: 29). Allerdings erlebten die *Polizeywissenschaften* im weiteren Verlauf des 19. Jahrhunderts ihren Niedergang, als unter einem „liberalen“ Staatsverständnis der alle staatlichen Funktionen umfassende *Polizei*-Begriff auf den engen, auf „Sicherheit und Ordnung“ gemünzten *Polizei*-Begriff schrumpfte. Mit dem gleichzeitigen Vordringen des Rechtspositivismus gewannen die rechtliche Regelung von Verwaltungshandeln und damit die juristische Ausbildung des Verwaltungspersonals die Oberhand – unter weitgehendem Verzicht auf empirisch fundierte und anwendungsbezogene „staatswissenschaftliche“ Kenntnisvermittlung.

Für die Generierung und Bereitstellung politik- und verwaltungsrelevanter Informationen setzten die deutschen Staaten im Verlauf des 19. Jahrhunderts stattdessen vermehrt auf die Einrichtung eigener statistischer Dienste (so 1805 in Preußen das *Preußische Statistische Büro*, vgl. Wollmann 2010: 258 und 1872 unmittelbar nach Reichsgründung das *Statistische Reichsamt*, vgl. Wollmann 1989: 236).

Hingegen drangen Ansätze einer anwendungsbezogenen empirischen Sozialforschung im letzten Drittel des 19. Jahrhunderts vor, als sich Vertreter der *historischen Schule der Nationalökonomie* unter Führung des an der Berliner Universität lehrenden *Gustav Schmoller* (vgl. Wollmann 2010: 258) – in reformpolitischer Absicht, die ihnen zeitgenössisch die spöttische Bezeichnung „Kathedersozialisten“ eintrug – der *sozialen Frage* zuwandten. Im Rahmen des 1872 gegründeten *Vereins für Socialpolitik* wurden empirische Untersuchungen (sog. *Enquêtes*), etwa zu

den (desolaten) Wohnungsverhältnissen der Arbeiterklasse, initiiert und durchgeführt (vgl. Wagner 1990: 80; Wagner/Wollmann 1991: 61; Wittrock/Wagner/Wollmann 1991: 34). Auf das Misslingen oder Fehlen staatlicher Politik gerichtet, hatten diese (auch im Ausland als beispielgebend anerkannten) frühen Ansätze empirischer Sozialforschung durchaus (modern gesprochen) „evaluative“ Züge.

Indessen wurde dieser Entwicklungsansatz einer (universitären) anwendungsbezogenen Politik- und Sozialforschung, wenn nicht sogar Wirkungsforschung, in der Folgezeit abgeschnitten. Zum einen zeigte die juristisch geprägte (und politisch konservative) staatliche Verwaltung an solcher Forschung kein Interesse. Zum andern kehrte die sich entwickelnde universitäre Sozialwissenschaft dem Praxisbezug des *Vereins für Socialpolitik* in dem Maße den Rücken, wie sie – nicht zuletzt unter dem Einfluss von *Max Weber* – ihre Anerkennung als *wissenschaftliche* Disziplin dadurch zu erreichen suchte, dass sie sich methodischer Stringenz und „werturteilsfreier“ Forschung verschrieb und in dieser Absicht Politik- und Praxisnähe vermied.

Zwar gab es in den 1920er Jahren seitens der Politik wieder erste Schritte, wissenschaftlichen Sachverstand für Politik und Verwaltung zu nutzen. So setzte der Reichstag 1926 einen *Ausschuss zur Untersuchung der Erzeugungs- und Absatzbedingungen der deutschen Wirtschaft* ein, in dessen Arbeit über 6.000 Wissenschaftler einbezogen wurden (vgl. Dierkes/Knie/Wagner 1988; Wollmann 1989: 238). Eine zukunftsweisende Idee des Ausschusses wurde von *Rudolf Hilferding*, der ein sozialdemokratisches Ausschussmitglied war, mit den Worten umschrieben: „Wir können uns der Einsicht nicht länger verschließen, dass Politik nur möglich ist auf wissenschaftlicher Grundlage... um endlich einmal aus dem politischen Dilettantismus herauszukommen“ (zit. nach Dierkes/Knie/Wagner 1988: 19). Jedoch wurden die Arbeiten des Ausschusses 1931 aus finanziellen Gründen eingestellt. Die kühne Vorstellung von einer „wissenschaftlich basierten Politik“ und damit einer neuen Rolle der Wissenschaft ging in den politischen und ökonomischen Wirren der späten Jahre der Weimarer Republik unter.

Insgesamt blieben diese Gegenüberstellung und wechselseitige Distanzierung einer Politik- und Verwaltungswelt, die sich in ihrem Politik- und Verwaltungshandeln wesentlich auf den Sachverstand ihrer (juristischen) Verwaltungseliten stützte, und einer universitären Wissenschaftswelt, die im Streben nach Wissenschaftlichkeit den Praxisbezug vermied, im beider- und wechselseitigen (Selbst-)Verständnis und Verhalten bis 1933 und auch nach 1945 in den ersten Jahren weithin bestimmend.

In eklatantem Gegensatz hierzu verlief die Entwicklung in den USA, wo während der 1920er Jahre an der *University of Chicago* unter dem prägenden Einfluss von *Charles Merriam* das Modell einer universitären sozialwissenschaftlichen Forschung entwickelt und praktiziert wurde. Hier wurde der Anspruch eines hohen methodischen und theoretischen Standards ausdrücklich und integral mit ihrem Praxis- und Anwendungsbezug verknüpft (vgl. Wittrock/Wagner/Wollmann 1991: 39; Wagner/Wollmann 1986b: 602). In den 1930er Jahren nutzte und verstärkte die Roosevelt Administration diese praxisbezogene Bereitschaft und Kompetenz universitärer Sozialforschung, als sie erstmals die Evaluation ihrer *New Deal Policy* beschloss und Sozialwissenschaftler in bislang unbekanntem Umfang in diese einbezog. Auch während des 2. Weltkriegs war die US-Regierung bestrebt, sozialwis-

senschaftlichen Sachverstand als für die Kriegsführung relevante Informationsressource zu mobilisieren. Vor diesem Hintergrund formulierte *Harold Lasswell* nach 1945 sein Konzept von *policy sciences* (vgl. Lasswell 1951), das in der Vorstellung und Forderung wurzelte, das Potenzial und den Fundus wissenschaftlichen Wissens in einem breiten interdisziplinären Umgriff für die Entscheidungsfindung politischer und gesellschaftlicher Probleme zu nutzen (vgl. Wittrock/Wagner/Wollmann 1991: 39).

3. Die Aufbruchphase der Evaluationsforschung in Deutschland (in den 1960er bis 1980er Jahren)

Nach 1945 knüpfte die Neubegründung bzw. der Wiederaufbau der politisch-administrativen und wissenschaftlichen Strukturen in der Bundesrepublik weitgehend an die bis 1933 kennzeichnende wechselseitige Distanz von Politik und Wissenschaft an. Aufseiten der Wissenschaft waren insbesondere die universitären Sozial- und Politikwissenschaften wesentlich von ihrer institutionellen und personellen Wieder- bzw. Neubegründung absorbiert (vgl. Hohn/Schimank 1990: 303ff.), die Politikwissenschaft zudem von dem normativen und eher „praxisabstinenten“ Selbstverständnis einer „Demokratiewissenschaft“. Auf der anderen Seite setzten Regierung und Verwaltung zur Informations- und Erkenntnisgewinnung nach wie vor überwiegend auf den Sachverstand der eigenen Verwaltungseliten und statistischen Dienste (vgl. Wollmann 1989: 239f.). Im Verlauf der 1950er Jahre suchten die Bundesministerien immerhin die Beratung durch (regierungsnahe) „wissenschaftliche Beiräte“. Nach dem Vorbild des US-amerikanischen *Council of Economic Advisers* wurde schließlich 1963 der (unabhängige) *Sachverständigenrat zur Beurteilung der gesamtwirtschaftlichen Entwicklung* eingerichtet.

Im Verlauf der 1960er Jahre wurde die politische Diskussion zunehmend von der Überzeugung bestimmt, dass die unmittelbare Nachkriegs- und Wiederaufbauphase zu Ende sei und zur Sicherung der internationalen Wettbewerbsfähigkeit durchgreifende Reformen geboten seien (hierzu und zum Folgenden vgl. Hellstern/Wollmann 1984b: 33ff.; Stockmann 2006a: 29ff.). Der von *Georg Picht* 1964 ausgelöste Alarmruf einer „Bildungskatastrophe“ wurde über den Bildungssektor hinaus als Menetekel wahrgenommen. Die 1969 gebildete sozial-liberale Koalition unter Kanzler Willy Brandt verkündete eine breit angelegte „Politik der inneren Reformen“, die eine umfassende Modernisierung der Politik- und Verwaltungsstrukturen und hierin die Einführung und Nutzung neuer Planungs-, Informations- und auch Evaluierungsverfahren einschloss.

Hierbei kamen wesentliche politikkonzeptionelle und -instrumentelle Impulse aus den USA, wo in den mittleren 1960er Jahren – unter den Präsidenten *Kennedy* und *Johnson* – weitreichende wohlfahrtsstaatliche Reformprogramme, insbesondere der *War on Poverty* in Gang gesetzt wurden. Zugleich wurde in den USA eine Modernisierung des politisch-administrativen Handlungssystems eingeleitet, zu der die Einführung von Informations- und Planungssystemen (*Programming, Planning,*

Budgeting System, PPBS) in den Ministerien ebenso gehörte wie die Evaluation der Reformprogramme. Dem lag die Leitvorstellung einer „rationalen“ Politik und eines Politikzyklus zugrunde, der über die Triade und Sequenz von Planung/Implementation/Wirkung abläuft und in dem Evaluation als Analyseverfahren und kybernetische Rückmeldeschleife eine strategische Schlüsselfunktion einnimmt. Die Vorstellung einer „verwissenschaftlichten“ Politik fand in *Donald Campbells* Ruf nach *reforms as experiments* (vgl. Campbell 1969) ihren beredten Ausdruck und in einer Abfolge von *social experiments* ihren wirkungsmächtigen Niederschlag (ausführlich vgl. Hellstern/Wollmann 1983b: 25ff.)

Der in den späten 1960er Jahren in der Bundesrepublik einsetzende Aufbruch der Politikevaluation und der Evaluationsforschung war von diesem Vorlauf und Beispiel der US-amerikanischen Entwicklung konzeptionell und institutionell, wie im Folgenden verdeutlicht werden soll, vielfältig beeinflusst.

3.1 Dramatische Neubestimmung des Verhältnisses von Politik/Verwaltung und (Sozial-)Wissenschaft

War in der deutschen Staats- und Wissenschaftstradition, wie oben skizziert, das Verhältnis zwischen der Politik- und Verwaltungswelt einerseits und der Wissenschaftswelt andererseits weitgehend durch wechselseitige Distanz gekennzeichnet, so wurde dieser Entwicklungsstrang im Laufe der 1960er Jahre dadurch einschneidend verändert, dass Regierung, Politik und Verwaltung die Beratung und Informationsgewinnung durch wissenschaftliche Forschung suchten, während umgekehrt sich auch die (Sozial- bzw. Politik-)Wissenschaft bereit und interessiert zeigte, entsprechende Beratungsforschung zu leisten.

Ein eindrucksvolles Beispiel für diesen Wandel und Bruch gab die *Projektgruppe Regierungs- und Verwaltungsreform* (PRV), die 1968 von der Bundesregierung gebildet wurde, in die eine Reihe von Sozialwissenschaftlern (darunter *Fritz W. Scharpf* und *Renate Mayntz*) berufen wurde und die eine in der deutschen Politik bislang kaum geläufige „Reformkoalition“ (*reform coalition*, siehe Wollmann 1984; Wittrock/Wagner/Wollmann 1991: 43) aus Politikern und Sozialwissenschaftlern bildete. Zwar blieben die Wirkungen der von der PRV erarbeiteten Gutachten und Empfehlungen zunächst gering. Jedoch sind längerfristige Wirkungen der PRV in den weiteren Forschungen und Publikationen der an ihr mitwirkenden Wissenschaftler zu erkennen.

Dies lässt sich beispielhaft an *Hans-Ulrich Derlien* zeigen, der vor dem Hintergrund seiner wissenschaftlichen Mitarbeit an der PRV seine Dissertation mit dem Titel „Erfolgskontrolle staatlicher Planung. Eine empirische Untersuchung über Organisation, Methode und Politik der Programmevaluation“ schrieb (vgl. Derlien 1976). Er spielte in der Folgezeit in der konzeptionellen Entwicklung und Verbreitung der (von ihm als „Programmforschung“ definierten) Evaluationsforschung eine wichtige Rolle (vgl. Derlien 1990a, 1990b; Derlien/Rist 2002) und war zudem 1981 maßgeblich an der Gründung der *Gesellschaft für Programmforschung* beteiligt (vgl. Derlien 1981).

3.2 Tiefgreifende Veränderung der Forschungslandschaft mit Vordringen der Evaluationsforschung

Da Politik und Verwaltung auf Bundes-, Landes- und Kommunalebene nur in beschränktem Umfang personell und qualifikatorisch im Stande waren, das neue Interesse an Evaluation durch eigene Untersuchungen (*in house*) einzulösen, gingen sie seit den späten 1960er Jahren dazu über, hierfür externe Forschungskapazitäten im Wege der „Auftragsforschung“ (vgl. Wollmann 2002b) zu gewinnen und entsprechende Fördermittel (im Bund sog. Ressortforschungsmittel) einzusetzen. Hierdurch wurde ein neuer „Forschungsfördermittelmittelmarkt“ geschaffen, der die Forschungslandschaft institutionell tiefgreifend veränderte. Wurde sozial- und politikwissenschaftliche Forschung bislang – dem Humboldt’schen Imperativ von der *Einheit von Lehre und Forschung* folgend – überwiegend an den Hochschulen (oder auch als angewandte Politik- und Sozialforschung an quasi-öffentlichen Forschungseinrichtungen wie dem 1969 gegründeten *Wissenschaftszentrum Berlin* und etwas später an dem 1984 in Köln gegründeten *Max-Planck-Institut für Gesellschaftsforschung*) betrieben, so löste der expandierende Auftragsforschungsmarkt die Gründung und Verbreitung von nicht universitären, überwiegend privatwirtschaftlich organisierten (vielfach Consulting- und Marktforschungs-)Unternehmen aus (vgl. Hellstern/Wollmann 1984b: 70ff.; Stockmann 2006a: 37).

3.3 Beispielfelder für den Aufstieg der Evaluationsforschung

Im Folgenden seien die Anfänge der Evaluationsforschung am Beispiel einiger Politikfelder gekennzeichnet.

Bildungspolitik: In der Bundesrepublik erhielt die Evaluationsforschung bereits in den mittleren 1960er Jahren einen kräftigen Impuls in der Bildungspolitik, als die Reform des Schulwesens in Bundes- und Landespolitik politische Dringlichkeit erhielt. Durch Vereinbarung zwischen Bund und Ländern wurde 1965 ein *Bildungsrat* eingerichtet, dessen mit zahlreichen Erziehungswissenschaftlern besetzter *Unterausschuss ‚Experimentalprogramme‘* zum Motor einer vom einschlägigen US-amerikanischen Vorlauf und Vorbild inspirierten *experimentellen* Politik in der Schulpolitik wurde. Die Bildungskommission verabschiedete im Juli 1968 Empfehlungen zur „Einrichtung von Schulversuchen mit Ganztagschulen“, in deren Rahmen zwischen Ende der 1960er bis Anfang der 1980er Jahre mit großem wissenschaftlichen Evaluationsaufwand in mehreren beteiligten Ländern, jeweils hälftig aus Mitteln des Bundes und des jeweiligen Sitzlandes finanziert, insgesamt 78 Einzelstudien durchgeführt und evaluiert wurden (vgl. Hellstern/Wollmann 1983b: 13ff.; Stockmann 2006a: 31). Zwar erwiesen sich die Evaluierungsstudien aufgrund unterschiedlicher methodischer Ansätze, Vorgehensweise, aber auch wegen der unterschiedlichen Interessen der beteiligten Länder als heterogen, so dass ein abschließendes und allseits geteiltes Urteil über die „getestete“ neue Schulform nicht zustande kam. Dennoch hatte „diese bis heute größte empirische Schulforschungswelle in Deutschland“

(Büeler 2006: 265) konzeptionell und methodisch eine „weit über die Bildungsforschung hinausgehende Wirkung“ (Stockmann 2006a: 31).

Weitere Felder experimenteller Politik: Nachdem der Ansatz *experimenteller Politik* in der deutschen Bildungspolitik bereits in den mittleren 1960er Jahren einen bemerkenswerten Einstieg gefunden hatte, wurde er ab den frühen 1970er Jahren in weiteren teilweise großangelegten, methodisch und empirisch aufwändigen ‚experimentellen‘ Vorhaben (z.B. Einphasen-Juristenausbildung, „Großversuch Tempo 100“) ins Werk gesetzt (vgl. Hellstern/Wollmann 1983a, 1983b für eine breite Übersicht und Diskussion). Darüber hinaus hielt *experimentelles* Vorgehen Einzug in Gesetzgebungsvorhaben auf Bundes- und auch auf Landesebene, wo gesetzliche Vorschriften befristet in Kraft gesetzt (*sunset legislation*) und evaluiert wurden (vgl. die Beiträge in Schreckenberger et al. 1986, darin Wollmann 1986 mit Beispielen).

Carl Böhret und Mitarbeiter verfolgten seit den frühen 1980er Jahren einen *experimentellen* Ansatz weiter, indem sie einzelne Gesetzgebungsvorhaben einem (realen) *Praxistest* unterzogen (vgl. Böhret/Hugger 1980) und im weiteren die Konzepte einer (ex-ante – und ex-post) *Gesetzesfolgenabschätzung* (vgl. Konzendorf 2011) entwickelten.

Entwicklungspolitik: In der frühzeitigen Begründung und Institutionalisierung von Evaluation nahm das *Bundesministerium für Wirtschaftliche Zusammenarbeit (BMZ)* eine Ausnahmestellung ein, insofern es bereits in den frühen 1970er Jahren begann, ein ausgefeiltes Evaluationssystem im Feld der Entwicklungspolitik zu entwickeln und bereits 1971 zunächst ein *Inspektionsreferat*, später *Zentrale Erfolgskontrolle* genannte Evaluationseinheit einrichtete (vgl. Lorz 1984: 293). Ihm war von Anfang an die Aufgabe gestellt, „die Wirksamkeit der deutschen Entwicklungszusammenarbeit zu überprüfen, indem die durchgeführten Projekte und Programme und die dafür eingesetzten Instrumente evaluiert werden“ (Stockmann 2006b: 379 mit weiteren Nachweisen).

Städtebaupolitik: Schließlich sei als instruktives Beispiel das *Bundesministerium für Bauwesen und Städtebau (BMBau)* hervorgehoben, das seit den frühen 1970er Jahren im Zuge der Expansion seiner städtebau- und wohnungsbaupolitischen Handlungsprogramme und Gesetze hierauf gerichtete Evaluationsvorhaben initiierte. So wurde eine methodische Vorstudie für die Evaluation städtebaulicher Maßnahmen beauftragt, einen Überblick über den Stand der einschlägigen internationalen Diskussion und Erfahrung zu erarbeiten (vgl. Hellstern/Wollmann 1978; Wollmann/Hellstern 1977). In Anknüpfung an die Ansätze ‚experimenteller Politik‘ wurde ein *Programm Experimenteller Wohnungs- und Städtebau* mit zahlreichen Studien- und Modellvorhaben aufgelegt, die Gegenstand von evaluativer *Begleitforschung* wurden (vgl. Wollmann 1990). Ferner vergab das *BMBau* zur Vorbereitung und Überprüfung seiner Gesetzgebungstätigkeit Untersuchungen, die konzeptionell im Schnittfeld von Ex-post-Evaluation, Implementations- und Rechtstatsachenforschung lagen (vgl. Wollmann 1982; vgl. Scharmer/Wollmann/Argast 1985³ als Bei-

3 Es sei persönlich angemerkt, dass die genannten (und auch andere) Evaluationsprojekte des *BM-Bau* seinerzeit am *IfS-Institut für Stadtforschung und Strukturpolitik* bearbeitet wurden, dessen Mitbegründer ich 1974 war (und dem ich als Mitgesellschafter nach wie vor verbunden bin, vgl. hierzu auch Wollmann 2017).

spiel einer konzeptionell und empirisch aufwändigen „Rechtstatsachenuntersuchung zur Baugenehmigungspraxis“. Zur konzeptionellen Überschneidung von Gesetzes-evaluation und Rechtstatsachenforschung vgl. auch Konzendorf 2009: 135).

3.4 Evaluationsforschung als „Auftragsforschung“ zwischen wissenschaftlicher Autonomie und Fremdbestimmung

Seit ihrer Entstehung und Entfaltung als eine Variante angewandter und praxisbezogener Politik- und Sozialforschung und zumal als Auftragsforschung (*contractual research*) wurde als ein forschungsethisches und wissenschaftlich-methodisches Schlüsselproblem der Evaluationsforschung die Frage ihrer intellektuellen und institutionellem Unabhängigkeit und wissenschaftlichen Objektivität diskutiert (vgl. Wollmann 2002b). Beide Ansprüche sind in mehrfacher Hinsicht herausgefordert und gefährdet. Während sich die (akademische) Grundlagenforschung in der Wahl ihrer Untersuchungsfelder, -fragen und -methoden (jedenfalls *idealiter*) von dem Grundverständnis einer Autonomie und Eigenbestimmung wissenschaftlicher Forschung und vom Imperativ einer (interessenfreien) Wahrheitssuche leiten lässt, erscheint Evaluationsforschung als *Auftragsforschung* wesentlich durch das Grundverhältnis von Auftraggebern und Auftragnehmern geprägt, in dem erstere vielfach die maßgeblichen Eckdaten (*terms of reference*) des zu vergebenden Forschungsprojektes (Gegenstandsbereich, Fragestellung, möglicherweise Methoden, zeitlicher Rahmen usw.) vorgeben, während sich die letzteren oft in einer (finanziellen) Abhängigkeit befinden, von der der Volksmund mit der Redensart „Wes Brot ich ess‘, des Lied ich sing“ spricht. Hierdurch kann für Evaluationsforschung als Auftragsforschung ein bis in ihre Erkenntnisprämissen reichender „Sog“ (*epistemic drift*, Elzinga 1985) und ihre „Kolonialisierung“ durch Politik- und Praxisinteressen erzeugt werden (vgl. Wittrock/Wagner/Wollmann 1991: 57ff.; Wollmann 2002b).

Auch die evaluative *Begleitforschung* – z.B. in den experimentellen Schulversuchen und im Experimentellen Städtebau – gab frühzeitig Veranlassung, ihre forschungsethischen und -methodischen Probleme zu diskutieren (vgl. Lutz 1983; Wollmann 1990). Hierbei wurden eine (der „klassischen“ Ex-post-Evaluation vergleichbare) distanziert-analytische und eine ‚intervenierende‘ (Lutz 1983) Variante unterschieden, in der die aktive Mitwirkung des Evaluators an dem zu evaluierenden Projekt (und an dessen Gelingen) den Vorrang hat und die wissenschaftliche Stringenz und Unabhängigkeit der Forschung bewusst nachgeordnet wird. Hiermit weist sie deutliche Überschneidungen mit der sozialwissenschaftlichen *Aktionsforschung* auf (vgl. Kritz 2004).

3.5 Zwischenbilanz der Aufbruchphase

In der Entwicklung der Evaluationsforschung seit den 1960er Jahren können – vereinfacht gesagt – zwei Phasen oder „Wellen“ unterschieden werden. In ihrer *kick-off*-Phase bis in die mittleren 1970er Jahre hatte Politikevaluation in Widerspiegelung ihres reformpolitischen und „planungseuphorischen“ Entstehungskontextes vor allem die Verbesserung der Wirksamkeit („*outcome*“) staatlicher Politik und Programme im Blick. Ab den mittleren 1970er Jahren – in Reaktion auf die im Gefolge des (ersten) Erdölpreisschocks nach 1973 einsetzende ökonomische und budgetäre Krise – wurde der (kurzlebige) Reformoptimismus von einer „planungsernüchterten“ und zunehmend neo-liberal inspirierten Politik der Haushaltskonsolidierung abgelöst. Der Evaluation wurde nunmehr vorrangig die Aufgabe zugewiesen, Kosten (*input*) zu reduzieren und zu minimieren (vgl. Hellstern/Wollmann 1984b: 48ff.; zuletzt Filsinger 2016: 2ff.).

In der Aufbruchphase der deutschen Evaluationsdiskussion übte die Rezeption der einschlägigen US-amerikanischen Literatur und Erfahrung einen prägenden Einfluss aus (vgl. Weiss 1974 und Rossi/Freeman/Hofmann 1988). Ferner sind ein (auf eine von Hellstern und Wollmann an der FU Berlin organisierte internationale Tagung zurückgehender englischsprachiger) Sammelband von Levine/Solomon/Hellstern/Wollmann 1981 (mit Aufsätzen u.a. von Rossi, Freeman/Solomon, Rein und Wholey), der Sammelband von Wulf 1972 sowie Aufsätze von Kromrey (1995, 2003) zu nennen. Als *Textbooks* zur Evaluationsforschung ist in dieser Phase auf Wittmann 1985 und Wottawa/Thierau 1990 (2003) zu verweisen.

Die von Hans-Ulrich Derlien 1981 gegründete *Gesellschaft für Programmforschung* (vgl. Derlien 1981) erwies sich als wichtiges Diskussionsforum, dessen Jahrestagungen mit Evaluation befasste Wissenschaftler und Verwaltungspraktiker regelmäßig zusammenführten und die in jährlichen Veröffentlichungen („Werkstattberichten“) mündeten.

Der Diskussionstand der 1980er Jahre wurde u.a. in Hellstern/Wollmann 1984a (zur Bundesebene), Hellstern/Wollmann 1984c, 1984d (zur kommunalen Ebene) und Hellstern/Wollmann 1983a, 1983b (zur „experimentellen Politik“) dokumentiert.

In international vergleichenden Darstellungen wurde die Bundesrepublik – neben Schweden und Großbritannien – zu den europäischen Vorreitern der Evaluation gerechnet (zum Stand in den 1980er Jahren vgl. Levine 1981; Wagner/Wollmann 1986a; Stockmann 2000 sowie zum Stand in den 1990er Jahren Furubo/Rist/Sandahl 2002; Furubo/Sandahl 2002; Derlien/Rist 2002; Wollmann 2003a, 2003c, 2007).

4. Konsolidierungsphase der Evaluationsforschung seit den 1990er Jahren

Seit den 1990er Jahren erhielt die weitere thematische, konzeptionelle und institutionelle Entfaltung und Konsolidierung der Evaluationsforschung durch mehrere Faktoren prägende Anstöße.

4.1 Anstöße und Beispielsfelder

New Public Management (NPM)

Das Konzept des *New Public Management* (NPM), das seit den 1980er Jahren die internationale Diskussion einer umfassenden Verwaltungsmodernisierung dominiert hat (vgl. Pollitt/Bouckaert 2011, Kuhlmann/Wollmann 2013) und seit den frühen 1990ern auch in Deutschland, zunächst vor allem auf der kommunalen Ebene, Eingang fand (vgl. Reichard 1994), löste für die Evaluierung politischen und administrativen Handelns einen weiteren Schub aus. In diesem Zusammenhang spielte der *Evaluationszyklus* – mit seiner Verknüpfung von Analyse und (kybernetischer) Rückmeldung – eine strategisch bedeutende Rolle. Dieser dient nicht nur als Verfahren zur laufenden analytischen Erfassung des *externen* Leistungsverhaltens der Organisation, sondern auch und gerade zur ständigen Beobachtung (*monitoring*) und Rückmeldung der (verwaltungs-)internen Handlungsabläufe. Hierbei finden die dem NPM eigentümlichen *kybernetischen* und auf *Indikatoren* gestützten Evaluationsverfahren nicht nur im Verwaltungssektor Anwendung, sondern wirken mit ihrem umfassenden Evaluierungsanspruch auch auf andere Politikfelder und deren Evaluation ein. Angesichts der von diesem konzeptionellen Anstoß ausgelösten Veränderungen ist von einer „dritten“ Welle der Evaluierung gesprochen worden (vgl. Wollmann 2003b). Ähnlich wurde NPM im internationalen Überblick als „the most important driving force“ (Meyer/Stockmann 2016: 346) für die aktuelle und weitere Expansion der Evaluationsforschung bezeichnet.

Hinsichtlich der Evaluation der Verwaltungsmodernisierung setzte in den frühen 1990er Jahren zunächst eine Diskussion über hierfür geeignete Konzepte und Methoden ein (vgl. Wollmann 2002a, 2006; S. Kuhlmann 2009), wobei auch die für die Evaluation von Reformpolitiken relevanten *Transaktionskosten* hervorgehoben wurden (vgl. Kuhlmann/Wollmann 2006). Für erste einschlägige empirische Evaluationsuntersuchungen sei auf Jaedicke/Thrun/Wollmann (2000), Bogumil/Grohs/Kuhlmann/Ohm (2007) sowie S. Kuhlmann et al. 2008 verwiesen.

Evaluation der EU-Förderprogramme

Kräftige Impulse für die weitere Entwicklung der Evaluationsforschung wurden seit den frühen 1990er Jahren in Deutschland (und EU-weit) ausgelöst, als die EU mit der Reform der Strukturfonds von 1988 den Programmansatz für ihre gesamte Strukturfondsförderung einführte und zugleich die Mitgliedsstaaten zur Evaluierung

EU-kofinanzierter Programme und Maßnahmen verpflichtete (vgl. Grajewski/Meyer 2014: 38). Beginnend mit der ersten fünfjährigen Förderperiode (1994-1999) wurde die periodische und systematische Evaluation der Strukturfondsprogramme – in der (geradezu lehrbuchartigen) Sequenz von *Ex-ante*-, *On-going*- (*mid term*) und *Ex-post*-Evaluation – vorgeschrieben (vgl. Leeuw 2006: 75ff.; Schwab 2009: 409ff.). An dem von der EU induzierten Evaluationssystem, innerhalb dessen ein kaum noch überschaubares internationales und nationales Netzwerk von Evaluationsinstituten und Evaluatoren entstanden ist, wirkt auch die deutsche *evaluation community* (insbesondere in Gestalt von „freien, meist privaten oder als sog. *An-Institute* organisierten Forschungseinrichtungen“ (Schwab 2009: 410) in erheblichem Umfang mit, sei es im Rahmen von Evaluationsvorhaben der Bundesländer oder im Rahmen von Forschungsaufträgen der EU-Kommission in der hierfür eigentümlichen Form internationaler Konsortien.

Evaluierung von Forschungseinrichtungen

Von den umfangreichen Evaluationen angestoßen, die der Wissenschaftsrat nach 1990 zur Restrukturierung („Abwicklung“) der Wissenschaftswelt der früheren DDR veranlasste, wurde in der Folgezeit auch in der „alten“ Bundesrepublik ein breiter Kranz von Forschungseinrichtungen evaluiert (vgl. St. Kuhlmann 2006, 2009 jeweils mit Nachweisen). So wurden in den späten 1990er Jahren sog. Systemevaluationen der *Institute der Max-Planck-Gesellschaft*, der *Fraunhofer-Gesellschaft*, der *Wissenschaftsgemeinschaft G.W. Leibnitz* (früher *Blaue Liste*-Institute) sowie der Großforschungseinrichtungen der *Helmholtz-Gemeinschaft* (mit jeweils Tausenden von Wissenschaftlern und Milliarden Euro-schweren Etats) durchgeführt.

Schulsektor: War der Schulsektor, wie weiter vorne skizziert, seit den mittleren 1960er Jahren im Rahmen der experimentellen Schulversuche ein Vorreiter in der Evaluationsentwicklung in Deutschland, so erlebte er in den frühen 2000er Jahren dadurch einen erneuten Evaluierungsschub, dass die Bundesländer an den von der OECD angestoßenen internationalen Schulvergleichsstudien (PISA) teilnahmen (vgl. Büeler 2006; Maag Merki 2009: 159). (Zu diesen und anderen Ansätzen der schulischen Evaluationen vgl. Maag Merki 2009: 161; Fraix/Renz 2014.)

Arbeitsmarktpolitik

Als „das bislang umfassende Projekt der Arbeitsmarktevaluierung in Deutschland“ (Wagner 2009: 122) sei hier die sog. *Hartz-IV*-Evaluation hervorgehoben, die im Zusammenhang mit den umfangreichen Arbeitsmarkt- und Sozialgesetzreformen der frühen 2000er Jahre (*Hartz-IV-Reformen*) durchgeführt wurde – von 100 Wissenschaftlern in knapp 20 wissenschaftlichen Einrichtungen mit einem Fördervolumen von 10,3 Mio. Euro (vgl. Heyer 2006: 468; Bangel/Brinkmann/Deeke 2006).

Entwicklungspolitik

Nachdem die vom Bundesministerium für Wirtschaftliche Zusammenarbeit (BMZ) seit den frühen 1970er Jahren durchgeführten oder beauftragten Evaluationen im Wesentlichen auf Einzelprojekte gerichtet waren, zielt seine Evaluationsstrategie

seit einigen Jahren darauf, das gewonnene Evaluationswissen – im Wege von Querschnittsevaluationen (sog. Meta-Evaluationen) – zu kumulieren und hieraus allgemeine Empfehlungen und Kriterien abzuleiten, die als Entscheidungshilfen für die Auswahl, Planung und Durchführung ähnlicher Projekte in der Zukunft dienen (vgl. Stockmann 2006a: 380). Mit der Gründung des eigenständigen *Deutschen Instituts für Entwicklungsevaluierung* im Jahr 2012 soll die Unabhängigkeit der Evaluation in diesem Politikfeld gesichert werden (vgl. Neubert/Mack/Roxin 2014: 67ff.).

Gesetzesfolgenabschätzung

Auf in den 1980er Jahren initiierte Arbeiten zurückgehend (vgl. Böhret/Hugger 1980), wurde 2004 die „Gesetzesfolgenabschätzung“ in §44 der Gemeinsamen Geschäftsordnung der Bundesministerien (GGO) verankert. Hierzu zählen neben den beabsichtigten Wirkungen und Nebenwirkungen auch die Kosten für die öffentlichen Haushalte, Kosten für die Wirtschaft und Auswirkungen auf Verbraucher (vgl. Konzendorf 2011: 137). Zur (Ex-ante-)Abschätzung der *Bürokratiekosten* wurde 2006 ein *Nationaler Normenkontrollrat* eingerichtet (vgl. Nationaler Normenkontrollrat 2016).

4.2 Institutionelle Konsolidierung der Evaluationsforschung

Im Verlauf der 1990er Jahre durchmaß die Evaluationsforschung mehrere wichtige Schritte der institutionellen und professionellen Konsolidierung.

Nachdem – dem Vorbild der 1984 gegründeten *American Evaluation Society* folgend – im Jahr 1988 die *European Evaluation Society* ins Leben gerufen worden war, wurde 1997 die *Deutsche Gesellschaft für Evaluation*, inzwischen *Gesellschaft für Evaluation* (DeGEval) (<http://www.degeval.de/home>) gegründet. An die ‚evaluation communities‘ in Deutschland und Österreich adressiert, war die DeGEval eine der ersten nationalen Fachverbände in Europa. Ein Jahr vorher wurde die *Schweizerische Evaluationsgesellschaft* (SEVAL) gegründet. Beide Gründungen sind Meilensteine auf dem Weg zur Institutionalisierung und Professionalisierung der Evaluationsforschung in Deutschland, Österreich und der Schweiz (vgl. Stockmann 2006a: 37). Inzwischen haben sich innerhalb der DeGEval 17 Arbeitskreise gebildet (siehe www.degeval.de/arbeitskreise, für eine Übersicht vgl. Böttcher et al. 2014a, 2014b: S. 215). Die jährlichen Konferenzen der DeGEval mit Workshops ihrer Arbeitskreise tragen wesentlich dazu bei, den fachlichen Diskurs und das professionelle Selbstverständnis der deutschen und österreichischen *evaluation communities* zu fördern. Gelegentlich organisieren DeGEval und SEVAL gemeinsame Tagungen, um vergleichende Untersuchungen und gemeinsame Publikationen anzuregen und hervorzuheben (vgl. das rund 600-seitige Handbuch Widmer/Beywl/Fabian 2009, das, im Anschluss an eine gemeinsame Tagung der beiden nationalen Fachverbände entstanden, einen breiten Überblick über den Stand der Evaluationsforschung in Deutschland, Österreich und der Schweiz gibt).

Als ein wichtiges Diskussionsforum hat sich über die Jahre die *Gesellschaft für Programmforschung* (GfP) erwiesen, die, wie weiter vorn erwähnt, 1981 auf Ini-

tative von *Hans-Ulrich Derlien* gegründet, seither (inzwischen über 30) jährliche Tagungen veranstaltet hat. Sie bringt mit Evaluation befassende Wissenschaftler und Praktiker zusammen und politik- und evaluationsrelevante Themen auf ihre Tagesordnung. Die Vorträge und Ergebnisse der Tagungen werden regelmäßig veröffentlicht (vgl. zuletzt Gesellschaft für Programmforschung 2009; Schimanke 2007; Schimanke/Fischer/Bucksteed 2006).

Seit 2002 erscheint zweimal jährlich die *Zeitschrift für Evaluation*, die sich als (im deutschen Sprachraum einzige) Fachzeitschrift für Evaluation Verbreitung und Anerkennung verschafft hat (vgl. Stockmann 2002, 2006a: 32).

Schließlich ist als ein signifikanter Beitrag zur Institutionalisierung und Professionalisierung der Evaluationsforschung in Deutschland das *Centrum für Evaluation*, CEval, hervorzuheben, das 2002 auf Initiative und unter maßgeblicher Mitwirkung von *Reinhard Stockmann* an der Universität des Saarlandes mit dem Ziel gegründet wurde, Lehre und Praxis auf dem Gebiet der Evaluation zu verbinden (vgl. Filsinger 2016: 5, vgl. auch <http://www.ceval.de/modx>). Anlässlich seines 10-jährigen Bestehens organisierte das CEval eine internationale Tagung, aus der ein Sammelband „The Future of Evaluation, Global Trends etc.“ hervorgegangen ist (vgl. Stockmann/Meyer 2016a).

In folgerichtiger Fortführung wurde 2004 – wiederum wesentlich auf Initiative von *Reinhard Stockmann* – an der Universität des Saarlandes und der Hochschule für Technik und Wirtschaft des Saarlandes ein gemeinsamer hochschulübergreifender und interdisziplinär ausgerichteter Aufbaustudiengang eingerichtet, der, in Deutschland erstmalig, mit dem akademischen Grad eines „*Master Evaluation*“ (MEval) abgeschlossen werden kann (siehe <http://www.master-evaluation.de/studiengang/profil>). In dieser Kombination von Forschungszentrum und Studiengang nimmt die Universität des Saarlandes im Feld der Evaluationsforschung in Deutschland eine herausgehobene Stellung ein.

4.3 Evaluationsforschung zwischen Unabhängigkeit und Fremdbestimmung

Wurden die Unabhängigkeit der Evaluationsforschung und ihre Gefährdung insbesondere als Auftragsforschung (*contractual research*, vgl. Wollmann 2002b) von Anfang an lebhaft und kritisch (etwa als *epistemic drift* oder *Kolonialisierung*) diskutiert, so rückten diese Fragen angesichts des weiteren rasanten Wachstums der Evaluationsforschung und ihrer fortschreitenden „teildisziplinären“ Spezialisierung und Fragmentierung nur noch dringlicher auf die Agenda.

In einer neueren Diskussion wird die Gefährdung der forschungsethischen und wissenschaftsmethodischen Unabhängigkeit der Evaluationsforschung als Auftragsforschung vor allem in ihrer „Entwissenschaftlichung“ (*de-scientification*, Meyer/Stockmann 2016: 356) als Verlust von methodischer Stringenz und wissenschaftlicher Rigorosität diskutiert. Abhilfe wird von einer verstärkten institutionellen und konzeptionellen Verknüpfung mit universitärer Forschung erwartet (für eine ausführ-

liche Diskussion von „evaluation between science and utility“ vgl. Stockmann/Meyer 2016b).

Wie schon in der früheren Diskussion über Begleitforschung, insbesondere über ihre ‚intervenierende‘ Variante (vgl. Lutz 1983), ergibt sich die Frage nach wissenschaftlicher Unabhängigkeit von Evaluationsforschung mit besonderer Schärfe im Zusammenhang mit Evaluierungsansätzen, in denen es vor allem um Anwendung und Akzeptanz der Evaluationsergebnisse geht und in denen auf eine enge Mitarbeit der zu evaluierenden Akteure („stakeholders“) und Projekte gesetzt wird. Solche Ansätze finden sich in terminologisch und auch konzeptionell leicht unterschiedlichen Varianten (etwa als „utilization-focused evaluation“ (Patton 1997), „pragmatic use evaluation“ (Mertens 2016: 217), „critical friend approach“ (Balthasar 2012: 173) oder als andere „participatory“ Varianten). In diesem Zusammenhang wird argumentiert, dass die Gefahr wissenschaftlicher „Unschärfe“ durch „increased use of mixed methods designs that are appropriate for answering complex questions“ (Mertens 2016: 218) gebannt werden könne.

Vor dem Hintergrund dieser vielfältigen forschungsethischen und methodischen Probleme hat die DeGEVal in Anlehnung an die von der *American Evaluation Society* 1994 beschlossenen „Guiding Principles for Evaluators“ im Jahr 2002 „Standards für Evaluation“ verabschiedet (vgl. DeGEVal 2002), die u.a. darauf zielen, im Verhältnis von (auftragnehmenden) Evaluationsforscher(inne)n und (auftragerteilenden) Politiker(inne)n/Administrator(inn)en beiderseits anerkannte forschungsethische und -methodische Maximen festzulegen und ihnen Geltung zu verschaffen (vgl. Stockmann 2006a: 39).

4.4 Zusammenfassung

Die weitere Expansion und Konsolidierung der Evaluationsforschung in Deutschland ist seit den 1990er Jahren durch eine Reihe von Faktoren angetrieben worden, von denen vor allem die internationale, wenn nicht globale Dominanz der von *New Public Management* (NPM) inspirierten Politik- und Verwaltungsmodernisierung (vgl. Meyer/Stockmann 2016: 331ff.) sowie die von der EU für ihre Strukturförderprogramme vorgeschriebenen Evaluierungsverfahren zu nennen sind. Die Gründung der *DeGEVal* vor nunmehr 20 Jahren ist als ein Meilenstein für Institutionalisierung, Qualifizierung und Professionalisierung der Evaluation in Deutschland und Österreich zu würdigen.

Der gegenwärtige Stand der Evaluationsforschung in Deutschland und Österreich ist in zwei im Kontext der *DeGEVal* entstandenen Sammelbänden eindrucksvoll dokumentiert, nämlich in Widmer/Beywl/Fabian 2009, wo die deutsche und österreichische Entwicklung in den Vergleich mit der schweizerischen gerückt wird, sowie in Böttcher et al. 2014a mit einschlägigen Berichten aus *DeGEVal*-Arbeitskreisen. Ferner sei an die jüngsten Tagungsberichte der *Gesellschaft für Programmforschung* erinnert, in denen aktuelle Themen aufgegriffen werden. Schließlich – fürwahr last but not least – ist auf die halbjährlich erscheinende *Zeitschrift für*

Evaluation zu verweisen als eine für Wissenschaft und Praxis unverzichtbare Informationsquelle zur aktuellen Entwicklung.

5. Postskriptum

Aus Platzgründen soll in dieser Übersicht darauf verzichtet werden, die wichtigen Fragen der *Verwendung und Nutzung von Evaluationsergebnissen* in Politik, Verwaltung und Gesellschaft aufzugreifen. Es muss genügen, auf die Aufsätze des unlängst aus einer DeGEval-Arbeitsgemeinschaft hervorgegangenen Sammelbandes Hense/Rädiker/Böttcher/Widmer 2013, auf Wollmann 2013 sowie jüngst auf einschlägige Aufsätze in Stockmann/Meyer 2016a, insbesondere auf Stockmann/Meyer 2016b hinzuweisen.

Danksagung

Der Autor hat den ‚reviewers‘ der Entwurfsfassung dieses Aufsatzes für ihre hilfreichen textlichen und inhaltlichen Hinweise und Anregungen zu danken.

6. Literatur

- Balthasar, Andreas (2012): Fremd- und Selbstevaluation kombinieren: der ‚Critical Friend Approach‘ als Option. In: Zeitschrift für Evaluation, 11 (2), S. 173-192.
- Bangel, Bettina/Brinkmann, Christian/Deeke, Axel (2006): Evaluation von Arbeitsmarktpolitik. In: Stockmann, Reinhard (Hg.): Evaluationsforschung. Münster: Waxmann (3. Aufl.), S. 311-344.
- Beyme, Klaus von (1985): Policy Analyse und traditionelle Politikwissenschaft. In: Hartwich, Hans-Hermann (Hrsg.): Policy-Forschung in der Bundesrepublik Deutschland. Ihr Selbstverständnis und ihr Verhältnis zu den Grundfragen der Politikwissenschaft. Wiesbaden: Westdeutscher Verlag, S. 7-29.
- Bogumil, Jörg/Grohs, Stephan/Kuhlmann, Sabine/Ohm, Anna K. (2007): Zehn Jahre Neues Steuerungsmodell. Berlin: Sigma.
- Böhret, Carl/Hugger, Werner (1980): Der Praxistest von Gesetzentwürfen: am Beispiel d. Referentenentwurfs e. Jugendhilfegesetzes. Baden-Baden: Nomos.
- Böttcher, Wolfgang/Kerlen, Christiane/Maats, Peter/Schwab, Oliver/Sheikh, Sonja (Hg.) (2014a): Evaluation in Deutschland und Österreich. Münster: Waxmann.
- Böttcher, Wolfgang/Kerlen, Christiane/Maats, Peter/Schwab, Oliver/Sheikh, Sonja (2014b): Arbeitsfelder und Herausforderungen der Evaluation. In: Böttcher, Wolfgang/Kerlen, Christiane/Maats, Peter/Schwab, Oliver/Sheikh, Sonja (Hg.): Evaluation in Deutschland und Österreich. Münster: Waxmann, S. 7-14.
- Büeler, Xaver (2006): Qualitätsevaluation und Schulentwicklung. In: Stockmann, Reinhard (Hg.): Evaluationsforschung. Münster: Waxmann (3. Aufl.), S. 260-288.
- Campbell, Donald (1969): Reforms as Experiments. In: American Psychologist, 24 (4), S. 409-429.
- DeGEval (2002): Standards für Evaluation. Köln: Geschäftsstelle DeGEval, www.degeval.de/?id=135.
- Derlien, Hans-Ulrich (1976): Die Erfolgskontrolle staatlicher Planung, Eine empirische Untersuchung über Organisation, Methode und Politik der Programmevaluation. Baden-Baden: Nomos.

- Derlien, Hans-Ulrich (1981): Stand und Entwicklung der Programmforschung in der öffentlichen Verwaltung. In: Derlien, Hans-Ulrich (Hg.): Programmforschung in der öffentlichen Verwaltung, Werkstattbericht 1 der Gesellschaft für Programmforschung. München, S. 5-45.
- Derlien, Hans-Ulrich (1990a): Program Evaluation in the Federal Republic of Germany. In: Rist, Ray C. (Hg.): Program Evaluation and the Management of Government. New Brunswick/London: Transaction, S. 37-52.
- Derlien, Hans-Ulrich (1990b): Genesis and Structure of Evaluation Efforts in Comparative Perspective. In: Rist, Ray C. (Hg.): Program Evaluation and the Management of Government. New Brunswick/London: Transaction, S. 147-176.
- Derlien, Hans-Ulrich/Rist, Ray C. (2002): Policy Evaluation in International Comparison. In: Furubo, Jan-Eric/Rist, Ray/Sandahl, Rolf (Hg.): International Atlas of Evaluation. New Brunswick: Transaction, S. 425-456.
- Dierkes, Meinolf/Knie, Andreas/Wagner, Peter (1988): Die Diskussion über das Verhältnis von Technik und Politik in der Weimarer Republik. In: Leviathan, 16 (1), S. 1-23.
- Elzinga, Aant (1985): Research Bureaucracy and the Drift of Epistemic Criteria. In: Wittrock, Björn/Elzinga, Aant (Hg.): The University Research System. Stockholm: Almqvist & Wiksell International, S. 191-220.
- Filsinger, Dieter (2016): Ten Years of the Center for Evaluation. In: Stockmann, Reinhard/Meyer, Wolfgang (Hg.): The Future of Evaluation. Hampshire: Palgrave Macmillan, S. 2-8.
- Frais, Michael/Renz, Monika (2014): Evaluation im Schulbereich. In: Böttcher, Wolfgang/Kerlen, Christiane/Maats, Peter/Schwab, Oliver/Sheikh, Sonja (Hg.): Evaluation in Deutschland und Österreich. Münster: Waxmann, S. 89-99.
- Freeman, Howard/Solomon, Marian A. (1981): The Next Decade of Evaluation Research. In: Levine, Robert A./Solomon, Marian A./Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Evaluation Research and Practice. Comparative and International Perspectives. Beverly Hills/London: Sage, S. 12-27.
- Furubo, Jan-Eric/Rist, Ray/Sandahl, Rolf (Hg.) (2002): International Atlas of Evaluation. New Brunswick: Transaction.
- Furubo, Jan-Eric/Sandahl, Rolf (2002): A Diffusion Perspective on Global Developments in Evaluation. In: Furubo, Jan-Eric/Rist, Ray/Sandahl, Rolf (Hg.): International Atlas of Evaluation. New Brunswick: Transaction, S. 1-26.
- Gesellschaft für Programmforschung (Hg.) (2009): Better Regulation: Bessere Institutionen und Normen als Voraussetzung für erfolgreiches Regieren. Lohmar: Eul.
- Grajewski, Regina/Meyer, Stefan (2014): Stand der Evaluation in der Strukturpolitik. In: Böttcher, Wolfgang/Kerlen, Christiane/Maats, Peter/Schwab, Oliver/Sheikh, Sonja (Hg.): Evaluation in Deutschland und Österreich. Münster: Waxmann, S. 37-59.
- Hellstern, Gerd-Michael/Wollmann, Hellmut (1978): Sanierungsmaßnahmen. Städtebauliche und stadtstrukturelle Wirkungen. Methodische Vorstudie. Bad Godesberg: Schriftenreihe des BMBau.
- Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.) (1983a): Experimentelle Politik – Reformstrohfeuer oder Lernstrategie. Opladen: Westdeutscher Verlag.
- Hellstern, Gerd-Michael/Wollmann, Hellmut (1983b): Bilanz – Reformexperimente, wissenschaftliche Begleitung und politische Realität. In: Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Experimentelle Politik – Reformstrohfeuer oder Lernstrategie. Opladen: Westdeutscher Verlag, S. 1-78.
- Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.) (1984a): Handbuch zur Evaluierungsforschung. Bd. 1, Opladen: Westdeutscher Verlag.
- Hellstern, Gerd-Michael/Wollmann, Hellmut (1984b): Evaluierung und Evaluierungsforschung – ein Entwicklungsbericht. In: Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Handbuch zur Evaluierungsforschung. Bd. 1, Opladen: Westdeutscher Verlag, S. 17-90.
- Hellstern, Gerd-Michael/ Wollmann, Hellmut (Hg.) (1984c): Evaluierung und Erfolgskontrolle in Kommunalpolitik und -verwaltung. Basel: Birkhäuser.
- Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.) (1984d): Evaluierung und Erfolgskontrolle auf der kommunalen Ebene. Ein Überblick. In: Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Evaluierung und Erfolgskontrolle in Kommunalpolitik und -verwaltung. Basel: Birkhäuser, S. 10-45.
- Hense, Jan/Rädiker, Stefan/Böttcher, Wolfgang/Widmer, Thomas (Hg.) (2013): Forschung über Evaluation. Bedingungen, Prozesse und Wirkungen. Münster: Waxmann.

- Heyer, Gerd (2006.): Zielsetzung und Struktur der „Hartz-Evaluation“. In: Zeitschrift für Arbeitsmarktforschung, 30 (3/4), S. 467-476.
- Hohn, Hans-Willy/Schimank, Uwe (1990): Konflikte und Gleichgewicht im Wissenschaftssystem. Frankfurt am Main: Campus.
- Jaedicke, Wolfgang/Thrun, Thomas/Wollmann, Hellmut (2000): Modernisierung der Kommunalverwaltung. Evaluierungsstudie zur Verwaltungsmodernisierung im Bereich Planen, Bauen und Umwelt. Stuttgart: Kohlhammer.
- Konzendorf, Götz (2009): Institutionelle Einbettung der Evaluationsfunktion in Politik und Verwaltung in Deutschland. In: Widmer, Thomas/Beywl, Wolfgang/Carlo, Fabian (Hg.): Evaluation. Ein systematisches Handbuch. Wiesbaden: VS, S. 27-33.
- Konzendorf, Götz (2011): Gesetzesfolgenabschätzung. In: Blanke, Bernhad/Nullmeierl, Frank/Reichard, Christoph/Wewer, Göttrik (Hg.): Handbuch zur Verwaltungsreform. Wiesbaden: VS (4. Aufl.), S. 135-143.
- Kritz, Jürgen (2004): Aktionsforschung. In: Nohlen, Dieter/Schultze, Rainer-Olaf (Hg.): Lexikon der Politikwissenschaft. Bd. 1, München: C.H. Beck, S. 10-11.
- Kromrey, Helmut (1995): Evaluation. Empirische Konzepte zur Bewertung von Handlungsprogrammen und die Schwierigkeiten ihrer Realisierung. In: Zeitschrift für Sozialforschung und Erziehungssoziologie, 15 (4), S. 313-336.
- Kromrey, Helmut (2003): Evaluation in Wissenschaft und Gesellschaft. In: Zeitschrift für Evaluation, 2 (1), S. 93-116.
- Kuhlmann, Sabine (2009): Die Evaluation von Institutionenpolitik in Deutschland: Verwaltungsmodernisierung und Wirkungsanalyse im föderalen System. In: Widmer, Thomas/Beywl, Wolfgang/Carlo, Fabian (Hg.): Evaluation. Ein systematisches Handbuch. Wiesbaden: Springer, S. 371-380.
- Kuhlmann, Sabine/Bogumil, Jörg/Grohs, Stefan (2008): Evaluating administrative modernization in German local governments. In: Public Administrative Review, 68 (5), S. 851-863.
- Kuhlmann, Sabine/Wollmann, Hellmut (2006): Transaktionskosten von Verwaltungsreformen – ein „missing link“ der Evaluationsforschung. In: Jann, Werner/Röber, Manfred/Wollmann, Hellmut (Hg.): Public Management. Grundlagen, Wirkungen, Kritik. Wiesbaden: Springer, S. 371-390.
- Kuhlmann, Sabine/Wollmann, Hellmut (2013): Verwaltung und Verwaltungsreformen in Europa. Wiesbaden: Springer.
- Kuhlmann, Stefan (2006): Evaluation in der Forschungs- und Innovationspolitik. In Stockmann, Reinhard (Hg.): Evaluationsforschung. Münster: Waxmann (3. Aufl.), S. 289-310.
- Kuhlmann, Stefan (2009): Evaluation von Forschungs- und Innovationspolitik in Deutschland – Stand und Perspektiven. In: Widmer, Thomas/Beywl, Wolfgang/Carlo, Fabian (Hg.): Evaluation. Ein systematisches Handbuch. Wiesbaden: Springer, S. 283-294.
- Lasswell, Harold D. (1951): The Policy Orientation. In: Lerner, Daniel/Lasswell, Harold D. (Hg.): Policy Sciences. Stanford University Press, S. 3-15.
- Leeuw, Frans (2006): Evaluation in Europe. In: Stockmann, Reinhard (Hg.): Evaluationsforschung. Münster: Waxmann (3. Aufl.), S. 64-84.
- Levine, Robert A. (1981). Program Evaluation and Policy Analysis in Western Nations: An Overview. In: Levine, Robert A./Solomon, Marian A./Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Evaluation Research and Practice. Comparative and International Perspectives. Beverly Hills/London: Sage, S. 12-27.
- Levine, Robert A./Solomon, Marian A./Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.) (1981): Evaluation Research and Practice. Comparative and International Perspectives. Beverly Hills/London: Sage.
- Lorz, Rainer E. (1984): Das Inspektionsreferat des Bundesministeriums für wirtschaftliche Zusammenarbeit. In: Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Handbuch zur Evaluationsforschung, Bd. 1. Opladen: Westdeutscher Verlag.
- Lutz, Burkhard (1983): Zur Problematik programmbegleitender Sozialforschung. In: Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Experimentelle Politik – Reformstrohfeuer oder Lernstrategie. Opladen: Westdeutscher Verlag, S. 258-261.
- Maag Merki, Katharina (2009): Evaluation im Bildungsbereich Schule in Deutschland. In: Widmer, Thomas/Beywl, Wolfgang/Carlo, Fabian (Hg.): Evaluation. Ein systematisches Handbuch. Wiesbaden: Springer, S. 157-162.
- Maier, Hans (1980): Die ältere deutsche Staats- und Verwaltungslehre. München: C.H. Beck.

- Mertens, Donna M. (2006): Institutionalizing Evaluation in the United States of America. In: Stockmann, Reinhard (Hg.): Evaluationsforschung. Münster: Waxmann (3. Aufl.), S. 47-63.
- Mertens, Donna M. (2016): The Role of Evaluation in Social Change: Perspectives on the Future of Evaluation from the USA. In: Stockmann, Reinhard/Meyer, Wolfgang (Hg.): The Future of Evaluation. Global Trends, New Challenges, Shared Perspectives. Hampshire: Palgrave Macmillan, S. 214-217.
- Meyer, Wolfgang/Stockmann, Reinhard (2016): Conclusion: Shared Perspectives or a United World of Evaluation. In: Stockmann, Reinhard/Meyer, Wolfgang (Hg.): The Future of Evaluation. Global Trends, New Challenges, Shared Perspectives. Hampshire: Palgrave Macmillan, S. 328-357.
- Nationaler Normenkontrollrat (2016): 10 Jahre NKR – gute Bilanz bei Bürokratieabbau und Folgekostenbegrenzung – alarmierender Rückstand bei E-Government, Jahresbericht 2016.
- Neubert, Susanne/Mack, Dorothee/Roxin, Helge (2014): Vom Methodenstreit zum Methodenmix – Chronologie und Stand der Entwicklungsevaluierung in Deutschland. In: Böttcher, Wolfgang/Kerlen, Christiane/Maats, Peter/Schwab, Oliver/Sheikh, Sonja (Hg.): Evaluation in Deutschland und Österreich. Münster: Waxmann, S. 61-72.
- Patton, Michael Quinn (1997): Utilization-Focused Evaluation. Thousand Oaks: Sage (3. Aufl.).
- Pollitt, Christopher/Bouckaert, Geert (2011): Public Management Reform. Oxford: Oxford University Press (3. Aufl.).
- Reichard, Christoph (1994): Umdenken im Rathaus. Berlin: edition sigma.
- Rein, Martin (1981): Comprehensive Program Evaluation. In: Levine, Robert A./Solomon, Marian A./Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Evaluation Research and Practice. Comparative and International Perspectives. Beverly Hills/London: Sage, S. 132-149.
- Rossi, Peter (1981): The Professionalization of Evaluation Research in the United States. In: Levine, Robert A./Solomon, Marian A./Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): Evaluation Research and Practice. Comparative and International Perspectives. Beverly Hills/London: Sage, S. 220-236.
- Rossi, Peter/Freeman, Howard (1993): Evaluation. A Systematic Approach. Thousand Oaks u.a.: Sage.
- Rossi, Peter/Freeman, Howard/Hofmann, Gerhard (1988): Programm Evaluation. Einführung in die Methoden angewandter Sozialforschung. Stuttgart: Enke.
- Scharmer, Eckart/Wollmann, Hellmut/Argast, Michael (1985): Rechtstatsachenuntersuchung zur Baugenehmigungspraxis. Schriftenreihe 03, Bad Godesberg: Bundesministerium für Raumordnung, Bauwesen und Städtebau.
- Schimanke, Dieter (2007): Qualität und Ergebnis öffentlicher Programme. Münster: Waxmann.
- Schimanke, Dieter/Fischer, Adrea/Bucksteed (Hg.) (2006): Wie lernt Politik? Münster: Waxmann.
- Schreckenberger, Waldemar/König, Klaus/Zeh, Wolfgang/Baden, Eberhard (Hg.) (1986): Gesetzgebungslehre. Grundlagen – Zugänge – Anwendung. Stuttgart u.a.: Kohlhammer.
- Schwab, Oliver (2009): Evaluierung von Raumordnungspolitik in Deutschland. In: Widmer, Thomas/Beywl, Wolfgang/Carlo, Fabian (Hg.): Evaluation. Ein systematisches Handbuch. Wiesbaden: Springer, S. 403-413.
- Stockmann, Reinhard (2000): Evaluation in Deutschland. In: Stockmann, Reinhard (Hg.): Evaluationsforschung. Opladen: Leske + Budrich.
- Stockmann, Reinhard (2002): Editorial. Zur Notwendigkeit und Konzeption einer deutschsprachigen „Zeitschrift für Evaluation“. In: Zeitschrift für Evaluation, 1 (1), S. 3-9.
- Stockmann, Reinhard (2006a): Evaluation in Deutschland. In: Stockmann, Reinhard (Hg.): Evaluationsforschung. Münster: Waxmann (3. Aufl.), S. 15-46.
- Stockmann, Reinhard (2006b): Evaluation staatlicher Entwicklungspolitik. In: Stockmann, Reinhard (Hg.): Evaluationsforschung. Münster: Waxmann (3. Aufl.), S. 378-414.
- Stockmann, Reinhard/Meyer, Wolfgang (Hg.) (2016a): The Future of Evaluation. Global Trends, New Challenges, Shared Perspectives. Hampshire: Palgrave Macmillan.
- Stockmann, Reinhard/Meyer, Wolfgang (2016b): Evaluation between Science and Utility. In: Stockmann, Reinhard/Meyer, Wolfgang (Hg.): The Future of Evaluation. Global Trends, New Challenges, Shared Perspectives. Hampshire: Palgrave Macmillan, S. 238-253.
- Wagner, Gerald (2009): Evaluation in der Arbeitsmarktpolitik in Deutschland. In: Widmer, Thomas/Beywl, Wolfgang/Carlo, Fabian (Hg.): Evaluation. Ein systematisches Handbuch. Wiesbaden: Springer, S. 117-126.
- Wagner, Peter (1990): Sozialwissenschaften und Staat. Frankreich, Italien, Deutschland 1870-1980. Frankfurt am Main: Campus.

- Wagner, Peter/Wollmann, Hellmut (1986a): Fluctuations in the Development of Evaluation Research. Do 'Regime Shifts' Matter? In: *International Social Science Journal*, 38 (2), S. 205-218.
- Wagner, Peter/Wollmann, Hellmut (1986b): Social Scientists in Policy Research and Consulting. Some Cross-National Comparisons. In: *International Social Science Journal*, 38 (4), S. 601-617.
- Wagner, Peter/Wollmann, Hellmut (1991): Beyond Serving State and Bureaucracy: Problem-Oriented Social Science in (West)Germany. In: *Knowledge and Policy*, 4 (1/2), S. 56-88.
- Weiss, Carol H. (1974): *Evaluierungsforschung*. Opladen: Westdeutscher Verlag (dt. Übersetzung von Weiss, Carol H. (1972): *Evaluation Research*. Englewood Cliffs: Prentice Hall).
- Wheley, Josef S. (1981): Using Evaluation to Improve Program Performance. In: Levine, Robert A./Solomon, Marian A./Hellstern, Gerd-Michael/Wollmann, Hellmut (Hg.): *Evaluation Research and Practice. Comparative and International Perspectives*. Beverly Hills/London: Sage, S. 92-106.
- Widmer, Thomas/Beywl, Wolfgang/Fabian, Carlo (Hg.) (2009): *Evaluation. Ein systematisches Handbuch*. Wiesbaden: Springer.
- Wittmann, Werner (1985): *Evaluationsforschung. Aufgaben, Probleme und Anwendungen*. Berlin u.a.: Springer.
- Wittrock, Björn/Wagner, Peter/Wollmann (1991): Social Science and the Modern State: Policy Knowledge and Political Institutions in Western Europe and the United States. In: Wagner, Peter/Weiss, Carol Hirschon/Wittrock, Björn/Wollmann, Hellmut (Hg.): *Social Science and the Modern State*. Cambridge: Cambridge University Press, S. 28-85.
- Wollmann, Hellmut (1982): Untersuchungsansätze und Nutzungschancen einer Rechtsstatsachenforschung im Städtebaurecht. In: *Informationen zur Raumentwicklung*, 1982, Bundesinstitut für Bau-, Stadt- und Raumforschung, S. 1-20.
- Wollmann, Hellmut (1984): Policy Analysis. Some Observations on the West German Scene. In: *Policy Sciences*, 17 (1), S. 27-48.
- Wollmann, Hellmut (1986): Gesetzgebung als experimentelle Politik – Möglichkeiten, Varianten und Grenzen erfahrungswissenschaftlich fundierter Gesetzgebungsarbeit. In: Schreckenberger, Walde-mar/König, Klaus/Zeh, Wolfgang/Baden, Eberhard u.a. (Hg.): *Gesetzgebungslehre. Grundlagen – Zugänge – Anwendung*. Stuttgart u.a.: Kohlhammer, S. 72-95.
- Wollmann, Hellmut (1989): Policy Analysis in West Germany's Federal Government: A Case of Unfinished Governmental and Administrative Modernization? In: *Governance*, 2 (3), S. 233-263.
- Wollmann, Hellmut (1990): Konzept und Methode der Begleitforschung. In: *Informationen zur Raumentwicklung*, Heft 10/11, Bundesinstitut für Bau-, Stadt- und Raumforschung, S. 563-575.
- Wollmann, Hellmut (2002a): Verwaltungspolitik und Evaluierung. Ansätze, Phasen und Beispiele im Ausland und in Deutschland. In: *Zeitschrift für Evaluation*, 2 (2), S. 75-101.
- Wollmann, Hellmut (2002b): Contractual Research and Policy Knowledge. In: *International Encyclopedia of the Social and Behavioral Sciences*, 5 (1), S. 11574-11578.
- Wollmann, Hellmut (Hg.) (2003a): *Evaluation in Public Sector-Reform*. Cheltenham: Edward Elgar.
- Wollmann, Hellmut (2003b): Evaluation in Public-Sector Reform. Towards a 'Third Wave' of Evaluation? In: Wollmann, Hellmut (Hg.): *Evaluation in Public Sector-Reform*. Cheltenham: Edward Elgar, S. 1-11.
- Wollmann, Hellmut (2003c): Evaluation in Public Sector Reform. Trends, Potentials and Limits in International Perspective. In: Wollmann, Hellmut (Hg.): *Evaluation in Public Sector-Reform*. Cheltenham: Edward Elgar, S. 231-258.
- Wollmann, Hellmut (2006): Evaluation und Verwaltungspolitik. Konzepte und Praxis in Deutschland und im internationalen Kontext. In: Stockmann, Reinhard (Hg.): *Evaluationsforschung*. Münster: Waxmann (3. Aufl.), S. 207-233.
- Wollmann, Hellmut (2007): Policy Evaluation and Evaluation Research. In: Fischer, Frank/Miller, Gerald J./Sidney, Mara S. (Hg.): *Handbook of Public Policy Analysis*. Boca Raton u.a.: CRC Press, S. 393-404.
- Wollmann, Hellmut (2009): Kontrolle in Politik und Verwaltung: Evaluation, Controlling und Wissensnutzung. In Schubert, Klaus/Brandelow, Nils (Hg.): *Lehrbuch der Politikfeldanalyse 2.0*. München: Oldenbourg (2. Aufl.), S. 379-400.
- Wollmann, Hellmut (2010): Soziologie zwischen Kaiserreich, Weimarer Republik und NS-Regime. In: Tenorth, Heinz-Elmar (Hg.): *Geschichte der Universität Unter den Linden 1810-2010*, Band 5. Berlin: Akademie-Verlag, S. 257-274.

- Wollmann, Hellmut (2013): Zur (Nicht-)Verwendung von Evaluationsergebnissen in Politik und Verwaltung. Eine vernachlässigte Fragestellung der Evaluationsforschung. In: Kropp, Sabine/Kuhlmann, Sabine (Hg.): Wissen und Expertise in Politik und Verwaltung, der moderne Staat. Sonderheft 1, Opladen: Verlag Barbara Budrich, S. 87-103.
- Wollmann, Hellmut (2017): Spurensuche. Warum ich Politikwissenschaftler wurde und worüber ich arbeitete und schrieb. In: Kuhlmann, Sabine/Schwab, Oliver (Hg.): Starke Kommunen – wirksame Verwaltung. Berlin: Springer, S. 241-280.
- Wollmann, Hellmut/Hellstern, Gerd-Michael (1977): Sozialwissenschaftliche Untersuchungsregeln und Wirkungsforschung. In: Haungs, Peter (Hg.): Res Publica, Studien zum Verfassungswesen. München: Fink Verlag, S. 415-466.
- Wottawa, Heinrich/Thierau, Heike (1990): Lehrbuch Evaluation. Bern: Huber.
- Wottawa, Heinrich/Thierau, Heike (2003): Lehrbuch Evaluation. Bern: Huber (3. Aufl.).