

Verwaltungspolitik und Evaluierung

Ansätze, Phasen und Beispiele im Ausland und in Deutschland.

Hellmut Wollmann, Humboldt-Universität, Berlin

0. Fragestellung

Aus guten Gründen stellte die *Deutsche Gesellschaft für Evaluierung* (DeGEval) ihre Jahrestagung 2001, die sie am 4. und 5.10.2001 an der Hochschule für Verwaltungswissenschaften in Speyer organisierte, unter das Generalthema „Verwaltungsmodernisierung und Evaluierung“ und findet diese Fragestellung in der ersten Nummer der „Zeitschrift für Evaluation“ besondere Aufmerksamkeit. Besitzt doch Evaluierung als Analyseverfahren, durch das die Wirkungen und Wirksamkeit, die Kosten und Nutzen politischen und administrativen Handelns erfaßt und in Politikformulierung und Verwaltungsvollzug „rückgemeldet“ werden sollen, in der Welle der Verwaltungspolitik und -modernisierung, die seit den frühen 1980er Jahren unter dem Banner des *New Public Management* die meisten OECD-Länder und seit den frühen 1990er Jahren auch die Bundesrepublik ergriffen hat, in doppelter Hinsicht einen herausragenden konzeptionellen und strategischen Stellenwert.

- Zum einen betrifft dies die Einführung und Anwendung von Evaluierungsverfahren als maßgebliche Aufgabe der Verwaltungspolitik. Insoweit geht es um „Evaluierung *in* der Verwaltungspolitik“.
- Zum andern handelt es sich um die Aufgabe und um Verfahren, die Verwaltungspolitik selber zum Gegenstand von Evaluierung zu machen. Insoweit geht es um „Evaluierung *von* Verwaltungspolitik“.

So strebte denn die Speyerer Tagung der Deutschen Gesellschaft für Evaluierung an, insbesondere in dem einleitenden Plenarvortrag von *Ernst Buschor* (vgl. Buschor 2002 in diesem Heft), in den Sitzungen des DeGEval-Arbeitskreises „Verwaltung und Evaluierung“ und mit dem abschließenden Podiumsgespräch (vgl. dessen Zusammenfassung von Jacoby 2002 in diesem Heft) die Verbindung von Evaluierung und Verwaltungspolitik als ein zentrales Thema sowohl der Evaluierungs- als auch verwaltungspolitischen Diskussion kenntlich zu machen und ihm Resonanz und Verbreitung zu schaffen.

Vor diesem Hintergrunde soll im nachstehenden Aufsatz ein Überblick über Entwicklung und Stand der Evaluierung *in* und *von* Verwaltungspolitik in einer Reihe relevanter anderer Länder und in Deutschland gegeben werden¹.

Der Aufsatz soll in drei Schritten entfaltet werden.

Zunächst seien einige konzeptionelle und methodische Bemerkungen zur Evaluierung von Verwaltungspolitik vorausgeschickt.

Sodann soll eine kurze Übersicht über die Entwicklung in einigen relevanten andern Ländern gegeben werden.

Schließlich soll die Entwicklung in Deutschland diskutiert werden.

1. Konzeptionelle und methodische Vorbemerkungen

1.1. Eigentümlichkeiten der Verwaltungspolitik und ihrer Evaluierung

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 (vgl. Wollmann 2000a: 200ff.) – im Gegensatz zu „normalen“ („substanziellen“) Politiken². Die letzteren zielen darauf, bestimmte (sozio-ökonomische, ökologische, 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 Evaluierung auf diese institutionellen Veränderungen fokussiert ist, untersucht sie deren Verlauf und (Zwischen-)Ergebnisse unter der Fragestellung, ob die

1 Der Aufsatz geht auf den schriftlichen Beitrag zurück, den der Verf. für eine Sitzung des Arbeitskreises *Evaluierung und Verwaltung* der *Deutschen Gesellschaft für Evaluierung* (DeGEval) im Rahmen von deren Jahrestagung am 4. und 5.10.2001 an der Deutschen Hochschule für Verwaltungswissenschaften in Speyer verfaßte, und ist eine überarbeitete und erweiterte Fassung von Wollmann 2002b, für die nicht zuletzt den in die Speyerer Tagung eingebrachten Beiträgen (Buschor 2002 – in diesem Band –, Widmer 2002 – in diesem Band –, Seyfried 2001b, Dieckmann 2001 sowie Jacoby 2002 – in diesem Band) zu danken ist.

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.

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 Evaluierung hier als *Institutionen-Evaluierung* bezeichnet (vgl. Wollmann 2000a: 200)⁴.

- Als weiteres Glied der Ursache-Wirkungs-Kette strebt die Verwaltungspolitik an, eine Veränderung/Verbesserung der administrativen Handlungsprozesse⁵ zu erreichen. Die auf das Verwaltungshandeln als (prozesshafte) „performance“ fokussierte Evaluierung wird hier als *Performanz-Evaluierung* bezeichnet (vgl. Wollmann 2000a: 201, ähnlich Thoenig 2001: 208)
- 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: 110f.). 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 verstanden⁶, 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 2001: 174) oder auch der politisch-demokratischen Strukturen insgesamt („effects on the political-administrative decision-making system“, vgl. Christensen/Lagreid 2001: 32f.) beziehen können.

1.2. Methodenprobleme

In der Evaluierung von Verwaltungspolitik stellen sich spezifische methodische Schwierigkeiten, die die aus der Evaluierung von „substanziellen“ Politikprogrammen, Maßnahmen usw. („Programmevaluierung“, vgl. Derlien 1981, Rist 1990) geläufigen methodischen Probleme (vgl. zuletzt insbesondere Pawson/Tilley 1997) 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 („Konzipierungs-

4 In der internationalen Diskussion wird diesbezüglich auch von „process Evaluierung“ gesprochen, vgl. Christensen/Laegreid/Wise 2001: 54: „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 2001: 173) sprechen.

6 Zur Unterscheidung von *outputs*, *outcomes*, *impacts* vgl. bereits Levy/Meltner/Wildavsky 1974.

problem) und zum ändern (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/Laegreid 2001, Wollmann 2000a: 201ff.):

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 Evaluierung haben soll (und kann) und welche Reformziele sich hieraus ergeben. Entlang der erwähnten Wirkungskette kann die Evaluierungsuntersuchung gerichtet werden

- auf Verlauf und Ergebnis der institutionellen Veränderungen (Institutionen-evaluierung),
- auf die veränderten Handlungsprozesse und -muster (Performanz-Evaluierung, „Prozess-Evaluierung“),
- auf die (verwaltungsökonomische) Effizienz des Reformprojekts, insbesondere die erzielte Kostenreduktion (vgl. Pollitt/Bouckaert 2001: 180: „saving of money“),
- auf die (materiellen) Leistungsverbesserungen (Output-Evaluierung)
- oder auf darüber hinausreichende (verteilungspolitische, „systemische“ usw.) Veränderungen (Outcomes-Evaluierung).

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 sich die Promotoren von Modernisierungsvorhaben je nach Zielsetzung, Neigung und Gusto bedienen. Damit steht die Evaluierung 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 Evaluierung auf bestimmte Teilelemente und deren Zielsetzungen zu beschränken und zu konzentrieren. Auch bei einer solchen (forschungspragmatisch gebotenen, jedoch eine umfassende Evaluierung aus dem Blick rückenden) Fokussierung stellt sich das für veraltungspolitische Reformen spezifische Problem, dass es sich in der Regel nicht um „eindimensionale“, sondern um „multiple“ Ziele und Kriterien handelt (vgl. Pollitt/Bouckaert 2001: 172), welche sich 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 lie-

genden „Handlungstheorien“) bemühte Einstiegsphase der Evaluierung den (Neben-) Effekt haben, zur Aufhellung der Politikformulierung und der ihr verborgenen Ziel- und Handlungsannahmen beizutragen (vgl. Seyfried 2001b).

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 2001: 207 „Overall external effects... hard to measure“). Angesichts dessen kann es kaum verwundern, daß sich Evaluierungsuntersuchungen vielfach auf solche Ziele als Meßlatte „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 „Meßlatte“ der Evaluierung 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 Evaluierung nach wie vor noch auf überraschende Grenzen (und Blockierungen) stößt⁷.

1.2.2. Kausalitätsproblem

Ist schon bei der (ex-post) Evaluierung „substanzieller“ Politiken die Anwendung des Repertoires der sozialwissenschaftlichen Methoden zur Gewinnung valider („kausaler“) Aussagen schwierig genug, so stellen sich die methodischen Probleme bei der Evaluierung von Verwaltungspolitik auch insoweit noch verschärfter dar (vgl. Pollitt/Bouckaert 2001: 170f., Wollmann 2000: 202ff.).

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

- 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 Evaluierungsuntersuchungen, die (z.B. in Australien und Neuseeland) auf eine umfassendere Evaluierung 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 2001: 171)

1.2.3. Quasi-Evaluierung

Angesichts der zahlreichen methodischen Probleme, mit denen die Evaluierung von verwaltungspolitischen Strategien und Maßnahmen konfrontiert ist, plädierte unlängst *Jean-Claude Thoenig* für das Konzept einer „Quasi-Evaluierung“ (vgl. Thoenig 2001: 205). Diese soll sich in erster Linie auf die eher deskriptive Erfassung der verwaltungspolitisch intendierten Veränderungen konzentrieren und beschränken und auf den der „klassischen“ (ex-post) Evaluierung eigentümlichen methodischen Aufwand und hierbei insbesondere auch auf den Anspruch, sozialwissenschaftlich valide „Kausalaussagen“ zu gewinnen, verzichten. Innerhalb der bekannten Varianten von Evaluierung (ex-ante Evaluierung, „ongoing“/formative Evaluierung/Monitoring und ex-post/summative Evaluierung) ist eine solche „Quasi-Evaluierung“ am ehesten als ein Fall „ongoing“/„formativen Evaluierung“ und des „Monitoring“ zu identifizieren, indem sie maßgeblich auf eine ständige Interaktion zwischen den für die Evaluierung initiiierenden und für sie verantwortlichen politischen und administrativen Akteuren und den mit ihrer Durchführung beauftragten externen Evaluatoren angelegt ist.

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

8 vgl. Pollitt/Bouckaert 2001: 170: „it is almost unknown for management reforms to be subjected to true experimental procedures“.

- Durch die konzeptionelle und methodische „Verschlankung“ der Evaluierung 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 Evaluierung) verantwortlichen politischen und administrativen Akteuren und den Evaluatoren zugeschnitten ist, eröffnet und fördert sie zwischen diesen einen verwaltungspolitischen „Dialog“ als „a gradual learning process in successive stages that forsters an information culture“ (Thoenig 2001: 205).

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

1.3. Zeitpunkt und Modalität der Evaluierung

Innerhalb der auf den Zeitpunkt der Evaluierung im Politikzyklus abhebenden Unterscheidung zwischen *ex-ante*, *formativer/ongoing* und *ex-post/summativer* Evaluierung (vgl. etwa Wollmann 2000a: 197f.) 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* Evaluierung 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: 95ff., Schuster 2001).

Ansätze der *ex post-Evaluierung* zielen in der Verwaltungspolitik darauf, die Wirkungen verwaltungspolitischer 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 („Eva-

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 2001: 205.

luierungsforschung“ im engeren Sinne), bis zu methodisch einfacheren Erhebungen (etwa in der Form der erwähnten „Quasi-Evaluierung“ oder von Praxis- und Erfahrungsberichten) reichen.

1.4. Träger der Evaluierung

Davon ausgehend, dass eine zweckmäßige Unterscheidung der Trägerschaft von Evaluierung danach vorzunehmen ist, wer die Evaluierung initiiert (und gegebenenfalls finanziert) und wer diese durchführt, kann von *interner* Evaluierung zunächst (im engeren Sinn) dann gesprochen werden, wenn das Evaluierungsverfahren in der operativen Verwaltungseinheit selbst initiiert, institutionalisiert und durchgeführt wird. Mit Bezug darauf ist vielfach von Formen der *Selbstevaluierung* oder – in Übernahme der angelsächsischen Terminologie – von *in house*-Evaluierung 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 Evaluierung besondere Aufmerksamkeit. Darüber hinaus sollen hier solche Evaluierungsansätze, die im Kontext der Verwaltung und Exekutive insgesamt initiiert bzw. durchgeführt werden (z.B. als „Erfolgskontrolle“ von Pilotvorhaben durch ein Ministerium), als *interne* Evaluierung (im weiteren Sinne) verstanden werden.

Demgegenüber kann von *externer Evaluierung* dann gesprochen werden, wenn diese Evaluierung 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 2002a), sei es als (eigenfinanziertes) wissenschaftliches Projekt. Zur Abgrenzung von der *internen* Evaluierung“ (im engeren wie weiteren Sinne) soll hier von *externer* Evaluierung (im weiteren Sinne) auch in Fällen die Rede sein, in denen die Initiierung, Durchführung oder Beauftragung der Evaluierung durch Institutionen 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* Evaluierung soll der Gliederung der nachfolgenden Darstellung dienen.

2. Evaluierung und Verwaltungspolitik im internationalen Überblick

2.1. Ansätze interner Evaluierung

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 Evaluierung zu verbinden (vgl. Naschold/Jann/Reichard 1999: 23 ff; Jann/Reichard 2001: 99).

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 Evaluierung verfolgt werden, wurde von einem kundigen Beobachter geradezu als „an obsessive concern for outputs and accountability“ bezeichnet (vgl. Halligan 2001: 80).

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 2001: 60ff.).

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 2001: 59), in dessen Rahmen diese Verwaltungseinheiten verpflichtet sind, jährliche *result*-gestützte Berichte zu erstatten. Diese jährlichen Berichte werden vom schwedischen Rechnungshof evaluiert und „benotet“ (vgl. Jann/Reichard 2001: 99). Hierbei machen Beobachter der „management by results“-Praxis in den skandinavischen Ländern darauf aufmerksam, dass hier die Funktion der indikatoren-gestützten Analyse- und Rückmeldeschleifen 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 fördern¹⁰.

Auch in *Norwegen* wiederum treten Finanzministerium und der Rechnungshof in den letzten Jahren immer stärker als Advokaten einer Performanz-Evaluierung („performance auditing“) hervor (vgl. Christensen/Lagreid/Wise 2001: 55).

In der *Schweiz* ist auf das 1997 in vier Ämtern der Bundesverwaltung eingeführte Reformkonzept zu verweisen, das sich an NPM-Vorstellungen orientiert (*FLAG* = **F**ührung durch **L**istungsauftrag und **G**lobalbudget, vgl. Ritz 1999: 39ff.) und in dem interne Informationschleifen und Berichtssysteme eine wesentliche Rolle spielen.

10 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“.

Benchmarking

Auf die Nutzung von Verfahren interner Evaluierung 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 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. Ein interkommunaler Wettbewerb soll dadurch angeregt werden, daß die ermittelten Daten der einzelnen Kommunalverwaltungen von der *Audit Commission* in Leistungstabellen jährlich landesweit bekannt gemacht werden (vgl. Schuster 2001: 70ff., Adamaschek/Banner 1997: 223ff.).

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 (vgl. Schuster 2001: 74; Adamaschek/Banner 197: 221ff.).

2.2.1 Ansätze externer Evaluierung

Verhältnismäßig frühe Beispiele von externer Evaluierung der eingeleiteten Verwaltungsreform liefert *Australien*, wo 1993 die bis dahin rund zehnjährigen Reformaktivitäten Gegenstand einer ersten umfassenden Evaluierung wurden. Diese wurden auf Veranlassung eines Parlamentsausschusses von einer *Task Force on Management Improvement*, einer quasi-unabhängigen Gruppe von Regierungsbeamten, unternommen (vgl. Halligan 2001: 78). Wesentliche Anstöße für die lebhaften Evaluierungsaktivitäten, die die Verwaltungsmodernisierung in Australien kennzeichnen, ebenso wie für die Verwaltungsmodernisierung insgesamt gehen vom Bundesfinanzministerium aus (vgl. Halligan 2001: 83, der vom Finanzministerium als dem „evaluation champion“ spricht). Zu Umfang und Intensität dieser Evaluierungstätigkeiten 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 2001: 80).

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 Evaluierungen durchgeführt, zunächst 1991 von der *Steering Group for Review of State Sector Reform* und dann 1996 dadurch, daß der international renommierte Wissenschaftler und Experte *Alan Schick* mit einer umfassenden Evaluierung beauftragt wurde (vgl. Halligan 2001: 78); 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 2001: 78).

Ein weiteres Beispiel lebhafter Evaluierungstätigkeiten 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 Evaluierungsverpflichtungen aufzunehmen, schrieb der *Civil Service Reform Act* von 1978 entsprechende Evaluierungen vor und enthielt zugleich die Ermächtigung zur Durchführung einschlägiger verwaltungspolitischer Experimente (vgl. Christensen/Lagreid/Wise 2001: 60). Darüber hinaus wurde unter dem *Government Performance and Results Act* von 1992 ein umfassendes System zur Evaluierung (*auditing*) von Verwaltungstätigkeiten etabliert. Außerdem gingen Evaluierungsanstöße vom *General Accounting Office* (einem traditionellen Promotor von Evaluierungen) und vom *Office of Management and Budget* aus.

In *Schweden* spielt der Rechnungshof auch in der Evaluierung der Verwaltungsmodernisierung eine maßgebliche Rolle. Ein Beispiel für Schwedens traditionelles System unabhängiger Kommissionen und für die Evaluierungskultur des Landes gibt die *Expertengruppe für den Öffentlichen Sektor*, die regelmäßige Berichte zur Leistungsstärke und Wirtschaftlichkeit der Regierungsaktivitäten vorlegt (vgl. Christensen/Lagreid/Wise 2001: 58).

In beachtlichem Umfang sind neuerdings in der *Schweiz* externe Evaluierungen 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 Evaluierung 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 Evaluierungen durchgeführt worden sind (vgl. Widmer 2001, 2002). Schließlich haben mehrere größere Städte Evaluierungsuntersuchungen in Auftrag gegeben (vgl. Widmer 2002 in diesem Heft mit Nachweisen).

Auffällig im internationalen Überblick ist, dass in *Großbritannien* zunächst kaum externe Evaluierungen 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 Evaluierungen merklich erhöht (vgl. Pollitt/Bouckaert 2001: 171).

Zusammenfassend ist freilich hervorzuheben, dass der Stand der externen Evaluierung 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) Evaluierung 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-Evaluierung“ zu bezeichnen sind), während die weitergehenden „materiellen“ Wirkungen allenfalls in der Dimension der Veränderung administrativer Handlungsprozesse und -muster („improved proces-

ses“), also als „Performanz-Evaluierung“ sowie als (im wesentlichen auf Kosteneinsparung verkürzte) „Effizienz“-Aussagen in den Blick gerückt werden¹¹. Hingegen bleibt die Wirkungsdimension der outputs und outcomes fast durchweg ausgeblendet. Sind (ausnahmsweise) umfassendere Evaluierungsaussagen anzutreffen, sind sie eher generell und „weich“ formuliert¹².

3. Verwaltungspolitik und Evaluierung in Deutschland

Die Evaluierungsfunktion als Instrument des Politik- und Verwaltungshandelns wurde in der deutschen Verwaltungspolitik erstmals in der „Reformwelle“ der 1960er und 1970er Jahre lebhaft diskutiert und in Ansätzen praktiziert. In dem der „Planungsdiskussion“ zugrunde liegenden Politikmodell und -zyklus (Planung, Implementation und Ergebnis) wurde der Evaluierung die strategische Aufgabe beigemessen, die Ergebnisse und Wirkungen des Politik- und Verwaltungshandelns zu ermitteln und diese „rückzumelden“. Der neue Informations-, Analyse- und Evaluierungsfokus der Verwaltungspolitik fand zum einen in der Installierung 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 Programmevaluierung; in dieser Phase der Evaluierung 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 Evaluierung und Evaluierungsforschung zu „substanziellen“ Politiken und politischen Programmen („Programmevaluierung“, „Programmforschung“, vgl. Derlien 1981, z.B. zu Städtebauförderung, Wohnungspolitik usw.), während (externe) Evaluierungen von verwaltungspolitischen Maßnahmen weitgehend fehlten¹³.

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 „Managementmo-

11 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 2001: 83.

12 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 Pollitt/Bouckaert 2001: 183).

13 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. Thieme/Prittowitz 1981.

dell und -zyklus“ (Zielsetzung, Implementation, Ergebnis) der Evaluierungsfunktion wiederum eine zentrale (und wegen ihrer strategischen Einbindung in eine umfassende Steuerungsfunktion noch gesteigerte) Rolle zu.¹⁴ (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 Wollmann 1999: 346ff.).

3.1. Ansätze interner Evaluierung

„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 Evaluierung 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 ergiebige Umweg, wenn nicht als eine Sackgasse erwiesen¹⁵. Jedoch haben Ansätze und Verfahren (indikatoren-gestützter) Evaluierung 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 gefunden¹⁶, auch wenn „von einem funktionsfähigen Controlling in deutschen Kommunalen Verwaltungen noch nicht die Rede sein“ kann (Bogumil)¹⁷

Inzwischen sind *Kosten- und Leistungsrechnung* und *Controlling* auch in der Verwaltungspolitik der *Bundesländer*¹⁸ und zuletzt auch des *Bundes* in der Absicht aufgegriffen worden, „einen Prozeß zu initiieren, der an strategischen Zielen ausge-

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

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

16 Vgl. hierzu zuletzt Kuhlmann 2002: 24 u.a. auf der Basis einer eigenen Auswertung der vom Deutsche Städtetag 1998 durchgeführten Befragung seiner Mitgliedsstädte.

17 vgl. J. Bogumil im DeGEval-Podiumsgespräch vom 5.10.2001, in dem Kurzbericht von Jacoby 2002 (in diesem Heft).

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

richtete Maßnahmen festlegt, diese auf ihre Wirksamkeit überwacht und durch einen permanenten ‚Soll-Ist-Vergleich‘ zu weiteren Verbesserungen führt“¹⁹.

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 Vergleichsringen* als ein Verfahren kommunaler Selbstevaluierung und (interkommunalen) Quasi-Wettbewerbs propagierte und finanziell förderte. 1997 waren bundesweit 150 Städte am kommunalen Leistungsvergleich der Stiftung beteiligt (vgl. Schuster 2001; Wollmann 2000a: 212f.) 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 Evaluierung als auch für den interkommunalen Leistungsvergleich nutzen (vgl. Kuhlmann 2002: 7). Ende 2001 bestanden bundesweit 62 regional organisierte thematische Vergleichsringe, an denen (Ende 2001) insgesamt 672 Kommunen mitwirken²⁰. Auch wenn mithin interkommunale 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 Selbstevaluierung reduzieren, und könnte ein wirklicher Leistungswettbewerb und eine „Kontrolle“ durch die Bürger gesichert werden.²¹

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 Kennzahlen zu erarbeiten²². Auch

19 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.: 18.

20 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.

21 vgl. hierzu eindringlich J. Bogumil (in: oben Fußnote 17).

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

in der Verwaltungen der Bundesländer²³ und des Bundes²⁴ findet *benchmarking* als neues Steuerungsinstrument wachsende Aufmerksamkeit und Anwendung.

Pilotvorhaben und deren (interne) Evaluierung

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 Evaluierung 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 (vgl. Banner 2001: 2099, zu den „free communes“ vgl. Baldersheim 1996: 293ff.), 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 (vgl. Wollmann 2000a: 217 mit Nachweisen). Allerdings wurde auf methodisch anspruchsvolle (externe) Evaluierung dieser Pilotvorhaben durchweg verzichtet (kritisch Klages 1996: 7ff.) 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- und Leistungsrechnung sowie kennzahlen- und kostenorientierten Führungssystemen (Controlling)“ in 30 „Pilot“-Behörden eingeleitet²⁵.

(Verwaltungsinterne) „Erfolgskontrolle“

Des weiteren ist zunehmend die Bereitschaft insbesondere in der Verwaltungspolitik der *Bundesländer* zu verzeichnen, zu den von ihnen eingeleiteten verwaltungs-

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

24 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>.

25 Vgl. Innenministerium BW 1999: 123ff. 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.

politischen Reformvorhaben Berichte und „Erfolgskontrolle“ vorzubereiten und vorzulegen, die methodisch überwiegend auf (verwaltungsinterne) Umfragen und Erfahrungsberichte gestützt sind. Beispielfhaft 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 2001b: 2)²⁶. 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 2001a).

In *Nordrhein-Westfalen* sei auf „erste Umfrage (des Innenministeriums) zum Stand der Binnenmodernisierung bei den obersten Landesbehörden“ (vgl. Behrens 2002: 12ff.) 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 wurde²⁷.

Einen bemerkenswert umfassenden Bericht zum „Stand der Umsetzung“ des (Berliner) „Verwaltungsgrundsatzgesetzes“ (als Gesetzesprogramms des 1994 eingeleiteten Reformprojekts)²⁸ 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 Bürgerorientierung usw., vgl. Senat von Berlin 2000: 3), gelangte in seinen Einzel- und Gesamtbetrachtungen zu überwiegend kritischen Einschätzungen²⁹.

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“

26 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. Vgl. auch Seyfried in Jacoby 2002 – in diesem Heft (vgl. oben FN 17).

27 Vgl. VM-Impulse, 2002/Heft 1: 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.

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

29 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).

zu schaffen³⁰. Ein erstes Beispiel gibt die (mittels verwaltungsinterner Befragungen vorgenommene) „Evaluation“ der Reformkomponente „Ideenmanagementsystem“³¹.

3.3. Ansätze externer Evaluierung

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

Als rares Beispiel einer von der Exekutive/Verwaltung initiierte und (als Auftragsforschung vergebene) externe Evaluierung 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 Evaluierung dieser lokalen Projekte beauftragte (vgl. Maaß/Reichard 1998, Maß 1999).

Als Projekt methodisch anspruchsvoller Evaluierungsforschung ist eine im Auftrag eines Bundesministeriums durchgeführte Untersuchung des öffentlichen Gesundheitsdienstes auf kommunaler Ebene hervorzuheben (vgl. Grunow/Grunow-Luther 2000). Des Weiteres sei eine (laufende) Forschungsarbeit genannt, in der der Einfluß der Deregulierung (Vereinfachung der immissionsschutzrechtlichen usw. Vorschriften) auf das Genehmigungsverfahren (Dauer usw.) untersucht werden soll³².

Während externe Evaluierungen, 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 Evaluierung 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: 218ff.), die – überwiegend im Wege schriftlicher Umfragen und deren Auswertung – vom *Deutschen Städtetag* (vgl. Grömig/Gruner 1998)³³, 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 Evaluierungsab-

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

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

32 Die Untersuchung wird an der Deutsche Hochschule für Verwaltungswissenschaften Speyer (Leitung: *Jan Ziekow*, Laufzeit 1.9.2001 bis 30.11.2003, Förderung: Innenministerium Baden-Württemberg) bearbeitet.

33 Für Sekundärauswertungen der DSt-Daten vgl. etwa Kuhlmann-Lorenz/Wegrich 2001: 25ff., 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, Banner 2001: 281.

sicht – 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 viel³⁴.

- Die meisten einschlägigen Forschungsvorhaben verfolgten einen *kommunalen* Fokus³⁵, 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 evaluierungsorientierter Verwaltungsforschung gemacht worden³⁶.
- *International vergleichende* Untersuchungen wurden bislang insbesondere von *Frieder Naschold* angestoßen und von dessen Forschungsgruppe³⁷ oder in einem von ihm initiierten Forschungsnetzwerk³⁸ 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 *Institutionenevaluierung* einzustufen. Demgegenüber wurden Fragen nach den weiterreichenden Wirkungen, sei es „Performanz“ oder gar „outcomes“, bislang mit geringerem Nachdruck verfolgt.

Abschließend sei auf einen bemerkenswerten Versuch *Gerhard Banners* verwiesen, eine recht umfassende Evaluierung der bisherigen Entwicklung und Leistung des *Neuen Steuerungsmodells* über einen „Vergleich des Erreichten mit den ursprünglichen Zielen des Neuen Steuerungsmodells“ (Banner 2001: 283) vorzunehmen und sich hierfür auf unterschiedliche (überwiegend methodisch „qualitativer“ und „weicher“) Informationsquellen zu stützen³⁹.

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

35 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).

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

37 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).

38 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).

39 Vgl. Banner 2001: 285: „Die nachfolgende Skizze stützt sich auf die Literatur und vor allem auf zahlreiche Gespräche mit Verwaltungsmitarbeitern und der kommunalen Reformszene nahe stehenden Forschern und Verbandsvertretern, und vielleicht ist 'teilnehmende Beobachtung' dafür die am ehesten zutreffende Bezeichnung“.

4. Abschließende Bemerkungen

Zur Entwicklung von Evaluierung *in* und *von* Verwaltungspolitik in Deutschland sei – mit einem vergleichenden Blick auf die internationalen Trends – zusammenfassend hervorgehoben:

Interne Evaluierungsverfahren

Die Schlüsselfunktion, die der Schaffung von Transparenz des Politik- und Verwaltungshandelns in den Modernisierungskonzepten des *New Public Management* und des *Neuen Steuerungsmodells* als dessen deutschen Derivat beigemessen wird, hat in die deutsche Modernisierungspraxis inzwischen in der zunehmenden Einführung von „output“-Steuerung, Kosten- und Leistungsrechnung, Controlling und Berichtswesen als Verfahren interner Evaluierung und Rückmeldung Ausdruck gefunden. Ausgehend und angestoßen von der kommunalen Ebene, die seit den frühen 1990er Jahren eine verwaltungspolitische Pilotrolle übernommen hat⁴⁰, setzen inzwischen auch die Landes- und Bundesverwaltung auf die Nutzung solcher Verfahren. Desgleichen werden inzwischen Leistungsvergleiche sowohl für (intra-organisatorische) Selbstevaluierung als auch zur Stimulierung von (inter-organisatorischem) Wettbewerb eingerichtet, hinsichtlich derer wiederum die kommunale Ebene eine Vorreiterrolle gespielt hat. Soweit zu Evaluierungsverfahren *in* der Verwaltungspolitik. Aber auch in der Evaluierung *von* Verwaltungspolitik sind insbesondere auf der Ebene der Länder sich mehrende Fälle und Ansätze von „Erfolgskontroll,-,Untersuchungen und (Wirkungs-)Berichten zu beobachten, die, auf interne Erhebungen (verwaltungsinterne Befragungen, Erfahrungsberichte) und damit auf „weichere“ Analyseverfahren gestützt, darauf gerichtet sind, Verlauf und (Zwischen-) Ergebnisse von verwaltungspolitischen Reformprojekten zu erfassen und einzuschätzen.

Allerdings bleiben diese in Deutschland inzwischen vordringenden Ansätze (Controlling/Berichtswesen, Leistungsvergleiche/benchmarking und interne, methodisch „weiche“ „Erfolgskontrollen“) nach wie vor deutlich hinter dem Diskussions- und Praxisstand zurück, den Ansätze der internen Evaluierung, insbesondere im Zusammenhang mit „performance management“ und „management by results“, in einer Reihe angelsächsischer und skandinavischer Länder erreicht haben.

Ein noch deutlicherer Rückstand im internationalen Vergleich ist im Bereich der *externen* Evaluierung zu konstatieren. Zwar hat sich in der Bundesrepublik inzwischen auf ein beachtliches Feld anwendungsbezogener (auch evaluationsorientierter) Politik- und Verwaltungsforschung zur Verwaltungspolitik entwickelt. Jedoch sind Beispiele „klassischer“ (ex-post) Evaluationsforschung bislang ausgesprochen rar, in der externe (unabhängige) Forschungseinrichtungen von der Regierung/Verwaltung (oder anderen öffentlichen Institutionen) beauftragt werden, die Zielerreichung, die Wirkungen usw. von verwaltungspolitischen Reformprojekten unter Einsatz des sozialwissenschaftlichen Methodenrepertoires zu untersuchen.

40 vgl. hierzu eindringlich Banner 2001: 294f.

Wie ist dieser (relative Rück-)Stand der externen Evaluierung von Verwaltungspolitik in Deutschland zu erklären?

Zunächst ist daran zu erinnern, dass die von New Public Management und Neuem Steuerungsmodell angetriebene Reformdiskussion die deutsche Politik- und Verwaltungswelt erst seit den frühen 1990er Jahren erfaßt hat und damit die zu evaluierenden Reformprojekte überwiegend recht jungen Datums sind und mithin eine „evaluierungsfähige“ Wirkungsphase und -strecke nur zeitlich beschränkt aufweisen.

Was die mögliche Initiierung und Beauftragung von externer Evaluationsforschung zu verwaltungspolitischen Projekten durch Politik und Verwaltung selber angeht, ist auf jene komplexe und widersprüchliche Interessenlage zu verweisen, in der sich die Politik- und Verwaltungsführung hinsichtlich der Durchführung von Politikevaluierungen insgesamt befinden. Mag sie sich einerseits zumal von externer Evaluierung die Gewinnung zusätzlicher Erkenntnisse und relevanten Steuerungswissens erwarten, so muß sie – unter den Gegebenheiten des parteipolitisch und konkurrenzdemokratisch geprägten Regierungssystems – andererseits befürchten, dass Evaluierungsuntersuchungen für die politische Opposition und die (politische und Fach-) Öffentlichkeit die Stichworte und Belege für kritische Auseinandersetzungen liefern⁴¹. In Fragen der Verwaltungspolitik dürfte das Zögern der Politik- und Verwaltungsspitze noch ausgeprägter sein, sich durch die Vergabe externer Evaluierungsuntersuchungen in die verwaltungs- und institutionenpolitischen Karten sehen zu lassen und „verwaltungsfremden“ unabhängigen Evaluatoren Zugang in das „Innenleben“ der Verwaltung zu verschaffen, wobei dieses Widerstreben zusätzliche Nahrung in einer Verwaltungstradition Nahrung finden könnte, in der Verwaltungs- und Institutionenpolitik noch immer als eine Domäne der Exekutive, wenn nicht gar, wie pointiert gesagt worden ist, „zu den Innenhöfen souveräner Staatsmacht“ (König/Füchtner 1998: 7) gerechnet wird.

Hinzu kommt, daß die *Parlamente* bislang auffallend geringes 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 Verwaltungspolitik⁴² 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. Möglicherweise wird auch die parlamentarische Zurückhaltung gegenüber der Verwaltungspolitik überdies auch von deren Wahrnehmung als eines traditionellen „Reservats der Exekutive“ geprägt. Allerdings scheinen das verwaltungspolitische Engagement der Parlamente (und damit auch das Interesse an einschlägigen Evaluierungen) neuerdings zu wachsen, wie etwa die Einrichtung von Ausschüssen für Verwaltungsreform in den Landtagen der Länder unterstreicht.

41 vgl. etwa Banner 2001: 281 und J. Bogumil in Jacoby 2002 (in diesem Heft).

42 vgl. auch Seyfried 2001 sowie in Jacoby 2002 in diesem Heft.

Schließlich haben sich bislang auch die *Rechnungshöfe* in Deutschland in geringerem Maße für Verwaltungspolitik und deren Evaluierung interessiert als dies in anderen Ländern zu beobachten war. Zwar haben die Rechnungshöfe des Bundes und der Länder sich seit den 1970er Jahren von der traditionellen „titelbezogenen Ausgabenkontrolle“ und „Ordnungsmäßigkeitsprüfung“ auf eine „politische Aufgabenkontrolle“ und „Erfolgskontrolle“⁴³ zubewegt. Jedoch ging es bei dieser „evaluativen“ Neuorientierung der Prüfaktivitäten der Rechnungshöfe bislang fast ausschließlich um „substanzielle“ Politiken (Umweltpolitik, Subventionsprogramme usw.) und Investitionsvorhaben. Freilich gibt es Hinweise darauf, dass Verwaltungspolitik neuerdings zunehmend in das Blickfeld und in die Prüftätigkeit der Rechnungshöfe gerät⁴⁴.

Abschließend sei noch einmal die Forderung nach vermehrter externer Evaluierung durch (unabhängige) Forschungseinrichtungen formuliert und begründet.

- Angesichts dessen, dass die aktuelle von den Vorstellungen des *New Public Management* und des *Neuen Steuerungsmodells* inspirierte Verwaltungspolitik um die Schaffung einer umfassenden Transparenz des Politik- und Verwaltungshandelns kreist, sollte es sich von selbst verstehen, auch und gerade die verwaltungspolitischen Reformprojekte selber zum Gegenstand von Evaluierungen zu machen. Nur durch die laufende und periodische Erfassung der Kosten und Wirkungen der verwaltungspolitischen Maßnahmen und ihre „Rückmeldung“ kann der notwendige Lern-, Anpassungs- und Innovationsprozeß gesichert werden.
- Dieses Evaluierungserfordernis ist umso dringlicher, als seit den frühen 1990er Jahren erhebliche finanzielle und personelle Aufwendungen in die Vorbereitung⁴⁵ und Durchführung der Reformprojekte auf der Kommunal-, Landes- und Bundesebene geflossen sind⁴⁶. Diese sollte Anlaß zu umfassenden Kosten-

43 vgl. Dieckmann 2001: 3 sowie Dieckmann in Jacoby 2002 in diesem Heft. Laut Dieckmann sind „beispielsweise die Jahresberichte des Hamburger Rechnungshofs der letzten Jahre... bis zu 50% der Prüfungen ... unter Erfolgskontrolle einzuordnen“.

44 vgl. Dieckmann 2001: 4 sowie Dieckmann in Jacoby 2002 in diesem Heft, der mitteilt, dass 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.

45 Vgl. insbesondere die zahlreichen Organisationsuntersuchungen, die von den Bundesländern zur Vorbereitung verwaltungspolitischer Reformen (insbesondere an Unternehmensberatungsfirmen) vergeben und finanziert haben. Vgl. IM BW 1999: 184ff. (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.

46 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 Mitarbeiter-

- Nutzen-Analysen und Kosten-Wirkungs-Bilanzen geben. Hierbei kommt der Aufhellung der „Kostenseite“ der verwaltungspolitischen Projekte, über die in der verwaltungspolitischen Praxis bislang auffallend wenig bekannt ist, zur Vorbereitung realitätsnäher Kosten-Nutzen-Aussagen besondere Bedeutung zu.
- Für die Durchführung der Evaluierung der (mit fortschreitender Durchführungsdauer zunehmend „evaluierbaren“) verwaltungspolitischen Projekte (auf Bundes von Verwaltungspolitik sollte in möglichst großem Umfange von externer Evaluierung unter Einschaltung (unabhängiger) Evaluierungsträger und Forschungseinrichtungen Gebrauch gemacht werden, um den verwaltungspolitischen Erkenntnisfortschritt und Lernprozess zu sichern. Zwar erscheint die Finanzierbarkeit externer Evaluierungsforschung, zumal vor dem Hintergrund der gegenwärtigen Krise der öffentlichen Finanzen, als prekär. Jedoch sollte es verwaltungs-, aber auch fiskalpolitisch geboten und verantwortbar sein, diese Aufwendungen gerade jetzt zu machen, da die verwaltungspolitischen Projekte in ihre „Reifezeit“ eintreten und eine reiche verwaltungspolitische „Erkenntnisdividende“ abzuwerfen versprechen, die bei einem Verzicht auf sorgfältige Evaluierung unwiederbringlich verspielt würde.

Bibliographie

- Adamaschek, B./Banner G. (1997): Der interkommunale Leistungsvergleich, in: M. Pröhl (Hg.): Internationale Strategien und Techniken für die Kommunalverwaltung der Zukunft, Gütersloh: 223ff.
- Baldersheim, H. (1996): Nordic Municipalities in Transition. From Free Communes to Free Choice?, in: Ch. Reichard/H. Wollmann (Hg.): Kommunalverwaltung im Modernisierungsschub?, Basel usw. 1996: 293ff.
- Banner, G. (1991): Von der Behörde zum Dienstleistungsunternehmen. Die Kommunen brauchen ein neues Steuerungsmodell, in: VOP, 1991, Heft 1: 6ff.
- Banner, G. (2001): Kommunale Verwaltungsmodernisierung: Wie erfolgreich waren die letzten zehn Jahre?, in: E. Schröter (Hg.): Empirische Policy- und Verwaltungsforschung, Opladen: 279-304.
- Böhret, C. (2001): Verwaltungspolitik als Führungsauftrag, in: Blanke, Bernhard u.a. (Hg.): Handbuch zur Verwaltungsreform, 2. Aufl., Opladen: Leske + Budrich: 43-48.
- Bundesregierung (2002): Moderner Staat, moderne Verwaltung, Bilanz 2002, Bonn 2002.
- Bürsch M./Müller, B. (1999): Verwaltungsreformen in den deutschen Bundesländern, Bonn-Bad Godesberg 1999.
- Buschor, E. (2002): Evaluation und New Public Management. In: Zeitschrift für Evaluation. Jg. 1, H. 1.
- Christensen, T./Laegreid, P./Wise, L. (2001): Assessing Public Management Reform in Norway, Sweden and the United States, in: Wollmann, Hellmut (ed.) (2001a): Evaluating Public Sector Reforms: An International and Comparative Perspective, Special issue of *Revista Internacional de Estudios Públicos*, September 2001, pp. 41-71.

qualifikation 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. Die Stadt Ludwigshafen entgelt die von ihr beauftragte Consulting-Firma mit 10,6 Mio. DM, vgl. Grunow 1998: FN 2.

- 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. (Hg.): *Programmforschung in der öffentlichen Verwaltung*, Werkstattbericht 1 der Gesellschaft für Programmforschung, München: 9ff.
- Dieckmann, R. (2001): Möglichkeiten und Grenzen der Evaluation (in) der Verwaltungspolitik, Ms. (Statement zur Podiumsdiskussion anlässlich der Tagung der Deutschen Gesellschaft für Evaluierung am 5. Oktober 2001 in Speyer).
- Engelniederhammer St./Köpp B./Reichard, Ch./Röber M./Wollmann, H. (1999): *Berliner Verwaltung auf Modernisierungskurs*. Berlin, 1999.
- Engelniederhammer, St./Köpp, B./Reichard, 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: 581ff.
- Grunow D. (1998): Lokale Verwaltungsmodernisierung „in progress“?, in: D. Grunow/H. Wollmann (Hg.): *Lokale Verwaltungsreform in Aktion*, Basel usw.: 1-25.
- Grunow D./Grunow-Lutter V. (2000): *Der öffentliche Gesundheitsdienst im Modernisierungsprozess*. Weinheim/München.
- Halligan, J. (2001): Public Sector Modernization in Australia and New Zealand – an Evaluative Perspective, in: Wollmann, Hellmut (ed.) (2001a): *Evaluating Public Sector Reforms: An International and Comparative Perspective*, Special issue of *Revista Internacional de Estudios Publicos*, September 2001, pp. 71-93.
- Hellstern, G.-M./Wollmann H. (Hg.) (1983): *Experimentelle Politik – Reformstrohfeuer oder Lernstrategie*, Opladen.
- Hellstern, G.-M./Wollmann, H. (1984a): *Evaluierung und Evaluierungsforschung – ein Entwicklungsbericht*, in: dies. (Hg.): *Handbuch zur Evaluierungsforschung*. Bd. 1, Opladen 1984 : 17-93.
- Hellstern, G.-M./Wollmann H. (Hg.) (1984b): *Evaluierung 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 (2001): *Bericht zum Stand der Verwaltungsmodernisierung*, unveröff. Ms., Düsseldorf 21.12.2001.
- Jacoby, K.P. (2002): Möglichkeiten und Grenzen von Evaluation in der Verwaltungspolitik. In: *Zeitschrift für Evaluation*. Jg. 1, H. 1.
- Jaedicke W./Thrun T./Wollmann, H. (2000): *Modernisierung der Kommunalverwaltung. Evaluierungsstudie 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. (2001): Best Practice in Central Government Modernization, in: Wollmann, Hellmut (ed.) (2001a): *Evaluating Public Sector Reforms: An International and Comparative Perspective*, Special issue of *Revista Internacional de Estudios Publicos*, September 2001, pp. 93-112.
- KGSt, Mitgliederbefragung (1997): *KGSt-Bericht 10/1998*, Köln.
- Kißler, L./Bogumil J./Greifenstein R./Wiechmann E. (1997): *Moderne Zeiten im Rathaus?* Berlin 1997.

- Klages, H. (1996): Methodenfragen des Verwaltungswettbewerbs, in: H. Hill/H. Klages (Hg.): *Jenseits der Experimentierklausel*, Düsseldorf: 7ff.
- Knöpfel, P./Bussmann, W. (1997): Die öffentliche Politik als Evaluierungsobjekt, in: W. Bussmann/U. Klöti/P. Knöpfel (Hg.): *Einführung in die PolitikEvaluation*, Basel/Frankfurt: 78ff.
- König K./Füchtner N.(1998): „Schlanker Staat“, *Verwaltungsmodernisierung im Bund*, Speyer 1998.
- Kuhlmann: (2002): *Evaluierung kommunaler Verwaltungsmodernisierung*, unveröff. Ms., (Humboldt-Universität zu Berlin, Institut für Sozialwissenschaften) März 2002.
- Kuhlmann-Lorenz, S./Wegrich, K. (2001): Kommunalverwaltung in den neuen Bundesländern, in: Schröter, E. (Hg.): *Empirische Policy- und Verwaltungsforschung*, Opladen, S 243-260.
- Kuprath, H. (2001): Das Berliner Modell einer (de-)zentralisierten Metropole aus der Sicht der Hauptverwaltung, in: Schröter, E. (Hg.): *Moderne Verwaltung für moderne Metropolen*, Opladen: Leske + Budrich: 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 (Hg.): *Lokale Verwaltungsreform in Aktion*, Basel usw.: 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: 9ff.
- Naschold F. (1995): *Ergebnissteuerung, Wettbewerb, Qualitätspolitik. Entwicklungspfade des öffentlichen Sektors in Europa. Modernisierung des öffentlichen Sektors*, Berlin 1995.
- Naschold F./Oppen M./Wegener A. (Hg.) (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.
- Pawson, R./Tilley N. (1997): *Realistic Evaluation*, London etc: Sage.
- Pollitt, Christopher (1995): *Justification by Works or by Faith? Evaluierung 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 (2001): *Evaluating Public Management Reforms: An International Perspective*, in: Wollmann, Hellmut (ed.) (2001a): *Evaluating Public Sector Reforms: An International and Comparative Perspective*, Special issue of *Revista Internacional de Estudios Publicos*, September 2001, pp. 167-193.
- 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 (Hg.): *Lokale Verwaltungsreform in Aktion*, Basel/New York: 85-102.
- 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: 15.
- Ritz, Adrian (1999): *Die Evaluierung von NPM*. Bern.

- Röber, M./Schröter, E. (2001): Berliner Politik- und Verwaltungsstrukturen: Neue Zeiten und alte Probleme, in: Schröter, E. (Hg.): *Moderne Verwaltung für moderne Metropolen*, Opladen: Leske + Budrich: 159-169.
- Schuster, F. (2001): *Der interkommunale Leistungsvergleich als Wettbewerbssurrogat*, Phil. Diss., Humboldt-Universität, Institut für Sozialwiss., verv. Ms. April 2001.
- Senat von Berlin (2000): *Vorlage über die Umsetzung des Verwaltungsreform-Grundsätze-Gesetzes*, verv. Ms., Berlin.
- Seyfried, P. (2001a): Erfolg der Bündelung von Aufgaben auf der unteren Verwaltungsebene, in: *Verwaltung und Management*: 156ff.
- Seyfried, P. (2001b): Braucht Verwaltungspolitik Evaluation? verv. Ms. Oktober 2001 (Podiumsdiskussion anlässlich der DeGEval-Jahrestagung am 5.10.2001 in Speyer).
- Stockmann, Reinhard (2000): Evaluation in Deutschland, in: Stockmann, Reinhard (Hg.) (2000a): *Evaluationsforschung*, Opladen: Leske + Budrich: 11-40.
- Thieme W./Prittitz G. (1981): *Durchführung und Ergebnisse der kommunalen Gebietsreform*, Baden-Baden.
- Thoenig, Jean Claude (2001): Evaluierung Public Sector Reforms. Learning from Practice, in: Wollmann, Hellmut (ed.) (2001a): *Evaluating Public Sector Reforms: An International and Comparative Perspective*, Special issue of *Revista Internacional de Estudios Publicos*, September 2001, pp. 193-215.
- Thom, St./Ritz A. (2000): *Public Management – Innovative Konzepte zur Führung im öffentlichen Sektor*, Wiesbaden 2000.
- Wegrich, K./Jaedicke, W./Lorenz, S./Wollmann H. (1996): *Kommunale Verwaltungspolitik in Ostdeutschland*, Basel usw: 1996.
- Widmer, Th. (2002): Staatsreformen und Evaluation. Konzeptionelle Grundlagen und Praxis bei den schweizer Kantonen. In: *Zeitschrift für Evaluation*. Jg. 1, H. 1.
- Wollmann, H. (1994): Evaluierungsansätze und -institutionen in Kommunalpolitik und -verwaltung. Stationen der Planungs- und Steuerungsdiskussion. In: Schulze-Böing, Matthias/Johrendt, Norbert, (Hg.): *Wirkungen kommunaler Beschäftigungsprogramme*. Basel usw.: 79ff.
- Wollmann, H. (1996): Verwaltungsmodernisierung: Ausgangsbedingungen, Reformanläufe und aktuelle Modernisierungsdiskurse, in: Ch. Reichard/H. Wollmann (Hg.): *Kommunalverwaltung im Modernisierungsschub?*, Basel usw. : 1-50.
- Wollmann, H. (1998): Licht und Schatten des Berliner Verwaltungsreformprojekts, in: D. Grunow/H. Wollmann (Hg.): *Lokale Verwaltungsreform in Aktion*, Basel usw : 219-242.
- Wollmann H. (1999): Politik- und Verwaltungsmodernisierung in den Kommunen. Zwischen Managementlehre und Demokratiegebot, in: *Die Verwaltung*, 1999, Heft 3: 346ff.
- Wollmann, Hellmut (2000a): Evaluierung und Verwaltungspolitik, in: Stockmann, Reinhard (Hg.): *Evaluationsforschung*, Opladen: Leske + Budrich: 195ff.
- Wollmann H. (2000b): Staat und Verwaltung in den 90er Jahren. Kontinuität oder Veränderungswelle?, in: R. Czada/H. Wollmann (Hg.): *Von der Bonner zur Berliner Republik*, Opladen 1999: 694-731.
- Wollmann, Hellmut (2002a): Contractual Research and Policy Knowledge, in: *International Encyclopedia of the Social and Behavioral Sciences*, vol. 5 (forthcoming).
- Wollmann, Hellmut (2002b): Evaluierung von Verwaltungspolitik: Reichweite und Grenzen. Ein internationaler Überblick, in: *Verwaltungsarchiv 2002* (im Erscheinen).
- Wollmann, Hellmut (2003): (Hg.): *Evaluating Public Sector Reforms*. Aldershot: Elgar (im Erscheinen)