

Evaluation und Verwaltungspolitik
Konzepte und Praxis in Deutschland und im internationalen Kontext

Hellmut Wollmann
Humboldt-Universität
Institut für Sozialwissenschaften
Unten den Linden 6
10099 Berlin
e-mail hellmut.wollmann@rz.hu-berlin.de

in: Stockmann, Reinhard (Hrsg.) 2004 , Evaluationsforschung, 2. Aufl.,

Opladen: Leske + Budrich, pp. 205-232

1. Konzeptionelle und methodische Eigentümlichkeiten der Evaluation von Verwaltungspolitik¹

1.1. Die mehreren „Schleifen“ von Evaluation

Zunächst sei darauf hingewiesen, dass sich *Verwaltungs-* oder *Institutionenpolitik* als ein komplexer Politiktyp erweist, der in mehreren Handlungsphasen und je spezifischen Ursache-Wirkungs-„Schleifen“ abläuft - im Gegensatz zu „normalen“ („substanziellen“) Politiken². Die letzteren zielen darauf, bestimmte (sozio-ökonomische, ökologogische, infrastrukturelle usw.) Veränderungen im Umfeld von Politik und Verwaltung unter Einsatz bestimmter rechtlicher Regelungen, finanzieller Mittel usw. zu erreichen (z.B. Sozial-, Beschäftigungs-, Umweltpolitik usw.). Ihnen liegt denn – vielfach eher implizit als explizit – eine verhältnismäßig „einfache“ Handlungstheorie zugrunde³, in der angenommen wird, dass die angestrebte Veränderung mit Hilfe der in Aussicht genommenen Handlungsmittel – sozusagen in *einer* Ursache/Wirkungs- und Mittel/Ziel-„Schleife“ – zu erreichen sei.

- In ihrer „ersten“ Schleife – und in ihrem institutionenpolitischen Kern – ist Verwaltungspolitik darauf gerichtet, die bestehenden politisch-administrativen (organisatorischen, personellen, qualifikatorischen, verwaltungskulturellen usw.) Strukturen durch den Einsatz geeigneter Strategien und Maßnahmen zu verändern. Soweit die Evaluation auf diese institutionellen Veränderungen fokussiert ist, untersucht sie deren Verlauf und (Zwischen-)Ergebnisse unter der Fragestellung, ob die intendierten (institutionenpolitischen) Ziele erreicht worden sind und welche Faktoren (methodisch gesprochen: als unabhängige Variablen) für den Grad der Zielerreichung bzw. -verfehlung (als abhängige Variable) bestimmend waren. Wegen des institutionenpolitischen Kerns dieses Reformschrittes wird die hierauf gerichtete Evaluation hier als *Institutionen-Evaluation* bezeichnet⁴.
- Als weiteres Glied der Ursache-Wirkungs-Kette strebt die Verwaltungspolitik an, eine Veränderung/Verbesserung der administrativen Handlungsprozesse⁵ zu erreichen. Die auf

¹ Die Ausarbeitung dieses Aufsatzes stützt sich auf Wollmann 2002a

² Zur Unterscheidung von „substanzieller“ und „institutioneller“ Politik vgl. auch Knöpfel/Bussmann 1997: 59 und Ritz 1999: 28

³ Zu den (Handlungs-) Theorien, von denen die politischen Akteure bei der Formulierung ihrer Politiken, politischen Programme und Interventionen vielfach (explizit oder implizit) ausgehen, vgl. Hofmann 1993 sowie – als „Klassiker“ der Implementationsforschung – Pressman/Wildavsky 1974

⁴ In der internationalen Diskussion wird diesbezüglich auch von „process Evaluation“ gesprochen, vgl. Christensen/ Laegreid/ Wise 2003: „Process evaluation tracks the extent to which programs or practices were put in place as intended and monitor how implementation has progressed“..

⁵ Vgl. Pollitt/Bouckaert, die insofern von „process improvement“ (Pollitt/Bouckaert 2000: 115) oder von „better processes“ (Pollitt/Bouckaert 2003) sprechen.

das Verwaltungshandeln als (prozesshafte) „performance“ fokussierte Evaluation wird hier als *Performanz-Evaluation* bezeichnet (vgl. ähnlich Thoenig 2003)

- In dem weiteren (und verwaltungspolitisch letztlich ausschlaggebenden) Glied der Ursache-Wirkungskette geht es um die angestrebten materiellen Wirkungen der Verwaltungsreformpolitik, deren unterschiedliche Reichweite vielfach als „output“, „outcomes“ und teilweise auch „impacts“ bezeichnet werden (vgl. etwa Pollitt/Bouckaert 2000: 110 f.). Während mit „outputs“ meist die unmittelbaren Verwaltungsleistungen und -„produkte“ gemeint werden (z.B. Zahl und Qualität der Entscheidungen, Dienstleistungen usw.), werden unter „outcomes“ (oder auch „impacts“) darüber hinausgehende (verteilungspolitische usw.) Wirkungen verstanden, die von den institutionellen Veränderungen, den veränderten Handlungsmustern („improved processes“) und/oder Ergebnissen („outputs“) des Verwaltungshandelns ausgelöst werden und sich u.a. auf die Verbesserung des administrativen Handlungssystems („system improvement“, Pollitt/Bouckaert 2000: 120, „improving the capacity of the administrative system“, Pollitt/Bouckaert 2003) oder auch der politisch-demokratischen Strukturen insgesamt („effects on the political-administrative decision-making system“, vgl. Christensen/Lagreid 2001: 32 f.) beziehen können. Vereinfacht kann hier von *output-* bzw. *outcome-*Evaluation gesprochen werden. .

1.2. Methodenprobleme

In der Evaluation von Verwaltungspolitik stellen sich spezifische methodische Schwierigkeiten, die die aus der Evaluation von „substanziellen“ Politikprogrammen, Maßnahmen usw. („ProgrammEvaluation“, vgl. Derlien 1981, Rist 1990) geläufigen methodischen Probleme (vgl. zuletzt insbesondere Pawson/Tilley 1997)⁶ noch deutlich übertreffen (grundlegend Pollitt 1995). Dabei geht es auch hier um die beiden konzeptionellen und methodischen Schlüsselfragen, zum einen die Ziele/Wirkungen (mit Hilfe geeigneter Indikatoren) zu konzipieren („Konzipierungsproblem) und zum andern (mit Hilfe geeigneter sozialwissenschaftlicher Methoden) die beobachteten Zielerreichungsgrade, Wirkungen usw. (als *abhängige Variable*) Zielerreichungen/Wirkungen der in Frage stehenden Intervention (Politik, Maßnahme) kausal zuzuordnen.

Von diesen Problemen seien hier die folgenden kurz diskutiert (ausführlicher Pollitt/Bouckaert 2001, Christensen/Lagreid 2001, Wollmann 2000a: 201 ff.):

⁶ Vgl. zuletzt insbesondere.

1.2.1. Konzipierungsproblem

Ermittlung der Ziele/Wirkungen und geeigneter Indikatoren

Entsprechend der soeben angesprochenen mehrgliedrigen Ziel-/Wirkungskette stellt sich als erstes Problem zu bestimmen, welche Reichweite die Evaluation haben soll (und kann) und welche Reformziele sich hieraus ergeben. Entlang der erwähnten Wirkungskette kann die Evaluationsuntersuchung gerichtet werden

- auf Verlauf und Ergebnis der institutionellen Veränderungen (Institutionen-Evaluation),
- auf die veränderten Handlungsprozesse und -muster (Performanz-Evaluation, „Prozess-Evaluation“),
- auf die (verwaltungsökonomische) Effizienz des Reformprojekts, insbesondere die erzielte Kostenreduktion (vgl. Pollitt/Bouckaert 2003: „saving of money“),
- auf die (materiellen) Leistungsverbesserungen (Output-Evaluation)
- oder auf darüber hinausreichende (verteilungspolitische, „systemische“ usw.) Veränderungen (Outcomes-Evaluation).

Die unter New Public Management firmierenden Modernisierungskonzepte stellen sich als ein Bündel von durchaus unterschiedlichen, teilweise widersprüchlichen Vorstellungen und Zielen dar. Pointiert wurde von einem „Einkaufskorb“ gesprochen („shopping basket“, Pollitt 1995), aus dem die Promotoren von Modernisierungsvorhaben je nach Zielsetzung, Neigung und Gusto bedienen. Damit steht die Evaluation eines konkreten Reformprojekts zunächst vor der Aufgabe, überhaupt erst einmal herauszufinden und zu klären, von welchen (organisations-, personal-, instrumenten- und „kultur“-bezogenen) Teilkonzepten und von welcher Mischung von Teilzielen sich die relevanten Akteure leiten lassen. Angesichts der möglichen Vielzahl der Reformkomponenten ist es mithin in der Regel geboten, die Evaluation auf bestimmte Teilelemente und deren Zielsetzungen zu beschränken und zu konzentrieren. Auch bei einer solchen (forschungspragmatisch gebotenen, jedoch eine umfassende Evaluation aus dem Blick rückenden) Fokussierung stellt als für veraltungspolitische Reformen spezifischen Problem, daß, dass es sich in der Regel nicht um „eindimensionale“, sondern um „multiple“ Ziele und Kriterien handelt (vgl. Pollitt/Bouckaert 2003), wechselseitig ergänzen, aber auch widersprechen und konterkarieren können (etwa: Kostenreduzierung versus Leistungsverbesserung, Verantwortungsdelegation versus Stärkung der politischen Verantwortung usw.). Angesichts der (im politischen Auseinandersetzungs- und Entscheidungsprozeß vielfach in Kauf genommen, wenn nicht beabsichtigen) Vagheit der Ziele und ihrer (zumindest latenten) Zielkonflikte und Widersprüche kann diese um eine Identifizierung der relevanten Handlungsziele (und der zugrunde liegenden „Handlungstheorien“) bemühte Einstiegsphase

deer Evaluation den (Neben-) Effekt haben, zur Aufhellung der Politikformulierung und der ihr verborgenen Ziel- und Handlungsannahmen beizutragen (vgl. Seyfried 2002),.

Ist die Bestimmung (und Auswahl) der Ziele (als intendierter Wirkungen) – sowie möglicherweise relevanter nicht-intendierter Folgen und Wirkungen - geleistet, stellt sich die Aufgabe, für diese untersuchungstaugliche, d.h. nach Möglichkeit operationalisierbare und meßbare Indikatoren zu konzipieren. Theoretische Überlegungen wie praktische Erfahrungen zeigen, dass die Formulierung brauchbarer Indikatoren um so schwieriger ist, je weicher und diffuser die Ziele sind, was wiederum um so ausgeprägter gilt, je weiter die Untersuchungsfrage die Wirkungskette abschreitet und insbesondere *outcomes*-Wirkungen zu erfassen sucht (vgl. Thoenig 2003 „Overall external effects... hard to measure“). Angesichts dessen kann es kaum verwundern, daß sich Evaluationsuntersuchungen vielfach auf solche Ziele als Meßplatte „stürzen“, die am ehesten operationalisiert zu sein scheinen (z.B. die Erfassung von *outputs* über die schiere Zahl von Entscheidungen, „Fällen“), während schwerer fassbare Wirkungsdimensionen (z.B. die Qualität der Entscheidungen und Leistungen) eher ausgeblendet bleiben.

Soweit Performanz-Indikatoren, Kennziffern festgelegt worden sind - sei durch externe Vorgaben (in Form von „citizen charters“), sei es durch Zielvereinbarungen/Kontrakte (extern zwischen politischer Führung und Verwaltung oder intern zwischen Verwaltungsführung und Verwaltungseinheiten), können diese als „Meßplatten“ der Evaluation dienen.

Sind Indikatoren auf eine brauchbare Weise definiert (operationalisiert), stellt sich als weitere Aufgabe, die für ihre empirische Darstellung erforderlichen Daten zu beschaffen. Zwar könnte vermutet werden, dass dies gerade im Rahmen von Verwaltungsreformprojekten dadurch erleichtert sein sollte, dass in nennenswertem Umfang „verwaltungsprozeßgenerierte“ Daten („Verwaltungsvollzugsdaten“) zur Verfügung stehen oder zumindest verhältnismäßig leicht beschaffbar zu sein scheinen. Die Praxis zeigt indessen, dass ungeachtet des Einzugs, den die Datenverarbeitung in der öffentlichen Verwaltung längst gemacht hat die Nutzung von „Verwaltungsvollzugsdaten“ für Zwecke der Evaluation nach wie vor nach auf überraschende Grenzen (und Blockierungen) stößt⁷.

1.2.2. *Kausalitätsproblem*

Ist schon bei der (ex-post) Evaluation „substanzieller“ Politiken die Anwendung des Repertoires der sozialwissenschaftlichen Methoden zur Gewinnung valider („kausaler“) Aussagen

⁷ Vgl. Jaedicke/Thrun/ Wollmann 2000: 154 (zum eingeschränkten Zugriff auf Verwaltungsdaten bei der Evaluation von Baugenehmigungsverfahren).

schwierig genug, so stellen sich die methodischen Probleme bei der Evaluation von Verwaltungspolitik auch insoweit noch verschärfter (vgl. Pollitt/ Bouckaert 2001: 170 f., ausführlicher auch Wollmann 2000a: 202 ff.).

- Anwendungsbeispiele für kontrollierte Experimente sind in der Verwaltungspolitik bislang kaum anzutreffen⁸.
- Auch quasi-experimentelles Vorgehen ebenso wie „kontrollierte“ Zeitreihen-Designs wurden bislang – wegen der erhöhten Schwierigkeiten, die „ceteris paribus“-Voraussetzungen in einem in methodisch befriedigender Weise herzustellen – kaum praktiziert.
- Für quantitative Untersuchungsverfahren erweist sich die Zahl der Untersuchungseinheiten (N) in der Regel als zu klein. Ähnliches gilt für die Durchführung methodisch disziplinierter „multi-cases“-Untersuchungen.
- In der Auswahl der Fälle weisen die Untersuchungen vielfach darin einen Bias auf, dass von vornherein „modernisierungspositive“ Fälle ausgewählt werden und damit die Ergebnisse „positiv verzerrt“ sind. Dies gilt nicht zuletzt für die „best practice“-Untersuchungen.
- Für die Durchführung von Vorher-Nachher-Untersuchungen zeigt sich, dass für die Sicherung der erforderlichen „Vorher-Daten“ kaum, jedenfalls unzureichend Sorge getragen worden ist. Von den (verhältnismäßig) wenigen Evaluationsuntersuchungen, die (z.B. in Australien und Neuseeland) auf eine umfassendere Evaluation von Verwaltungsreformprogrammen angelegt waren, wurden die meisten in deren Spätphase begonnen, so daß auf verlässliche „Vorher“-Daten und -Informationen kaum zurückgegriffen werden konnte (vgl. Pollitt/Bouckaert 2003)

1.2.3. Quasi-Evaluation

Angesichts der zahlreichen methodischen Probleme, mit denen die Evaluation von verwaltpolitischen Strategien und Maßnahmen konfrontiert ist, plädierte unlängst *Jean-Claude Thoenig* für das Konzept einer „Quasi-Evaluation“ (vgl. Thoenig 2003). Diese soll sich in erster Linie auf die eher deskriptive Erfassung der verwaltpolitisch intendierten Veränderungen konzentrieren und beschränken und auf den der „klassischen“ (ex-post) Evaluation eigentümlichen methodischen Aufwand und hierbei insbesondere auch auf den Anspruch, sozialwissenschaftlich valide „Kausalaussagen“ zu gewinnen, verzichten. Innerhalb der bekannten Varianten von Evaluation (ex ante-Evaluation, „ongoing“/formative Evaluati-

⁸ vgl. Pollitt/Bouckaert 2003: „it is almost unknown for management reforms to be subjected to true experimental procedures“

on/Monitoring und ex-post/summative Evaluation) ist eine solche „Quasi-Evaluation“ am ehesten als ein Fall „ongoing“/„formativen Evaluation“ und des „Monitoring“ zu identifizieren, indem sie maßgeblich auf eine ständige Interaktion zwischen den für die Evaluation initiiierenden und für sie verantwortlichen politischen und administrativen Akteuren und den mit ihrer Durchführung beauftragten externen Evaluatoren angelegt ist.

Die Vorteile dieses „pragmatischen“ Evaluationsansatzes werden von *Thoenig* in zweierlei gesehen:

- Durch die konzeptionelle und methodische „Verschlankung“ der Evaluation werden deren Initiierung durch Politik und Verwaltung und ihre Durchführung politisch, zeitlich und auch finanziell überhaupt erst ermöglicht⁹.
- Dadurch, daß dieses Konzept auf die Interaktion zwischen den für die Reformmaßnahme (und deren Evaluation) verantwortlichen politischen und administrativen Akteuren und den Evaluatoren zugeschnitten ist, eröffnet und fördert sie zwischen diesen einen veraltungspolitischen „Dialog“ als „a gradual learning process in successive stages that fosters an information culture“ (Thoenig 2003)..

Dabei kann Quasi-Evaluation durchaus als ein *Einstieg* und eine „vertrauensbildende Maßnahme“ auf dem Weg zu umfassenderen, auf das klassische Repertoire sozialwissenschaftlicher Methoden gestützten (ex-post) Evaluationen durch unabhängige externer Evaluationsforschung gesehen werden, auf die – auch aus der Sicht *Thoenigs* – grundsätzlich und mittelfristig nicht verzichtet werden sollte.

1.3. Zeitpunkt und Modalität der Evaluation

Innerhalb der auf den Zeitpunkt der Evaluation im Politikzyklus abhebenden Unterscheidung zwischen *ex-ante*, *formativer/ongoing* und *ex-post/summativer* Evaluation verdienen aus dem Blickwinkel der Verwaltungspolitik und angesichts des hohen Stellenwerts, den Ansätze des „performance managements“ in der NPM-Diskussion haben, jene Varianten von *ongoing* Evaluation besondere Aufmerksamkeit, die auf die laufende Erfassung von Kosten und Leistungen und von Wirkungen und deren umgehende Rückmeldung in den Entscheidungs- und Vollzugsprozeß gerichtet sind. Hierzu sind insbesondere die unterschiedlichen Varianten des *Controlling* zu rechnen (vgl. Richter 2001: 392ff.), die untersuchungsmethodisch überwiegend als deskriptives Monitoring von Veränderungen begriffen werden können und deren „kybernetische Rückmeldung“ vor allem über indikatoren- und kennziffern-gestützte Berichtssysteme erfolgen kann. Auch Ansätze des *benchmarking* können hierzu gezählt werden,

die – im Wege laufender oder periodischer, auf Kennziffern gestützter Datenerhebungen – dem intra- sowie inter-organisatorischen Leistungsvergleich und (Quasi-) Wettbewerb dienen (vgl. Nullmeier 2001: 95 ff., Schuster 2001).

Ansätze der *ex post-Evaluation* zielen in der Verwaltungspolitik darauf, die Wirkungen verwaltpolitischer Interventionen an deren Ende oder an markanten Zwischenstationen zu erfassen. Entsprechend dieser Fragestellung sind sie vielfach methodisch darauf zugeschnitten, nicht nur die Veränderungen und den Zielerreichungsgrad („Soll-Ist“-Vergleich, „Erfolgskontrolle“), sondern auch die sie (kausal) bedingenden Faktoren zu identifizieren („Wirkungskontrolle“). Der methodische Anspruch kann hierbei von Untersuchungen, die das verfügbare sozialwissenschaftliche Methodenrepertoire möglichst kompetent und vollständig nutzen („Evaluationsforschung“ im engeren Sinne), bis zu methodisch einfacheren Erhebungen (etwa in der Form der erwähnten „Quasi-Evaluation“ oder von Praxis- und Erfahrungsberichten) reichen.

1.4. Träger der Evaluation

Davon ausgehend, dass eine zweckmäßige Unterscheidung der Trägerschaft von Evaluation danach vorzunehmen ist, wer die Evaluation initiiert (und gegebenenfalls finanziert) und wer diese durchführt, kann von *interner* Evaluation zunächst (im engeren Sinn) dann gesprochen werden, wenn das Evaluationsverfahren in der operativen Verwaltungseinheit selbst initiiert, institutionalisiert und durchgeführt wird. Mit Bezug darauf ist vielfach von Formen der *Selbstevaluation* oder – in Übernahme der angelsächsischen Terminologie – von *in house-Evaluation* die Rede. Angesichts der strategischen Bedeutung, die in der neueren NPM- und NSM-inspirierten Verwaltungsreformdiskussion den Verfahren des verwaltungsinternen *Controlling* beigemessen wird, verdienen Varianten der internen Evaluation besondere Aufmerksamkeit. Darüber hinaus sollen hier solche Evaluationsansätze, die im Kontext der Verwaltung und Exekutive insgesamt initiiert bzw. durchgeführt werden (z.B. als „Erfolgskontrolle“ von Pilotvorhaben durch ein Ministerium), als *interne* Evaluation (im weiteren Sinne) verstanden werden.

Demgegenüber kann von *externer Evaluation* dann gesprochen werden, wenn diese Evaluation von einer (formal unabhängigen) Forschungseinrichtung durchgeführt wird, sei es als Untersuchung, mit der diese von der zu evaluierenden Verwaltungseinheit, von einer anderen Institution oder Forschungsförderungseinrichtung beauftragt worden ist („Auftragsforschung“, *contractual research*, vgl. Wollmann 2002b), sei es als (eigenfinanziertes) wissenschaftliches Projekt. Zur Abgrenzung von der *internen* Evaluation“ (im engeren wie weiteren Sinne) soll hier von *externer* Evaluation (im weiteren Sinne) auch in Fällen die Rede sein, in denen die Initiierung, Durchführung oder Beauftragung der Evaluation durch Institutionen

⁹ Thoenig merkt hierzu trocken an: „This is a perfectly good way to start. There is no surer way of stifling evaluation at birth than to confine it in the ghetto of methodology“, Thoenig 2003

außerhalb der Verwaltung/Exekutive (z.B. durch Parlament, Rechnungshof, unabhängige Reformkommission usw.), also außerhalb des Bannkreises der zu evaluierenden Verwaltungseinheit erfolgt.

Diese (in ihrer analytischen Trennschärfe allerdings begrenzte) Unterscheidung zwischen Ansätzen und Varianten *interner*“ und *externer* Evaluation soll der Gliederung der nachfolgenden Darstellung dienen.

2. Verwaltungspolitik und Evaluation in Deutschland

Die Evaluationsfunktion als Instrument des Politik- und Verwaltungshandelns wurde in der deutschen Verwaltungspolitik erstmals in der „Reformwelle“ der 1960er und 1970er Jahren lebhaft diskutiert und in Ansätzen praktiziert. In dem der „Planungsdiskussion“ zugrunde liegenden Politikmodell und -zyklus (Planung, Implementation und Ergebnis) wurde der Evaluation die strategische Aufgabe beigemessen, die Ergebnisse und Wirkungen des Politik- und Verwaltungshandelns zu ermitteln und diese „rückzumelden“. Der neue Informations-, Analyse- und Evaluationsfokus der Verwaltungspolitik fand zum einen in der Installierung von von *verwaltungsinternen* Informations-, Indikatoren- und Berichtssystemen seinen Ausdruck, worin sich insbesondere die kommunale Ebene als Schrittmacher erwies (vgl. Hellstern/ Wollmann 1984b). Zum andern zeigte sich dies in der raschen Expansion der von Bundes- und Landesministerien in Auftrag gegebenen und finanzierten *externen* Politik- und Programm-Evaluation; in dieser Phase der Evaluation erwies sich die Bundesrepublik – neben *Schweden* – als eine Vorreiterin unter den westeuropäischen Ländern (vgl. etwa Derlien 1976, Levine 1981, Hellstern/Wollmann 1984a). Allerdings ging es hierbei fast ausschließlich um Evaluation und Evaluationsforschung zu „substanziellen“ Politiken und politischen Programmen („ProgrammEvaluation“, „Programmforschung“, vgl. Derlien 1981, z.B. zu Städtebauförderung, Wohnungspolitik usw.), während (externe) Evaluationen von verwaltungspolitischen Maßnahmen weitgehend fehlten¹⁰.

In der verwaltungspolitischen Veränderungswelle, die die deutsche Politik- und Verwaltungswelt seit den frühen 1990er Jahren unter dem konzeptionellen Banner des *Neuen Steuerungsmodells* (als der deutschen Variante und Spielart des *New Public Management*)

ergriffen hat, weist das zugrunde liegende „Managementmodell und -zyklus“ (Zielsetzung, Implementation, Ergebnis) der Evaluationsfunktion wiederum eine zentrale (und wegen ihrer

¹⁰ Hierfür ist bezeichnend, daß die kommunalen Gebietsreformen, die in den 1960er und 1970er Jahren in den Bundesländern durchgeführt wurden und das einschneidendste verwaltungspolitische Reformprojekt dieser Periode bildete, nicht etwa von den Ländern (extern) evaluiert, sondern von einem (von der VW-Stiftung finanzierten) Verbund von Wissenschaftlern (unter evaluativer Fragestellung) untersucht wurden. Für eine Zusammenfassung vgl. vgl. Thieme/ Prittwitz 1981.

strategischen Einbindung in eine umfassende Steuerungsfunktion noch gesteigerte) Rolle zu.¹¹ (Auf die Gründe, weshalb sich die deutsche verwaltungspolitische Reformdebatte erst in den frühen 1990er Jahren an die seit den frühen 1980er Jahren international dominierende NPM-Diskussion angeschlossen hat, soll und kann an dieser Stelle nicht eingegangen werden, ausführlich hierzu etwa Wollmann 1996.)

2.1. Ansätze interner Evaluation

„Performanz-Management“, Controlling

In dem *Neuen Steuerungsmodell*, das die *KGSt* (in Anlehnung an *New Public Management* und an das „Tilburger Modell“) seit 1991 formulierte und propagierte (vgl. bahnbrechend Banner 1991) und das sich zunächst unter den (west-)deutschen Kommunen „wie ein Buschfeuer“ (Reichard 1994: 7) verbreitete, spielte die Einführung von Kosten- und Leistungsrechnungen, von (output)-Budgetierung, Controlling und Berichtswesen – und damit Ansätze interner Evaluation und des – in der internationalen Terminologie – *Performance Management* eine zentrale strategische Rolle. Zwar hat der Eifer, mit dem sich die Kommunen, oft unter dem Eindruck externer Beratung, auf die Ausarbeitung von „Produktdefinitionen“ und „-katalogen“ (als vermeintlicher Schlüsselgrößen des *Neuen Steuerungsmodells*) geworfen und konzentriert haben, vielfach als ein kosten- und zeitraubender und reformpolitisch letztlich wenig kaum ergiebiger Umweg, wenn nicht als eine Sackgasse erwiesen.¹² Jedoch haben Ansätze und Verfahren (indikatoren-gestützter) Evaluation und Rückmeldung in (meist konzeptionell und instrumentell abgeschwächten und „weicheren“) Varianten von Kosten- und Leistungsrechnung, Controlling und Berichtswesen inzwischen durchaus verbreitet Eingang in die Modernisierungspraxis Kommunalverwaltungen gefunden. Insoweit hat sich „Controlling“ als systematische verwaltungsinterne „Rückmeldeschleife“ zum „Schlüsselkonzept der aktuellen Politik- und Verwaltungsreform“ (Wollmann 2003a: 345) entwickelt. Allerdings ist hinzuzufügen, daß in der Praxis der deutschen Kommunalverwaltung bislang von einem funktionsfähigen Controlling kaum die Rede sein dürfte (Bogumil)¹³

Inzwischen sind *Kosten- und Leistungsrechnung* und *Controlling* auch in der Verwaltungspolitik der *Bundesländer*¹⁴ und zuletzt auch des *Bundes* in der Absicht aufgegriffen worden, „einen Prozeß zu initiieren, der an strategischen Zielen ausgerichtete Maßnahmen festlegt,

¹¹ Zur Analyse und Interpretation der zwei „Wellen“ („lange Wellen“) der jüngeren deutschen Verwaltungspolitik („Planungsdiskussion“ und „Neuen Steuerungsmodell“) vgl. Wollmann 2000b: 725 ff. ff. ; zu Analogie und Unterschieden zwischen den diesen „Wellen“ zugrunde liegenden Schlüsselkonzepten „Politik-, bzw. „Managementzyklus“ vgl. Wollmann 1994.

¹² Vgl. Reichard 1998: 85 ff. sowie zuletzt bemerkenswert (selbst-)kritisch Banner 2001: 287

¹³ vgl. J. Bogumil im DeGEval-Podiumsgespräch vom 5.10.2001, zitiert in Jacoby 2002, Tagungsbericht, S. 199 ff.

¹⁴ Zum Beispiel des Landes *Baden-Württemberg* vgl. Innenministerium Baden-Württemberg 1999: 128 ff.

diese auf ihre Wirksamkeit überwacht und durch einen permanenten 'Soll-Ist-Vergleich' zu weiteren Verbesserungen führt¹⁵.

Vergleichsringe

Im Laufe der frühen 1990er Jahre haben auch Ansätze des *benchmarking* in die deutsche Verwaltungswelt Eingang gefunden. Wiederum nahm zunächst die kommunale Ebene diese verwaltungspolitische Spur auf. Die Initiative ging zuerst von der *Bertelsmann-Stiftung* aus, die 1990/1991 mit einem Projekt „Grundlagen einer leistungsfähigen Kommunalverwaltung“ die Bildung von *interkommunalen Vergleichsringe* als ein Verfahren kommunaler SelbstEvaluation und (interkommunalen) Quasi-Wettbewerbs propagierte und finanziell förderte. 1997 waren bundesweit 150 Städte am kommunalen Leistungsvergleich der Stiftung beteiligt (vgl. Schuster 2001a, 2001b). An diese Vorarbeiten wurde Ende 1996 von der *KGSt* ein *IKO-Netzwerk* ins Leben gerufen. Dieses soll als „internes Informationssystem der Kommunen“ fungieren, die sich zu themen- bzw. arbeitsbereichsspezifischen und zum Teil regional organisierten Vergleichsringen zusammenschließen, gemeinsam aussagefähige Kosten- und Leistungsindikatoren („Kennziffern“) ausarbeiten und sowohl für die interne Evaluation als auch für den interkommunalen Leistungsvergleich nutzen (vgl. Kuhlmann 2002, 2003). Ende 2001 bestanden bundesweit 62 regional organisierte thematische Vergleichsringe, an denen (Ende 2001) insgesamt 672 Kommunen mitwirken¹⁶. Auch wenn mithin interkommunaler Leistungsvergleiche in der kommunalen Verwaltungspolitik inzwischen einen gefestigten Platz haben dürften, weisen sie darin eine erhebliche Schwäche auf, daß ihre Ergebnisse bislang weitgehend nur *verwaltungsintern* (für verwaltungsinterne Steuerung und Kontrolle) vorgesehen und verwendet werden, während sie den kommunalen Parlamenten und der kommunalen Öffentlichkeit nicht oder eingeschränkt, jedenfalls keineswegs verbindlich zugänglich sind. Aber erst durch eine verbindliche Öffentlichkeit der durch die Leistungsvergleiche ermöglichten Verwaltungstransparenz könnte verhindert werden, daß sich die kommunalen Leistungsvergleiche auf eine verwaltungsinterne SelbstEvaluation reduzieren, und könnten ein wirklicher Leistungswettbewerb und eine „Kontrolle“ durch die Bürger gesichert werden.

Parallel zu den Aktivitäten des *IKO-Netzes* war beispielsweise in *Baden-Württemberg* seit 1998 eine kommunale Arbeitsgemeinschaft damit befaßt, in Facharbeitskreisen mit Praktikern aus Städten, Gemeinden, Kreisen und Verbänden den (im Jahr 1996 vorgelegten) kommunalen Produktplan fortzuschreiben und Definitionen für steuerungsrelevante produktbezogene

¹⁵ Bundesregierung 2002: 18. Das Verfahren der Kosten- und Leistungsrechnung erfaßt derzeit (Stand Ende 2001) 89 Prozent des Bundespersonals – gegenüber 8,5 Prozent im Jahr 1998. ebda., S. 18.

¹⁶ vgl. *KGSt IKO-Netz*, Aktuelle Übersicht der *IKO-Netz Vergleichsringe*, www.iko-netz.de Vgl. auch Auswertung und Übersicht bei Kuhlmann 2002: 8 Tabelle 1

Kennzahlen zu erarbeiten¹⁷. Auch in der Verwaltungen der Bundesländer¹⁸ und des Bundes¹⁹ findet *benchmarking* als neues Steuerungsinstrument wachsende Aufmerksamkeit und Anwendung. (ausführlich zur Indikatoren- und Kennzifferngestützten Leistungsmessung kommunalen/staatlichen Handelns vgl. Kuhlmann/Bogumil/Wollmann 2003).

Pilotvorhaben und deren (interne) Evaluation

Nachdem Ansätze „experimenteller Politik“ in den 1960er und 1970er Jahren in einer ganzen Reihe von („substanziellen“) Politikfeldern (Schul-, Verkehrspolitik usw.) eine – auch im internationalen Vergleich – beachtliche Blüte erlebt hatten (vgl. Hellstern/Wollmann 1983), in der Folgezeit aber weitgehend verebbt waren, erfuhren sie in der Verwaltungspolitik der 1990er Jahre eine gewisse Wiederbelebung; ihre Evaluation erfolgte indessen weitgehend in verwaltungsinternen Verfahren und Formen.

Einen Anstoß hier gab in fast allen Bundesländern die Einführung von Experimentierklauseln, als 1994 – offenbar durch das Vorbild der in den skandinavischen Ländern in den 1980er Jahren initiierten „free communes“-Experimente (zu letzteren vgl. Baldersheim 1996: 293 ff.), die einschlägigen Landesgesetze, insbesondere die Gemeinde- und Kreisordnungen, in der Absicht novelliert wurden, den Gemeinden die rechtliche Möglichkeit zu geben, eine auf zwei Jahre befristete Freistellung von bestimmten kommunal-, insbesondere haushaltsrechtlichen Vorschriften zu beantragen und innerhalb der hierdurch eröffneten Spielräume neue kommunale Handlungsmuster zu erproben (ausführlicher. Wollmann 2000a: 217 mit Nachweisen). Allerdings wurde auf methodisch anspruchsvolle (externe) Evaluation dieser Pilotvorhaben durchweg verzichtet (kritisch Klages 1996: 7 ff.) und begnügten sich die Innenministerien der Länder mit einer methodisch eher „handgestrickten“ Auswertung verwaltungsinterner Berichtsbögen (vgl. Wollmann 2000: 218 mit Nachweisen).

Dieser Ansatz, reformpolitische Veränderungen im Wege von „Pilotvorhaben“ zu erproben und (im Wege verwaltungsinterner Erhebungen) zu überprüfen, wurde in einigen Ländern in der Folgezeit weitergeführt. So wurden beispielsweise in *Baden-Württemberg* 1995 „Pilotversuche mit dezentraler Budgetverantwortung ... kombiniert mit der Einführung von Kosten-

¹⁷ Zu dem im Juni 2001 veröffentlichten (neuen) „Kommunalen Produktplan“ vgl. Innenministerium Baden-Württemberg 2001)

¹⁸ Zur Entwicklung in *Baden-Württemberg* vgl. Innenministerium Baden-Württemberg 1999: 122, in *Nordrhein-Westfalen* vgl. Innenministerium NRW 2001: 16 f.

¹⁹ Seit Sommer 2000 werden in 11 Bundesministerien und 8 Bundesbehörden drei Benchmarking-Ringe zu drei Themen („Produkten“, nämlich Personalwesen, Informationstechnik und Zuwendungswesen) aufgebaut, vgl. Bundesregierung 2002: 21 sowie <http://www.staat-modern.de/projekte/index.html>

und Leistungsrechnung sowie kennzahlen- und kostenorientierten Führungssystemen (Controlling)“ in 30 „Pilot“-Behörden eingeleitet²⁰.

(Verwaltungsinterne) „Erfolgskontrolle“

Des Weiteren ist zunehmend die Bereitschaft insbesondere in der Verwaltungspolitik der *Bundesländer* zu verzeichnen, zu den von ihnen eingeleiteten verwaltungspolitischen Reformvorhaben Berichte und „Erfolgskontrolle“ vorzubereiten und vorzulegen, die methodisch überwiegend auf (verwaltungsinterne) Umfragen und Erfahrungsberichte gestützt sind. Beispielhaft sei hier kurz auf drei Bundesländer eingegangen.

In *Baden-Württemberg* sieht die Landesregierung (seit den späten 1980er Jahren) „Berichtspflichten“ über die Durchführung (und Zielerreichung) maßgeblicher verwaltungspolitischer Reformprojekte vor (vgl. Seyfried 2002)²¹. So führte das Landesinnenministerium „zur Evaluation (der Eingliederung der unteren Sonderbehörden, H.W.) eine Umfrage über die Regierungspräsidien bei allen Landratsämtern und Bürgermeisterämtern der Stadtkreise“ durch (vgl. Innenministerium BW 1999: 80; Seyfried 2001)

In *Nordrhein-Westfalen* sei auf „erste Umfrage (des Innenministeriums) zum Stand der Binnenmodernisierung bei den obersten Landesbehörden“ (vgl. Behrens 2002: 12 ff.) und auf „Evaluationsseminar“ verwiesen, das am 17. und 18. 1. 2001 unter der Ägide des Innenministeriums zu der am 1.1.2001 in Kraft getretenen Umstrukturierung der Bezirks- und Landesoberbehörden durchgeführt wurde und in dem eine auf Praktiker- und Erfahrungsberichte gestützte erste (evaluierende) Einschätzung unternommen wurde²².

Einen bemerkenswert umfassenden Bericht zum „Stand der Umsetzung“ des (Berliner) „Verwaltungsgrundsatzgesetzes“ (als Gesetzesprogramms des 1994 eingeleiteten Reformprojekts)²³ legte in *Berlin* der Senat im Oktober 2000 vor (vgl. Senat von Berlin 2000). Im wesentlichen auf „standardisierte Abfragen“ bei den Dienststellen der Senats- und Bezirksebene anhand maßgeblicher Reformziele (Neugestaltung der Binnenstruktur, Verbesserung der Bür-

²⁰ Vgl. Innenministerium BW 1999: 123 ff. Nach zweieinhalbjähriger Laufzeit führte das Landesfinanzministerium im Wege einer internen Untersuchung eine „Evaluation der Nutzungspotentiale, die sich bis dahin bei den Pilotämtern ergeben hatten“, ebda.

²¹ vgl. zuletzt das (vom Ministerrat des Landes BW am 9.4.2002 beschlossene) Arbeitsprogramm der Verwaltungsreform in der 13. Legislaturperiode des Landtags mit der Festlegung von „Berichten“ zu einigen abgeschlossenen bzw. laufenden Reformprojekten.

²² Vgl. VM-Impulse, 2002/Heft 1, S. 1, www.im.nrw.de/inn/132.htm. Das „Evaluationsseminar“ wurde unter der Federführung der *Deutschen Hochschule für Verwaltungswissenschaften Speyer* und unter der Moderation von *Hermann Hill* durchgeführt.

²³ Zu Konzept und Verlauf des Berliner Reformprojekts vgl. etwa Röber/Schröter 2001: 159 ff., Kuprath 2001: 215 ff., Wollmann 1998, Engelniederhammer/Köpp/Reichard/Röber/Wollmann 1999

gerorientierung usw., vgl. Senat von Berlin 2000: 3), gelangte in seinen Einzel- und Gesamtbetrachtungen zu überwiegend kritischen Einschätzungen²⁴.

Auf der *Bundesebene* wird im Rahmen des von der rot-grünen Bundesregierung Ende 1999 beschlossenen Reformprogramms „Moderner Staat - Moderne Verwaltung“ die Absicht verfolgt, „die Voraussetzungen einer modernen Erfolgskontrolle“ zu schaffen²⁵. Ein erstes Beispiel gibt die (mittels verwaltungsinterner Befragungen vorgenommene) „Evaluation“ der Reformkomponente „Ideenmanagementsystem“²⁶.

3.3. Ansätze externer Evaluation

Während die Nutzung verwaltungsinterner Evaluationsverfahren zu verwaltungspolitischen Maßnahmen in den letzten Jahren im Vordringen ist, sind externe (durch unabhängige Forschungseinrichtungen als Evaluationsforschung durchgeführte) Evaluationen noch immer nur vereinzelt anzutreffen.

Als rares Beispiel einer von der Exekutive/Verwaltung initiierte und (als Auftragsforschung vergebene) externe Evaluation ist die Untersuchung zu nennen, die im Zusammenhang damit durchgeführt wurde, daß die Landesregierung von *Brandenburg* zwischen 1995 und 1997 verwaltungspolitische Modellversuche in 8 Kommunen des Landes förderte und ein Forschungsinstitut mit der Begleitung und Evaluation dieser lokalen Projekte beauftragte (vgl. Maaß/Reichard 1998, Maß 1999).

Als Projekt methodisch anspruchsvoller Evaluationsforschung ist eine im Auftrag eines Bundesministeriums durchgeführte Untersuchung des öffentlichen Gesundheitsdienstes auf kommunaler Ebene hervorzuheben (vgl. Grunow/ Grunow-Luther 2000). Während externe Evaluationen, die von der Exekutive/Verwaltung selber veranlasst und als Auftragsforschung bearbeitet werden, bislang mithin bemerkenswert selten sind, kann immerhin auf eine recht breite Palette von Untersuchungen verwiesen werden, die – teils explizit, teils implizit – auf eine Evaluation von verwaltungspolitischen Maßnahmen gerichtet sind und von Institutionen außerhalb der Exekutive/Verwaltung initiiert, gefördert und/oder durchgeführt werden.

Hierzu sind zum einen Erhebungen zu rechnen (ausführlicher Wollmann 2000a: 218 ff.), die – überwiegend im Wege schriftlicher Umfragen und deren Auswertung – vom *Deutschen*

²⁴ Vgl. etwa aus der Gesamtbewertung: „Die Auswirkungen der Neustrukturierung werden vielfach nicht positiv beurteilt, weil die betriebswirtschaftlichen Instrumente – dies gilt besonders für den ministeriellen Bereich aber auch für bezirkliche Bereiche – noch nicht überzeugen. Die Kosten- und Leistungsrechnung bietet zzt. z.B. zur frist- und qualitätsgerechten Erledigung der Fachaufgaben, des Personaleinsatzes und der Organisation der Arbeitsabläufe noch keine nennenswerte Steuerungsunterstützung“ (Senat von Berlin 2000: 24).

²⁵ So Bundesinnenminister *Otto Schily* am 1.12.99 bei der Vorstellung des Reformprogramms.

²⁶ Vgl. „Idee 21 – Ideen für die Bundesverwaltung im 21. Jahrhundert“, www.staat-modern.de.

Städtetag (vgl. Grömig/Gruner 1998)²⁷, der *KGSt* (vgl. KGSt 1997) und vom *Deutschen Institut für Urbanistik* (Difu) (vgl. Mäding 1998) unternommen worden sind.

Zum andern ist auf der (anwendungsbezogenen) empirischen Politik- und Verwaltungsforschung zuzurechnende Untersuchungen zur Verwaltungsmodernisierung zu verweisen, die – meist mit eher impliziter denn expliziter Evaluationsabsicht – überwiegend von an Hochschulen oder an anderen (unabhängigen) Forschungseinrichtungen beheimateten Forschungsgruppen unternommen und zum erheblichen Teil von der (in Fragen der Verwaltungspolitik und -modernisierung besonders engagierten) *Hans-Böckler-Stiftung* gefördert worden sind. Zum Profil dieser Projektfamilie an dieser Stelle nur so viel²⁸.

- Die meisten einschlägigen Forschungsvorhaben verfolgten einen *kommunalen Fokus*²⁹, was darauf zurückzuführen ist, daß die jüngste Welle der Verwaltungsmodernisierung zunächst auf der kommunalen Ebene vordrang. Auf Landesebene ist bislang lediglich das 1994 im Stadtstaat Berlin in Gang gesetzte Reformprojekt zum Gegenstand Evaluationsorientierter Verwaltungsforschung gemacht worden³⁰.
- *International vergleichende* Untersuchungen wurden bislang insbesondere von *Frieder Naschold* angestoßen und von dessen Forschungsgruppe³¹ oder in einem von ihm initiierten Forschungsnetzwerk³² bearbeitet.
- In ihrer Fragestellung sind die vorliegenden empirischen Untersuchungen überwiegend mit dem Verlauf und Stand der verwaltungspolitischen Maßnahmen befaßt und mithin (in der hier verwendeten Typologie) in erster Linie als *InstitutionenEvaluation* einzustufen. Demgegenüber wurden Fragen nach den weiterreichenden Wirkungen von Verwaltungsmodernisierung, seien es „Performanz“ oder gar „outcomes“, bislang nur in wenigen Studien, etwa in Jaedicke/Thrun/Wollmann 2000³³, verfolgt.

²⁷ Für Sekundärauswertungen der DSt-Daten vgl. etwa Kuhlmann-Lorenz/ Wegrich 2001: 25ß ff-, zuletzt Kuhlmann 2002. Kritisch zur Verlässlichkeit der (Selbst-)Einschätzungen – und auskünfte von kommunaler Praktikern in Rahmen einer (schriftlichen) Befragung durch „ihren“ (Interessen-)Verband vgl. Grunow 1998: 5, Wollmann 2000: 219, unlängst auch Banner 2001: 281

²⁸ Vgl. Wollmann 2000: 219 ff.. mit detaillierteren Nachweisen von Gegenstand, Methoden usw. der einschlägigen Forschungsprojekte.. Vgl. auch die „Transparenzstudie“ von Jann/Reichard 1998

²⁹ vgl.. insbesondere Kißler/ Bogumil/ Greifenstein/ Wiechmann 1997 (Untersuchung des Modernisierungsprozesses in drei Stadtverwaltungen; Förderung durch Hans-Böckler-Stiftung) und Jaedicke/Thrun/Wollmann 1999 (Untersuchung von kommunalen Bauplanungs- und Baugenehmigungsämtern; Förderung durch Wüstenrot-Stiftung).

³⁰ Vgl Engelniederhammer/Köpp/Reichard/Röber/ Wollmann 1999, 2000 (Förderung durch das vom Berliner Senat finanzierte universitäre Programm Berlin-dienlicher Forschung).

³¹ Vgl. Naschold 1995 (Vergleichende Untersuchung der Modernisierungspolitik in mehreren OECD-Ländern, Förderung: Hans-Böckler-Stiftung) sowie Naschold/Oppen/Wegener 1997 (Vergleichende Untersuchung von innovativen Städten in mehreren Ländern, Förderung: Hans-Böckler-Stiftung).

³² Vgl. Naschold/Jann/Reichard 1999, Jann/Reichard 2001 (Vergleichende Untersuchung der Regierungs- und Verwaltungsmodernisierung auf der zentralen Regierungsebene in 8 Ländern, Förderung: Hans-Böckler-Stiftung).

³³ Es handelte sich um eine von der Wüstenrot-Stiftung ausdrücklich als Evaluation vergebene und geförderte Studie zur kommunalen Bau- und Planungsverwaltung

- In diesem Zusammenhang sei auch auf die methodisch wie inhaltlich bemerkenswerte Evaluationsstudie (als „Vergleich des Erreichten mit den ursprünglichen Zielen des Neuen Steuerungsmodells“) von Banner verwiesen, der sich auf unterschiedliche (überwiegend methodisch „qualitative“ und „weiche“) Informationsquellen stützt (vgl. Banner 2001: 285 f.).
-
- Schließlich sei ein verwaltungswissenschaftliches Forschungsprojekt genannt, in dem, wiederum von der Hans-Böckler-Stiftung gefördert, die erste systematische Evaluation des Verlaufs und der Ergebnisse der kommunalen Verwaltungsmodernisierung („10 Jahre NSM“) unternommen werden soll.³⁴

2.3. Evaluation und Verwaltungspolitik im internationalen Überblick

Im folgenden internationalen Überblick ist an dieser Stelle nur eine sehr geraffte und entsprechend lückenhafte Skizze möglich. Für eingehendere aktuelle Abhandlungen sei auf die Länderberichte in Wollmann 2003b sowie hierin auf den Überblicksaufsatz Wollmann 2003d verwiesen.

3.1. Ansätze interner Evaluation

Performance management

In einer wachsenden Zahl von OECD-Ländern ist die Einführung und Anwendung von Ansätzen des „*performance management*“ und „*management for results*“ zu beobachten, in denen es wesentlich darum geht, die Steuerung des Verwaltungshandelns durch Leistungs-/Output-Indikatoren mit deren laufender Kontrolle und Rückmeldung als Verfahren interner Evaluation zu verbinden (vgl. Naschold/ Jann/ Reichard 1999: 23 ff; Jann/ Reichard 2003).

Der bislang intensivste Einsatz eines auf *performance management* beruhenden Managementsystems ist bislang wohl in *Neuseeland* anzutreffen, wo zwischen den Ministerien und den ihnen nachgeordneten Organisationseinheiten (*agencies*) Zielvereinbarungen (*contracts*) über die von diesen zu erbringenden *outputs* abgeschlossen werden und diese gehalten sind, regelmäßige output-gestützte Berichte zu erstatten. Der Nachdruck, mit dem in *Neuseeland* das *performance management*-System und die mit ihm verbundene interne Evaluation verfolgt werden, wurde von einem kundigen Beobachter geradezu als „an obsessive concern for outputs and accountability“ bezeichnet (vgl. Halligan 2003).

³⁴ Das Projekt wird von Jörg Bogumil, Werner Jann, Leo Kißler, Christoph Reichard und Hellmut Wollmann geleitet. Laufzeit 2004-2005.

In den USA wurde durch den *Civil Service Reform Act* von 1978 in jedem Bundesministerium ein *Office of Inspector General* eingerichtet, das die Aufgabe hat, die (internen) Verwaltungstätigkeiten zu beobachten (*auditing*) und darüber zu berichten. Unter dem *Government Performance and Results Act* von 1992 wurde darüber hinaus ein umfassendes System für die laufende Beobachtung (*auditing*) der Tätigkeit der Regierungs- und Verwaltungseinheiten vorgeschrieben (vgl. Christensen/ Lagreid/ Wise 2003).

Ein weiteres anschauliches Beispiel liefert Schweden, wo insbesondere für die (seit langem bestehenden) unabhängigen Behörden (*Ämbetsverk*) ein *management by results*-System eingeführt worden ist (vgl. Christensen/ Lagreid/ Wise 2003), in dessen Rahmen diese Verwaltungseinheiten verpflichtet sind, jährliche *result*-gestützte Berichte zu erstatten. Hierbei machen Beobachter der „management by results“-Praxis in den skandinavischen Ländern darauf aufmerksam, daß hier die Funktion der indikatoren-gestützten Analyse- und Rückmelde-schleifen weniger in der Kontrolle der Einhaltung von Leistungsvereinbarungen (Kontrakten) als vielmehr (als Ausdruck der Verwaltungstradition und –kultur dieser Länder) darin zu sehen ist, Interaktions- und Lernprozesse zwischen den relevanten Akteuren zu fördern³⁵.

Auch in Norwegen wiederum treten Finanzministerium und der Rechnungshof in den letzten Jahre immer stärker als Advokaten einer Performanz-Evaluation („performance auditing“) hervor (vgl., Christensen/ Laegreid/ Wise 2003).

In der Schweiz ist auf das 1997 in vier Ämtern der Bundesverwaltung eingeführte Reformkonzept Konzept zu verweisen, das sich an NPM-Vorstellungen orientiert (*FLAG = Führung durch Leistungsauftrag und Globalbudget*, vgl. Ritz 1999: 39 ff.) und in dem interne Informationsschleifen und Berichtssysteme eine wesentliche Rolle spielen.

Benchmarking

Auf die Nutzung von Verfahren interner Evaluation als Vehikel und Stimulans für intra- und insbesondere inter-organisatorischen (Quasi) Wettbewerb zielen die benchmarking-Ansätze, deren Vorbilder im privaten Sektor gesehen wurden und in der Modernisierungspraxis bislang insbesondere auf der kommunalen Ebene verfolgt worden sind. Beispiele hierfür liefert vor allem Großbritannien, wo in einem ersten Schritt 1992 – im Rahmen der von Konservativen Regierung unter John Major ergriffenen „citizen charter“-Initiative – die Kommunen durch Gesetz verpflichtet worden sind, ihre Leistungen nach einem bestimmten Kennzahlensystem („Performance Indicators“), das landeseinheitlich von der Audit Commission ausgearbeitet und vorgegeben wird, nachzuweisen und zu veröffentlichen. Unter dem „Best value regime“, das die New Labour Government unter Tony Blair eingeführt hat und in dem die

Audit Commission als „Wachhund“ der Regierung eine Schlüsselrolle spielt, hat die indikatorengestützte Leistungskontrolle der Kommunen und ihrer Verwaltungen einen „centrally prescriptive and potentially even more interventionist“ (Wilson/Game 2002: 337) Charakter angenommen. In den USA wurde Mitte der 1990er Jahre von der International City/County Management Association (IMCA) ein Projekt „Comparative Performance Management Program“ angestoßen, das zunächst 44 Städte und Counties umfaßte und auf einen indikatoren-gestützten Leistungsvergleich zielt (Adamaschek/Banner 197: 221 ff.).

3.2. Ansätze externer Evaluation

Verhältnismäßig frühe Beispiele von externer Evaluation der eingeleiteten Verwaltungsreform liefert Australien, wo 1993 die bis dahin rund zehnjährigen Reformaktivitäten Gegenstand einer ersten umfassenden Evaluation wurden. Diese wurden auf Veranlassung eines Parlamentsausschusses von einer *Task Force on Management Improvement*, einer quasi-unabhängigen Gruppe von Regierungsbeamten, unternommen (vgl. Halligan 2003). Wesentliche Anstöße für die lebhaften Evaluationsaktivitäten, die die Verwaltungsmodernisierung in Australien kennzeichnen, ebenso wie für die Verwaltungsmodernisierung insgesamt gehen vom Bundesfinanzministerium aus (vgl. Halligan 2003, der vom Finanzministerium als dem „evaluation champion“ spricht). Zu Umfang und Intensität dieser Evaluationstätigkeiten ist von einer Regierungskommission kritisch angemerkt worden, daß der australische Weg, den man geradezu als „evaluating everything that moves“ bezeichnen könne, „paints a picture of evaluation overkill“ (State Service Commission, zit. nach Halligan 2003).

Auch in Neuseeland, wo in den frühen 1980er Jahren eine durchgreifende Staats- und Verwaltungsmodernisierung (unter Anleitung einer theoretisch und politisch „radikalen“ NPM-Variante) in Gang gesetzt worden ist, wurden verhältnismäßig früh externe Evaluationen durchgeführt, zunächst 1991 von der *Steering Group for Review of State Sector Reform* und dann 1996 dadurch, daß der international renommierte Wissenschaftler und Experte Allan Schick mit einer umfassenden Evaluation beauftragt wurde (vgl. Halligan 2003); dieser kam zum Schluß, die neuseeländischen Reformen seien „more comprehensive and rigorous than those introduced in other countries, they have been neither complete nor perfect“ (zit. nach Halligan 2003)..

Ein weiteres Beispiel lebhafter Evaluationstätigkeiten in der Verwaltungsmodernisierung bieten die USA. In offensichtlicher Anknüpfung an die in die 1960er Jahren zurückreichende gesetzgeberische Praxis, bei der Verabschiedung einschlägiger Gesetze in diese Evaluations-

³⁵ Vgl. Naschold/Jann/Reichard 1999: 32: „Contracts in the Danish context are used in a „soft version“, signalling governance by contact rather than governance by contract in the hard version of the term“.

verpflichtungen aufzunehmen, schrieb der *Civil Service Reform Act* von 1978 entsprechende Evaluationen vor und enthielt zugleich die Ermächtigung zur Durchführung einschlägiger verwaltungspolitischer Experimente (vgl. Christensen/ Lagreid/ Wise 2003). Darüber hinaus wurde unter dem *Government Performance and Results Act* von 1992 ein umfassendes System zur Evaluation (*auditing*) von Verwaltungstätigkeiten etabliert. Außerdem gingen Evaluationsanstöße vom *General Accounting Office* (einem traditionellen Promotor von Evaluationen) und vom *Office of Management and Budget* aus.

In Schweden spielt der Rechnungshof auch in der Evaluation der Verwaltungsmodernisierung eine maßgebliche Rolle. Ein Beispiel für Schwedens traditionelles System unabhängiger Kommissionen und für die Evaluationskultur des Landes gibt die *Expertengruppe für den Öffentlichen Sektor*, die regelmäßige Berichte zur Leistungsstärke und Wirtschaftlichkeit der Regierungsaktivitäten vorlegt (vgl. Christensen/ Laegreid/ Wise 2003).

In beachtlichem Umfange sind neuerdings in der Schweiz externe Evaluationen auf allen Ebenen in Gang gekommen. Dies gilt für die Bundesebene, wo, wie bereits erwähnt, das an NPM-Vorstellungen orientierte Reformprojekt FLAG in vier Pilotämtern eingeführt worden ist und eine (unabhängige) Forschungseinrichtung mit einer externen Evaluation beauftragt worden ist (vgl. Ritz/ Thom/ Rieder 2002: 15; Thom/ Ritz 2000). Zum andern trifft dies in bemerkenswertem Maße auf die Verwaltungsreformpolitik der kantonalen Ebene zu, wo in der Mehrheit der 26 Kantone ebenfalls umfangreiche Reformprojekte initiiert und eine beträchtliche Zahl externer Evaluationen durchgeführt worden sind (vgl. Widmer 2002). Schließlich haben mehrere größere Städte Evaluationsuntersuchungen in Auftrag gegeben (vgl. Widmer 2002 mit Nachweisen).

Auffällig im internationalen Überblick ist, daß in Großbritannien zunächst kaum externe Evaluationen durchgeführt wurden, obgleich hier die Konservativen Regierungen unter *Margret Thatcher* und *John Major* in den 1980er Jahren tiefgreifende NPM-inspirierte Verwaltungsreformen (u.a. die Schaffung der *Executive Agencies* und die Einführung des *Compulsory Competitive Tendering, CCT*) eingeleitet hatten. Erst mit der neuen Labour Regierung unter *Tony Blair* hat sich die Zahl der Evaluationen merklich erhöht (vgl. Pollitt/ Bouckaert 2003).

Zusammenfassend ist freilich hervorzuheben, daß der Stand der externen Evaluation der Verwaltungsmodernisierung von internationalen Beobachtern als insgesamt nach wie vor defizitär eingeschätzt wird, jedenfalls gemessen an den hohen Erwartungen, die in der jüngsten Modernisierungswelle unter dem Vorzeichen des New Public Management in die Schaffung größtmöglicher Transparenz des Verwaltungshandelns gesetzt und geweckt worden ist. Insbesondere weist der gegenwärtige Stand der (externen) Evaluation eine auffällige analytische Schwäche darin auf, daß sich die vorliegenden einschlägigen in erster Linie mit dem Verlauf und den (Zwischen-)Ergebnissen der institutionellen Veränderungen befassen (also in unserer Terminologie als „Institutionen-Evaluation“ zu bezeichnen sind), während die weitergehen-

den „materiellen“ Wirkungen allenfalls in der Dimension der Veränderung administrativer Handlungsprozesse und -muster („improved processes“), also als „Performanz-Evaluation“ sowie als (im wesentlichen auf Kosteneinsparung verkürzte) „Effizienz“-Aussagen in den Blick gerückt werden³⁶. Hingegen bleibt die Wirkungsdimension der outputs und outcomes fast durchweg ausgeblendet. Sind (ausnahmsweise) umfassendere Evaluationsaussagen anzutreffen, sind sie eher generell und „weich“ formuliert³⁷.

4 Evaluation in der Verwaltungspolitik – eine „dritte Welle“ von Evaluation?

Die Entwicklung der Evaluation in der Verwaltungspolitik vermittelt national wie international ein ambivalentes und „asymmetrisches“ Bild.

Auf der einen Seite sind der Grundgedanke der Evaluation und evaluative Verfahren in dem Maße in den Mittelpunkt der Verwaltungspolitik gerückt, wie die systematische Erfassung von Kosten und Leistungen der Verwaltungstätigkeit als zentraler Baustein und Hebel von Verwaltungsreform als ständiger Aufgabe erkannt worden sind. Die bereits im Politikzyklus der Planungsdiskussion der 1960er und 1970er Jahre angelegte Ablauftriade von Zielformulierung, Implementation und Wirkungskontrolle ist im analogen Managementmodell der NPM-Diskussion vollends einen strategischen Stellenwert erlangt (vgl. Wollmann 1994). Controlling als permanente Rückmeldeschleife ist zum „Schlüsselkonzept in der aktuellen Politik- und Verwaltungsreform geworden“ (Wollmann 2003a: 345). Vor dem Hintergrunde des Vordringens verwaltungsinterner Evaluationsverfahren (Kosten-Leistungsrechnungen, Controlling usw.) und deren systematische Einbindung in die Verwaltungsführung könnte sogar von einer „dritten Welle“ der Evaluation gesprochen werden (vgl. Wollmann 2003c) – nachdem, vereinfacht gesprochen, die „erste Welle“ von Evaluation in den 1960er und 1970er Jahren, im wesentlichen in Form von externer Evaluation auf die Verbesserung der Zielerreichung von Politik und in der „zweiten Welle“ in den späten 1970er und 1980er Jahren, vorrangig eine Reduktion der Kosten von Politik erreicht werden sollte (zum Konzept der früheren „Wellen“ von Evaluation vgl. Wagner/Wollmann 1986, Derlien 1990).

³⁶ So berichtet die neuseeländische State Services Commission „evidence of a good deal of evaluation activity, (but) most tends to be focused on process, implementation, and efficiency, and used for internal management purposes... Evaluation against outcomes is considered too hard“, zit. nach Halligan 2003.

³⁷ vgl. die (erste) umfassende Evaluation, die in Neuseeland von einer dort eingesetzten Steering Group 1991 durchgeführt wurde: „that there were ‘substantial benefits’ in terms of system capability“, zit. nach Politt/Bouckaert 2001: 183).

Auf der anderen Seite steht sowohl im nationalen wie im internationalen Kontext systematische Evaluationsstudien zu Verlauf und Ergebnis der Verwaltungsreformen durch unabhängige externe Evaluationsforschung nach wie vor weitgehend aus. Dieses Defizit ist umso auffälliger, als die Schaffung von Transparenz der Wirkungen, Kosten, Nutzen usw. des Verwaltungshandelns eine Kernbotschaft der NPM-inspirierten Verwaltungsmodernisierung ist, die sich damit zuvörderst auch selber zum Gegenstand von Evaluationsuntersuchungen machen sollte.

Die Gründe für dieses Fehlen unabhängiger externer Evaluationsforschung im Feld der Verwaltungspolitik liegen auf der Hand. Während insbesondere die Verwaltungsführung in den Verwaltungen ein deutliches Eigeninteresse daran hat, die Steuerungsfähigkeit und Kostentransparenz der Verwaltung durch die Einführung und Nutzung *verwaltungsinthener* evaluativer Verfahren zu verstärken und zu festigen, ist den Politik- und Verwaltungsspitzen in der Regel kaum daran gelegen, sich durch die Vergabe externer Evaluierungsuntersuchungen in ihre verwaltungs- und institutionenpolitischen Karten sehen zu lassen und „verwaltungs-fremden“ unabhängigen Evaluatoren Zugang und Einblick in das „Innenleben“ der Verwaltung zu gewähren.

Hinzu kommt, daß – jedenfalls in Deutschland - die *Parlamente* bislang wenig Interesse daran gezeigt haben, an der Verwaltungspolitik gestaltend mitzuwirken oder zumindest (innerhalb ihrer Kontrollfunktion gegenüber der Exekutive) auf deren Evaluierung zu drängen.. Der niedrige politisch-parlamentarische Aufmerksamkeitspegel für Verwaltungspolitik dürfte wesentlich darauf zurückzuführen sein, daß diese – im Gegensatz zu „substanziellen“ Politikfeldern (Sozial-, Umweltpolitik usw.), die in politisch-gesellschaftlichen Interessen- und Klientelgruppen verwurzelt sind – über eine solche politisch-gesellschaftliche Verankerung und Einbettung nicht verfügt und deshalb aus der Sicht der Parteien und Parlamentarier keine wähler- und wahlbezogene Relevanz besitzt..

Weiterhin haben sich auch die *Rechnungshöfe* in Deutschland in geringerem Maße für Verwaltungspolitik und deren Evaluierung interessiert gezeigt als dies in anderen Ländern (z.B. in den USA und in Schweden, vgl. Wollmann 2003d) zu beobachten ist. Zwar haben die Rechnungshöfe des Bundes und der Länder seit den 1970er Jahren von der traditionellen „titelbezogenen Ausgabenkontrolle“ und „Ordnungsmäßigkeitsprüfung“ auf eine „politische Aufgabenkontrolle“ und „Erfolgskontrolle“³⁸ zubewegt. Jedoch ging es bei dieser „evaluativen“ Neuorientierung der Prüfaktivitäten der Rechnungshöfe bislang fast ausschließlich um „substanzielle“ Politiken (Umweltpolitik, Subventionsprogramme usw.) und Investitionsvor-

³⁸ vgl. Dieckmann 2002 in Jacoby 2002: 116 ff. Laut Dieckmann sind „beispielsweise die Jahresberichte des Hamburger Rechnungshofs der letzten Jahre... bis zu 50% der Prüfungen ... unter Erfolgskontrolle einzuordnen“.

haben. Freilich gibt es Hinweise darauf, daß Verwaltungspolitik neuerdings zunehmend in das Blickfeld und in die Prüftätigkeit der Rechnungshöfe gerät³⁹.

In der dringend gebotenen, jedoch bislang weitgehenden ausgebliebenen externen unabhängigen Evaluation von Verwaltungspolitik sollte es vor allem um die folgenden Evaluationsfragen gehen:

- Zum einen sollte der bisherige Fokus auf *Institutionenevaluation*, also auf die beobachtbaren *institutionellen und instrumentellen* Veränderungen überwunden und um die Fragen nach dem veränderten Verwaltungshandeln („Performanz“) und nach den „Ergebnissen“ („outputs, outcomes“) – auch und gerade aus der Sicht der Bürger als den letztlich Adressaten (und Nutznießern) der Verwaltungspolitik – erweitern.

Zum andern sollten die Modernisierungsmaßnahmen selber zum Gegenstand von „Nutzen-Kosten-Analysen“ gemacht werden. Hierbei ist daran zu erinnern, daß seit den frühen 1990er Jahren zwar erhebliche finanzielle und personelle Aufwendungen in die Vorbereitung⁴⁰ und Durchführung der Reformprojekte auf der Kommunal-, Landes- und Bundesebene, nicht zuletzt in die Honorierung von Unternehmensberatern, geflossen sind⁴¹, daß jedoch über diese Kosten auffallend wenig bekannt geworden ist⁴². Die verwaltungspolitisch dringend gebote-

³⁹ vgl. Dieckmann 2002 in in Jacoby 2002: 117, wonach die „Bewertung der Ziele und Umsetzung der einzelnen Elemente des Neuen Steuerungsmodells nach 5 Jahren Reformprozeß“ unlängst in die Prüffelder des Hamburger Rechnungshofes aufgenommen worden ist.

:

⁴⁰ Vgl. insbesondere die zahlreichen Organisationsuntersuchungen, die von den Bundesländern zur Vorbereitung verwaltungspolitischer Reformen (insbesondere an Unternehmensberatungsfirmen) vergeben und finanziert haben. Vgl. IM BW 1999: 184 ff. (mit einer Liste der vergebenen Organisationsuntersuchungen) sowie IM NRW 2001: 4, wonach im Auftrag des beim Finanzministerium angesiedelten „Arbeitsstabes Aufgabenkritik“ „externe Sachverständige in 88 Organisationsgutachten die gesamte Landesverwaltung unter organisatorischen, aufbau- und verfahrenskritischen Aspekten“ untersuchten. vgl. hierzu auch Bürsch/Müller 1999: 9

⁴¹ Über den Umfang der finanziellen Aufwendungen, die die Kommunen, die Länder und der Bund in die Vorbereitung und Durchführung der Verwaltungsmodernisierung (Honorierung von Organisationsgutachten, Beratungs- und Begleitungstätigkeit usw., aber auch eigene personelle und sächliche Aufwendungen) gesteckt haben, ist auffallend wenig bekannt. Die wenigen eher zufälligen Informationen lassen wohl kaum mehr als „die Spitze des Eisbergs“ erkennen. So wandte das Land Berlin allein 1994/1995 rund 20 Mio. DM für externe Beratungsleistungen von Consulting-Firmen auf. Hinzu kamen weitere 5,3 Mio. für Mitarbeiterqualifikation und 3,6 Mio. DM für IuK-Technik (vgl. Wollmann 1998: 220). Für externe Beratungsleistungen (einschließlich Produkt-Beratung) gab der ostdeutsche Landkreis Ludwigslust jährlich 120.000 DM, der (ostdeutsche) Landkreis Barnim einmalig 420.000 DM aus, vgl. Wegrich/Jaedicke/Lorenz/Wollmann 1996: 216, S. 216. Die Stadt Ludwigshafen entgelt die von ihr beauftragte Consulting-Firma mit 10.6 Mio. DM, vgl. Grunow 1998_1 FN 2

⁴² Öffentlich zugängliche Informationen z.B. hinsichtlich der aufgewendeten Honorarkosten (für Organisationsgutachten, konzeptionelle Vorbereitung und Durchführung von Reformprojekten) fehlen weithin. Hier nur einige beispielhafte (und eher zufällig verfügbare) Informationen, die kaum mehr als die „Spitze des Eisbergs“ erkennen lassen. So gab das Land Berlin allein 1994/1995 rund 20 Mio. DM für externe Beratungsleistungen von Consulting-Firmen aus. Hinzu kamen weitere 5,3 Mio. für Mitarbeiterqualifikation und 3,6 Mio. DM für IuK-Technik (vgl. Wollmann 1998: 220). Für externe Beratungsleistungen (einschließlich Produkt-Beratung) gab der ostdeutsche Landkreis Ludwigslust jährlich 120.000 DM, der (ostdeutsche) Landkreis Barnim einmalig 420.000 DM aus, vgl. Wegrich/Jaedicke/Lorenz/Wollmann 1996: 216, S. 216. Die

nen Kosten-Nutzen-Analysen sollen darauf gerichtet sein, einerseits eine möglichst vollständige Transparenz der „Kostenseite“ der Reformen, insbesondere ihrer „Transaktionskosten“ (z.B. die Hardware-, Software-, Beratungs-, Personalkosten und Daten-Erst-Beschaffungskosten ihrer Einführung und andererseits die „Erträge“ (Effizienz-, Performan-Gewinne für die Verwaltung, aber auch und gerade Leistungsverbesserung aus der Sicht des Bürgers) zu ermitteln. Nur eine solche analytisch sorgfältige „Kosten-Nutzen-Bilanz“ der Modernisierungsmaßnahmen erlaubt robuste Aussagen über deren „Erfolg oder Mißerfolg“.

Bibliographie

Adamaschek, B./Banner G. 1997, Der interkommunale Leistungsvergleich, in: M. Pröhl (Hrsg.), Internationale Strategien und Techniken für die Kommunalverwaltung der Zukunft, Gütersloh, S. 223 ff.

Baldersheim, H. 1996, Nordic Municipalities in Transition. From Free Communes to Free Choice?, in: Ch. Reichard/ H. Wollmann (Hrsg.), Kommunalverwaltung im Modernisierungsschub?, Basel usw. 1996, S. 293 ff.

Banner, G. 1991, Von der Behörde zum Dienstleistungsunternehmen. Die Kommunen brauchen ein neues Steuerungsmodell, in: VOP, 1991, Heft 1, S. 6 ff

Banner, G. 2001, Kommunale Verwaltungsmodernisierung: Wie erfolgreich waren die letzten zehn Jahre?, in: E. Schröter (Hrsg.), Empirische Policy- und Verwaltungsforschung, Opladen, S. 279-304

Böhret, C. 2001, Verwaltungspolitik als Führungsauftrag, in: Blanke, Bernhard u.a. (Hrsg.), Handbuch zur Verwaltungsreform, 2. Aufl., Opladen: Leske + Budrich, S. 43-48

Bogumil, J. 2002, Statement, in: in: Jacoby, K.P. 2002, Tagungsbericht, Möglichkeiten und Grenzen von Evaluation in der Verwaltungspolitik, in: Zeitschrift für Evaluation, Bd. 1, Heft 1, S. 119 ff.

Bundesregierung 2002, Moderner Staat, moderne Verwaltung, Bilanz 2002, Bonn 2002

Bürsch M. /Müller, B. 1999, Verwaltungsreformen in den deutschen Bundesländern, Bonn-Bad Godesberg 1999

Christensen, T./ Laegreid, P./ Wise, L. 2003, Evaluating public management reforms in central government: Norway, Sweden and the United States of America, in: in: Wollmann, Hellmut (ed.) 2003, Evaluation in Public Sector Reforms, Concepts and Practice in International Perspective, Aldershot: Edward Elgar (forthcoming)

Christensen, T./Laegreid, P. 2001, A Transformative Perspective of Administrative Reforms, in: T. Christensen/ P. Laegreid (eds.), New Public Management, Aldershot etc.: Ashgate, p. 13-43

Derlien, H.-U. 1976, Die Erfolgskontrolle staatlicher Planung, Baden-Baden

Derlien, H.-U. 1981, Stand und Entwicklung der Programmforschung in der öffentlichen Verwaltung, in: ders. (Hrsg.), Programmforschung in der öffentlichen Verwaltung, Werkstattbericht 1 der Gesellschaft für Programmforschung, München, S. 9 ff.

Derlien, H.-U. 1990, Genesis and structure of evaluation efforts in comparative perspective, in: Ray, R. (ed.), Program Evaluation and the Management of Government, New Brunswick, pp. 147-176

Dieckmann, R. 2002, Statement, in: Jacoby, K.P. 2002, Tagungsbericht, Möglichkeiten und Grenzen von Evaluation in der Verwaltungspolitik, in: Zeitschrift für Evaluation, Bd. 1, Heft 1, S. 116 ff.

Engelniederhammer St./Köpp B./Reichard, Ch./Röber M./Wollmann, H. 1999, Berliner Verwaltung auf Modernisierungskurs. Berlin.1999

Engelniederhammer, St./ Köpp, B./ Reichadr, Ch./Röber, M./ Wollmann, H. 2000, Hauptweg und Nebenwege. Eine Zwischenbilanz zur Verwaltungsreform Berlin, Berlin; Berlin Verlag

Grömig E./ Gruner K.1998, Reform in den Rathäusern, in: Der Städtetag, 1998, Heft 8, S. 581 ff.

Grunow D. 1998, Lokale Verwaltungsmodernisierung „in progress“?, in: D. Grunow/ H. Wollmann (Hrsg.), Lokale Verwaltungsreform in Aktion, Basel usw., S. 1-25

Grunow D./Grunow-Lutter V. 2000, Der öffentliche Gesundheitsdienst im Modernisierungsprozess. Weinheim/München

H., Klages 1996, Methodenfragen des Verwaltungswettbewerbs, in: H. Hill/ H. Klages (Hrsg.), Jenseits der Experimentierklausel, Düsseldorf, S. 7 ff.

Halligan, J. 2003, Public-sector reform and evaluation in Australia and New Zealand, in; Wollmann, Hellmut (ed.) 2003, Evaluation in Public Sector Reforms, Concepts and Practice in International Perspective, Aldershot: Edward Elgar (forthcoming)

Hellstern, G.-M./Wollmann H. (Hrsg.) 1983, Experimentelle Politik – Reformstrohfeuer oder Lernstrategie, Opladen

Hellstern, G.-M./Wollmann, H. 1984a, Evaluation und Evaluationsforschung – ein Entwicklungsbericht, in: dies. (Hrsg.), Handbuch zur Evaluationsforschung. Bd. 1, Opladen 1984 , S. 17-93

Hellstern, G.-M./Wollmann H. (Hrsg.) 1984b, Evaluation und Erfolgskontrolle in Kommunalpolitik und -verwaltung, Opladen

Hofmann, J. 1993, Implizite Theorien in der Politik, Opladen: Westdeutscher Verlag

Innenministerium Baden-Württemberg 1999, Zwischenbilanz der Verwaltungsreform, Schriftenreihe der Stabsstelle für Verwaltungsreform, Bd. 21, Stuttgart 1999

Innenministerium Baden-Württemberg 2001, Produkte – Ziele – Kennzahlen. Kommunalen Produktplan Baden-Württemberg, (Schriftenreihe des Innenministeriums Baden-Württemberg, Heft 7), Stuttgart

Innenministerium Nordrhein-Westfalen, Bericht zum Stand der Verwaltungsmodernisierung, unveröff. Ms., Düsseldorf 21.12.2001

Jacoby, Klaus-Peter 2002, Tagungsbericht. Möglichkeiten und Grenzen von Evaluation in der Verwaltungspolitik, in; Zeitschrift für Evaluation, vol. 1, Heft 1, S. 115-126

Jaedicke W. / Thrun T./ Wollmann, H. 2000, Modernisierung der Kommunalverwaltung. Evaluationsstudie zur Verwaltungsmodernisierung im Bereich Planen, Bauen und Umwelt (Hrsg. Wüstenrot.-Stiftung) Stuttgart usw. 2000

Jann, W. /Reichard, Ch. 1998, Transparenzstudie. Wissenschaftliche Begleitforschung zu Verwaltungsreformprojekten in Deutschland (im Auftrag der Hans-Böckler-Stiftung), verv. Ms. KWI-Universität Potsdam

Jann, W./ Reichard, Ch. 2003, Evaluating best practice in central government modernization, in: Wollmann, Hellmut (ed.) 2003, Evaluation in Public Sector Reforms, Concepts and Practice in International Perspective, Aldershot: Edward Elgar (forthcoming)

KGSt, Mitgliederbefragung 1997, KGSt-Bericht 10/1998, Köln.

Kißler, L./ Bogumil J./Greifenstein R. / Wiechmann E. 1997, Moderne Zeiten im Rathaus? Berlin 1997

Knöpfel, P./ Bussmann, W.1997, Die öffentliche Politik als Evaluationsobjekt, in: W. Bussmann/ U. Klöti/ P. Knöpfel (Hrsg.) Einführung in die PolitikEvaluation, Basel/Frankfurt, S. 78 ff.

König K./Füchtner N., „Schlanker Staat“, Verwaltungsmodernisierung im Bund, Speyer 1998

Kuhlmann, S. 2002, Evaluation kommunaler Verwaltungsmodernisierung, unveröff. Ms., (Humboldt-Universität zu Berlin, Institut für Sozialwissenschaften) März 2002

Kuhlmann, S. 2003 Benchmarking auf dem Prüfstand: Kosten, Nutzen und Wirkungen interkommunaler Leistungsvergleiche in Deutschland, in: Verwaltungsarchiv, Bd. 94, Heft 1, S. 99-126

.

—

Kuhlmann, S./ Bogumil, Jörg/ Wollmann, Hellmut (Hgs.) 2003, Leistungsmessung kommunalen/staatlichen Handelns. Konzepte und Praxis, Opladen: Leske + Budrich (in Vorbereitung)

Kuprath, H. 2001, Das Berliner Modell einer (de-)zentralisierten Metropole aus der Sicht der Hauptverwaltung, in: Schröter, E. (Hrsg.), Moderne Verwaltung für moderne Metropolen, Opladen: Leske + Budrich, S. 215 - 222

Levine, R. A. 1981, Program Evaluation and Policy Analysis in Western Nations: An Overview, in: R.A. Levine/M.A. Solomon/G-M Hellstern/ H. Wollmann (eds.), Evaluation Research and Practice. Comparative and International Perspectives, Beverly Hills/London: Sage 1981, pp. 27-60

Levy, F./ Meltsner, A.J./ Wildavsky, A. 1974, Urban Outcomes, Berkeley: University of California Press

Maaß Ch. 1999, Brandenburger Kommunen auf dem Weg, Handbuch, Potsdam 1999

Maaß Ch./Reichard Ch. 1998, Von Konzepten zu wirklichen Veränderungen? Erfahrungen mit der Einführung des neuen Steuerungsmodells in Brandenburger Modellkommunen, in: D. Grunow/ H. Wollmann (Hrsg.), Lokale Verwaltungsreform in Aktion, Basel usw., S. 267-275

- Mäding, H. 1998, Empirische Untersuchungen zur Verwaltungsmodernisierung aus dem Deutschen Institut für Urbanistik, in: Deutscher Städtetag, Verwaltungsmodernisierung, DSt-Beiträge zur Kommunalpolitik, Reihe A, Heft 27, Köln, S. 9 ff
- Naschold F. 1995, Ergebnissteuerung, Wettbewerb, Qualitätspolitik. Entwicklungspfade des öffentlichen Sektors in Europa. Modernisierung des öffentlichen Sektors, Berlin 1995
- Naschold F./Oppen M./Wegener A. (Hrsg.) 1997, Innovative Kommunen. Internationale Trends und deutsche Erfahrungen. Stuttgart usw.
- Naschold, Frieder/Jann, Werner/Reichard, Christoph, (1999): Innovation, Effektivität, Nachhaltigkeit. Internationale Erfahrungen zentralstaatlicher Verwaltungsreform. Berlin.
- Nullmeier, Frank 2001, Kennzahlen und Performance Measurement, in: Blanke, Bernhard/ von Bandemer, Stephan/ Nullmeier, Frank/ Wewer, Göttrik (Hrsg.), Handbuch zur Verwaltungsreform, 2. Aufl., Opladen: Westdeutscher Verlag, S. 383 ff.
- Pawson, R. /Tilley N. 1997, Realistic Evaluation, London etc: Sage
- Pollitt, Christopher, (1995), Justification by Works or by Faith? Evaluation the New Public Management. In: Evaluation, vol. 1, no. 2, p. 133-154.
- Pollitt, Christopher/Bouckaert, Geert, (2000): Public Management Reform. A Comparative Analysis. Oxford.
- Pollitt, Christopher / Bouckaert, Geert 2003, Evaluating Public Management Reforms: An International Perspective, in: Wollmann, Hellmut (ed.) 2003, Evaluation in Public Sector Reforms, Concepts and Practice in International Perspective, Aldershot: Edward Elgar (forthcoming)
- Pressman, J. ./Wildavsky, A. 1974, Implementation, Berkeley 1974.
- Reichard, Ch. 1994, Umdenken in den Kommunen, Berlin 1994
- Reichard, Ch. 1998, Der Produktansatz im Neuen Steuerungsmodell – von der Euphorie zur Ernüchterung, in: D. Grunow/ H. Wollmann (Hrsg.), Lokale Verwaltungsreform in Aktion, Basel/ New York, S. 85 – 102
- Richter, Walter 2001, Controlling und Berichtswesen, in: Blanke, Bernhard/ von Bandemer, Stephan/ Nullmeier, Frank/ Wewer, Göttrik (Hrsg.), Handbuch zur Verwaltungsreform, 2. Aufl., Opladen: Westdeutscher Verlag, SS. 392 ff.
- Rist, Ray C. (ed.) 1990, Program Evaluation and the Management of Government, New Brunswick/London
- Ritz, A./ Thom St./Rieder N. 2002, Was bewirken NPM-Reformen? Die Evaluation von New Public Management-Reformen auf Bundesebene, in: Neue Züricher Zeitung vom 20.2.2002, S. 15
- Ritz, Adrian 1999: Die Evaluation von NPM. Bern.
- Schuster, F. 2001a, Der interkommunale Leistungsvergleich als Wettbewerbssurrogat, Phil. Diss., Humboldt-Universität, Institut für Sozialwiss., verv. Ms. April 2001
- Schuster 2001b, Benchmarking als Ersatz für Wettbewerb. Können interkommunale Leistungsvergleiche ein Motor für Veränderungen sein? in: T. Edeling u.a. (Hrsg.), Reorganisationsstrategien in Wirtschaft und Verwaltung, Opladen: Leske + Budrich, S. 201-228.
- Röber, M./ Schröter, E. 2001, Berliner Politik- und Verwaltungsstrukturen: Neue Zeiten und alte Probleme, in: Schröter, E. (Hrsg.), Moderne Verwaltung für moderne Metropolen, Opladen: Leske + Budrich, S. 159-169

Senat von Berlin 2000, Vorlage über die Umsetzung des Verwaltungsreform-Grundsätze-Gesetzes, verv. Ms., Berlin

Seyfried, P. 2001, Erfolg der Bündelung von Aufgaben auf der unteren Verwaltungsebene, in: Verwaltung und Management, S. 156 ff.

Seyfried, P. 2002, Braucht Verwaltungspolitik Evaluation?, zitiert in Jacoby, Tagungsbericht, in; Zeitschrift für Evaluation, vol. 1, Heft 1, S. 115 ff.

Stockmann, Reinhard 2000, Evaluation in Deutschland, in: Stockmann, Reinhard (Hrsg.) 2000a, Evaluationsforschung, Opladen: Leske + Budrich, S. 11-40

Thieme W./ Prittwitz G. 1981, Durchführung und Ergebnisse der kommunalen Gebietsreform, Baden-Baden

Thoenig, Jean Claude 2003, Learning from evaluation practice: The case of public-sector reforms, in: Wollmann, Hellmut (ed.) 2003, Evaluation in Public Sector Reforms, Concepts and Practice in International Perspective, Aldershot: Edward Elgar (forthcoming),

Thom, St./Ritz A. 2000, Public Management – Innovative Konzepte zur Führung im öffentlichen Sektor, Wiesbaden 2000

Wagner, P./ Wollmann. H. 1986, Fluctuation in the development of evaluation research, in: International Social Science Journal, vol. , pp. 205-218

Wegrich, K./Jaedicke, W. /Lorenz, S./Wollmann H. 1996, Kommunale Verwaltungspolitik in Ostdeutschland, Basel usw: 1996

Widmer, Thomas 2002, Staatsreformen und Evaluation: Konzeptionelle Grundlagen und Praxis bei den Schweizer Kantonen, in: Zeitschrift für Evaluation, vol.1., Heft 1, S. 101-114

Wollmann, H. 1994, Evaluationsansätze und -institutionen in Kommunalpolitik und -verwaltung. Stationen der Planungs- und Steuerungsdiskussion. In: Schulze-Böing, Matthias/Johrendt, Norbert, (Hrsg.): Wirkungen kommunaler Beschäftigungsprogramme. Basel usw., S. 79 ff.

Wollmann, H. 1996 Verwaltungsmodernisierung: Ausgangsbedingungen, Reformanläufe und aktuelle Modernisierungsdiskurse, in: Ch. Reichard/ H. Wollmann (Hrsg.), Kommunalverwaltung im Modernisierungsschub?, Basel usw. , S. 1-50

Wollmann, H. 1998, Licht und Schatten des Berliner Verwaltungsreformprojekts, in: D. Grunow/ H. Wollmann (Hrsg.), Lokale Verwaltungsreform in Aktion, Basel usw , S. 219 –242

Wollmann, H. 1999, Politik- und Verwaltungsmodernisierung in den Kommunen. Zwischen Managementlehre und Demokratiegebote, in: Die Verwaltung, Heft 3, S. 346 ff.

Wollmann, Hellmut 2000a, Evaluation und Verwaltungspolitik, in: Stockmann, Reinhard (Hrsg.), Evaluationsforschung, 1. Auflage, Opladen: Leske + Budrich, S. 195 ff.

Wollmann H. 2000b, Staat und Verwaltung in den 90er Jahren. Kontinuität oder Veränderungswelle?, in: R. Czada/ H. Wollmann (Hrsg.), Von der Bonner zur Berliner Republik, Opladen 1999, S. 694- 731

Wollmann, H. 2002a, Verwaltungspolitik und Evaluierung. Ansätze, Phasen und Beispiele im Ausland und in Deutschland, in: Zeitschrift für Evaluation, Jhg. 2, H. 1, S. 75-101

Wollmann, Hellmut 2002b, Contractual Research and Policy Knowledge, in: International Encyclopedia of the Social and Behavioral Sciences, vol. 5

Wollmann, H. 2002c, Evaluierung von Verwaltungspolitik: Reichweite und Grenzen. Ein internationaler Überblick, in: Verwaltungsarchiv

Wollmann, Hellmut 2003a, Kontrolle in Politik und Verwaltung. Evaluation, Controlling und Wissensnutzung, in: Schubert, Klaus/ Bandelow, Nils C. (Hrsg.), Lehrbuch der Politikfeldanalyse, München/Wien: Oldenbourg, S. 335 ff.

Wollmann, Hellmut (ed.) 2003b, Evaluation in Public Sector Reforms, Concepts and Practice in International Perspective, Aldershot: Edward Elgar (forthcoming)

Wollmann, Hellmut 2003c, Evaluation in public-sector reform: Towards a „third wave“ of evaluation?, in: Wollmann, Hellmut (ed.) 2003, Evaluation in Public Sector Reforms, Concepts and Practice in International Perspective, Aldershot: Edward Elgar (forthcoming)

Wollmann, Hellmut 2003d, Evaluation in public sector reform. Approaches and patterns in cross-country perspective in: in: Wollmann, Hellmut (ed.) 2003, Evaluation in Public Sector Reforms, Concepts and Practice in International Perspective, Aldershot: Edward Elgar (forthcoming)