
Die zweite Welle der Programmevaluation

Als die GfP 1978 gegründet wurde, geschah dies noch im Hinblick auf die Planungsversuche der 70er Jahre und im Bemühen um eine Weiterentwicklung von Methodik und Institutionalisierung der in Bonn beobachtbaren Evaluationsansätze. Ähnlich wie man nach der Entwicklung des integrierten Planungs- und Budgetierungssystems PPBS 20 Jahre später bei kontinuierlicher Betrachtung „Programmhaushalte der zweiten Generation“ (Rürup/Färber 1980) feststellen konnte, meine ich, in den letzten Jahren international eine zweite Welle der Programmevaluation wachsen zu sehen.

Ich möchte im folgenden (1) diese neue Entwicklung skizzieren und mit den Aufschwüngen der Programmevaluation Ende der 60er Jahre vergleichen. Ferner (2) möchte ich versuchen, die unterschiedlichen situativen, politischen und verfassungsstrukturellen Faktoren herauszuarbeiten, die für beide Institutionalisierungsschübe ausschlaggebend waren. Schließlich (3) möchte ich der Frage nachgehen, welches die zentralen Promotoren dieser Entwicklung waren.

1. Die Institutionalisierung der Programmevaluation in historischer Perspektive

Überblickt man den heutigen Stand der Evaluation, soweit sie in den Zentralregierungen in Amerika und Europa institutionalisiert ist, stellt man sehr schnell unterschiedliche Entwicklungsniveaus und Reifegrade fest. Dies gilt sowohl hinsichtlich der praktischen Erfahrung mit diesem Instrument der Rationalisierung von Politik als auch hinsichtlich der Häufigkeit, der Regelmäßigkeit und dem

methodischen Niveau, auf dem Evaluationen betrieben werden.

Gegenwärtig würde ich den Stand der Evaluation in den einzelnen Ländern, für die mir Informationen zur Verfügung stehen, folgendermaßen umreißen:

- Ohne Zweifel kann das am höchsten entwickelte Evaluationssystem in den USA beobachtet werden. Mit rund 20 Jahren Erfahrung in der Programmevaluation hat das amerikanische System ein hohes Maß methodischer Professionalität erreicht. Diese Professionalisierung erlaubt es auch, in den USA die Programmevaluation von traditionellen Techniken der Kontrolle und der Rechnungsprüfung deutlich zu unterscheiden. Dazu hat beigetragen, daß das amerikanische Evaluationssystem fest sowohl innerhalb der Administration als auch im Kongreß – und dazu gehört auch der US-Rechnungshof – institutionalisiert ist; darüber wird im Anschluß an mein Referat Ray Rist zu uns sprechen. Es dürfte auch zutreffen, daß die amerikanischen Evaluationsbemühungen entsprechende Entwicklungen in anderen Ländern, wenn nicht ausgelöst, so doch zumindest unterstützt haben.
- In eine zweite Gruppe lassen sich Schweden, Kanada und die Bundesrepublik einordnen, denen gemeinsam ist, daß sie bis in die 60er Jahre zurückreichend im Gefolge der amerikanischen Entwicklung erste Versuche mit der Programmevaluation unternommen haben. Allerdings blieb die Programmevaluation trotz Institutionalisierung in verschiedenen Politikbereichen in diesen Ländern relativ fragmentiert, und die Zahl der Evaluationsstudien, die durchgeführt wurden, ist im Vergleich zu den USA relativ niedrig.
- Während die Programmevaluation sich in den USA kontinuierlich weiterentwickelte und in Schweden und der Bundesrepublik ihr Mitte der

70er Jahre erreichtes Niveau halten konnte, sind in Kanada die frühen Evaluationsversuche zunächst gescheitert. Dies gilt in gewisser Weise auch für Großbritannien, wo das Programme Analysis and Review System in den 70er Jahren wieder eingestellt wurde. Für Großbritannien wie auch für Kanada läßt sich jedoch feststellen, daß Ende der 70er Jahre neue Versuche politischer Planung und eben auch der Programmevaluation unternommen worden sind. Dabei ging Kanada den Weg, die Evaluationsfunktion einem Comptroller General sowie den einzelnen Ministerien zu übertragen, während die Evaluation in Großbritannien keinen organisatorischen Kristallisationspunkt erhielt, sondern eher Teil der allgemeinen, neuen Managementphilosophie wurde und primär im Rahmen des Budgetprozesses zum Tragen kommt.

- Dieser zweite Anlauf, der in Kanada und Großbritannien unternommen wurde, fällt zeitlich und, wie noch zu zeigen sein wird, systematisch zusammen mit Bemühungen zur Institutionalisierung der Programmevaluation in Ländern wie Norwegen, Dänemark, den Niederlanden, Frankreich und sogar der Schweiz. Auch für diese Länder ist kennzeichnend, daß sie die Programmevaluation stärker im Zusammenhang mit dem Budgetprozeß gesehen haben, als dies während der ersten Generation der Programmevaluation üblich war.

Die Wurzeln des amerikanischen Evaluationsystems werden üblicherweise in Präsident Johnsons „Krieg gegen die Armut“ und den dabei entwickelten Sozialprogrammen Mitte der 60er Jahre gesehen. Wir sollten uns dabei auch in Erinnerung rufen, daß die Programmevaluation in engem logischen Zusammenhang mit dem von 1965–69 regierungsweit betriebenen Planning-Programming-Budgeting-System gesehen werden muß. Die Departments der amerikanischen Bundes-

regierung, vor allem das Office of Economic Opportunity, bildeten seinerzeit Evaluationseinheiten, um die Ergebnisse der neuen Wohlfahrtsprogramme zu überprüfen, und der amerikanische Kongreß verfügte in der Economic Opportunity Act von 1967, daß das General Accounting Office die Effektivität der einzelnen Programme ebenfalls untersuchen und dem Kongreß darüber berichten sollte. Während der Präsidentschaft Nixons wurde die Evaluationsfunktion innerhalb der Exekutive sogar noch gestärkt, indem dem Bureau of the Budget (1970 Office of Management and Budget genannt) die Aufgabe zugeteilt wurde, die Evaluationsbemühungen der einzelnen Ressorts vor allem auch mit der Verbreitung methodischen Wissens zu unterstützen. Seit jener Zeit haben beide Zweige, Administration und Kongreß, ihre Bemühungen stetig fortgesetzt, die Evaluation zu einem Routineverfahren zu entwickeln. Vor allem dem Kongreß ist hierbei eine wesentliche Rolle zugefallen, sei es mit der Einführung der sogenannten Sunset-Legislation, also der Terminierung von Programmen, wenn keine positive Evaluation vorliegt, sei es mit der Einrichtung einer speziellen Einheit im General Accounting Office im Jahre 1980, die heute den Namen „Program Evaluation and Methodology Division“ trägt und rund 100 Wissenschaftler umfaßt. Diese Evaluationskapazitäten im Kongreß sollten später umso bedeutsamer werden, als die Reagan-Administration seit 1980 geradezu dramatisch die Zahl der Evaluationseinheiten in den Ministerien reduziert hat.

Auch in der Bundesrepublik war, wie wir uns erinnern werden, die Entwicklung der Programmevaluation eng mit der Intensivierung sozialpolitischer und ökonomischer Interventionsprogramme im Zuge der Reformpolitik nach 1969 verbunden (Derlien 1986). Allerdings muß man im Vergleich zu den USA sagen, daß die Evaluationslandschaft in Bonn stärker fragmentiert ist und diese Funktion bei

weitem nicht so fest und zentral institutionalisiert ist wie in den USA. In einigen Ressorts finden wir Evaluationseinheiten, während andere die Evaluation im Rahmen der Ressortforschung nicht besonders hervorheben. Hinzu trat bei uns die programmspezifische Verankerung von Evaluationspflichten in Einzelgesetzen, aber auch in Zusammenhang mit den vielfältigen Modellversuchen im Bildungs-, Gesundheits- und Sozialsektor sowie beim Programm „Humanisierung der Arbeit“ wurde Evaluation betrieben, wenngleich nicht auf dem hohen methodischen Niveau, wie wir es bei den großen amerikanischen Sozialexperimenten beobachten konnten. Nicht unerwähnt bleiben darf auch die Institutionalisierung der Evaluation über den § 7b der Bundeshaushaltsordnung im Jahre 1973, wodurch der Finanzminister immerhin in die Lage versetzt wird, von einzelnen Ministerien derartige Studien anzufordern. Wie der Finanzminister hiervon jedoch selten Gebrauch zu machen scheint, fehlt es in Bonn auch an einer zentralen Stelle im Regierungsapparat, die eine dem Office of Management and Budget vergleichbare Initiatorenfunktion übernommen hätte. Daß sich in Bonn keine Washington vergleichbare Evaluationskultur hat entwickeln können, dürfte auch mit der ambivalenten Rolle des Bundestages in dieser Sache zusammenhängen. Auf der einen Seite stellen wir eine Aufgeschlossenheit, ja sogar Forderungen in dieser Richtung in bestimmten Politikbereichen fest, die nicht zuletzt auch in gesetzliche Berichtspflichten der Exekutive einmündeten, auf der anderen Seite scheint es kaum eine intensive parlamentarische Diskussion der rund 200 Berichte zu geben, die jährlich dem Bundestag vorgelegt werden und von denen eine Reihe evaluativen Charakters ist. Im Vergleich zu Washington fällt natürlich auf, daß der wissenschaftliche Dienst des Bundestages hierfür wohl auch quantitativ und qualitativ nicht gerüstet ist.

Auch der Bundesrechnungshof hat bei der Institutionalisierung der Evaluation in den Bundesministerien keine aktive Rolle gespielt, noch hat er selbst bis vor kurzem diese Funktion für sich akzeptiert. Erst 1985 – und das ist bemerkenswert im Hinblick auf gleichgerichtete ausländische Entwicklungen – hat der Rechnungshof eine Evaluationsfunktion im Prinzip anerkannt, den Stand der Dinge in den Bundesministerien untersucht und sich entschlossen, die Evaluationsbemühungen dort zu kräftigen.

Wie die amerikanische und die deutsche Regierung so war auch die schwedische Regierung zunächst Ende der 60er Jahre primär damit beschäftigt, neue Wohlfahrtsprogramme auf den Weg zu bringen, überließ dabei jedoch die Kontrollfunktion den traditionellen fiskalischen Kontrollverfahren und -institutionen. Daß die Programmevaluation in Schweden zunächst unterentwickelt war, mag daran gelegen haben, daß die Reformpolitiken dort auf ausführliche und intensive vorgängige Analysen in Ausschüssen unter Einschluß von Sozialwissenschaftlern gestützt waren, um den für die schwedische Politik typischen breiten Konsens zu erzielen. Da allerdings die Konzeption des PPBS in Schweden eine größere Bedeutung als in der Bundesrepublik hatte, wurden frühzeitig schon Mitte der 60er Jahre betriebswirtschaftliche Wirtschaftlichkeitsanalysen und eine Programmkostenrechnung eingeführt. Dabei spielte der schwedische Rechnungshof eine prominente Rolle. Rund 100 Ökonomen und Sozialwissenschaftler waren mit derartigen Wirtschaftlichkeitsberechnungen, von denen auch eine Reihe als Evaluationen eingestuft werden können, befaßt. Ende der 70er Jahre setzte dann in Schweden angesichts ökonomischer Stagnation eine Phase globaler Überprüfungen existierender Politiken ein. Sowohl die zwischenzeitlich (1976–82) amtierende bürgerliche Regierung als auch die Sozialdemokraten haben den Bedarf derartiger gründlicher Programmüber-

prüfungen betont. Evaluationseinheiten wurden in den obersten Behörden eingerichtet, und Expertengruppen in den Ministerien haben einen Teil ihres Interesses an der Planung auf die Evaluation verlagert. Das Parlament jedoch, der schwedische Reichstag, hat in dieser Entwicklung keine Rolle gespielt.

Die kritische Bewertung ganzer Politiken mehr als die Evaluation einzelner Programme war auch die Aufgabe des britischen Programme Analysis and Review Verfahrens, das zunächst dem Central Policy Review Staff zugeordnet war, der wiederum dem Premierminister zuarbeitete, bevor das Verfahren auf die Treasury übertragen wurde. Die einzelnen Ministerien blieben jedoch von den Aktivitäten des Central Policy Review Staff relativ unbeeinflusst. Obwohl das Verfahren 1979 eingestellt und der Stab 1983 aufgelöst worden waren, reaktivierte die Regierung Thatcher die früheren Evaluationsbemühungen im Rahmen ihrer Betonung der Kosten-Effektivität, des „value for money“, und ihres allgemeinen managerialistischen Politikansatzes. Die sogenannte Financial Management Initiative von 1983 beinhaltete zwar weiterhin die Kostenkontrolle, verlangte aber auch von neuen Projekten und Programmen, daß sie anschließend evaluiert würden. In dieser Entwicklungslinie ist auch die Schaffung einer Management-Einheit im Jahre 1985 zu sehen, deren Aufgabe es ist, die Evaluation als Routineprozedur in den einzelnen Ministerien zu verankern. Parallel dazu hat das britische Parlament in den letzten Jahren die Evaluation stärker betont und das Public Accounts Committee, den Rechnungsprüfungsausschuß, gestärkt. Dies läßt sich vor allem daran ablesen, daß das 1983 anstelle des Exchequer and Audit Department eingerichtete National Audit Office diesem Rechnungsprüfungsausschuß zuarbeiten soll. Zur Zeit läßt sich allerdings noch nicht abschließend beurteilen, ob diese Institutionalisierungen eher eine Stärkung traditioneller

fiskalischer Kontrollen bewirken werden oder sich in Richtung auf ein hiervon unterscheidbares Evaluationssystem bewegen.

Historisch und systematisch betrachtet ähnelt die Entwicklung in Kanada am stärksten der in den USA beobachteten. Nicht nur führte Kanada 1969 das PPBS ein (und behielt es), das Finanzministerium (Treasury Board) errichtete auch eine spezielle Planungseinheit, der Evaluationsexperten angehörten, die die Versuche der Ministerien zur Evaluation ihrer Maßnahmen unterstützen sollten. Allerdings wurden diese frühen Bemühungen Ende der 70er Jahre in Kanada als im großen und ganzen ineffektiv angesehen. Wiederum war es das Parlament, das 1977 mit Unterstützung des Auditor General, des Obersten Rechnungshofes, die Regierung dazu drängte, ein eigenes Office of Comptroller General einzurichten. Diese Behörde erhielt eine spezielle Einheit für Programmevaluation, deren Aufgabe es ist, dieses Verfahren in der Bundesregierung als normalen Teil des Managementzyklus durchzusetzen. Organisatorisch ist an der kanadischen Institutionalisierung interessant, daß die Staatssekretäre der Ministerien (Deputy Ministers) für die Evaluation verantwortlich gemacht und in jedem Ministerium Evaluationseinheiten eingerichtet wurden. Gleichzeitig wurde mit der Einführung des Policy and Expenditure Management System im Jahre 1979 die Evaluation in den Prozeß der Haushaltsaufstellung eingebunden, indem dem Finanzministerium — ähnlich wie in Bonn die Möglichkeit gegeben wurde, Evaluationsergebnisse und Evaluationsplanung bei der Haushaltsaufstellung anzufordern. Für den Grad der Institutionalisierung der Evaluation in Kanada ist heute wie auch für die USA charakteristisch, daß es Regierungshandbücher gibt, in denen über die Methodik von Evaluationsstudien aufgeklärt wird.

Während die USA, Schweden und die Bundesrepublik ihre frühen Evaluationsbemühungen bis

heute fortgesetzt und z.T. ausgebaut haben, wenn auch auf unterschiedlichem Niveau, und während Kanada und Großbritannien Ende der 70er Jahre einen zweiten Anlauf unternommen haben, die Evaluation zu institutionalisieren, gibt es, wie gesagt, eine dritte Gruppe von Ländern, in denen die Praxis der Evaluation zunächst sehr begrenzt wie in Holland oder fast unbekannt wie in Dänemark und Norwegen war, bevor auch diese Länder sich in den Sog der zweiten Institutionalisierungswelle haben hineinziehen lassen. In den Niederlanden forderte schon 1980 eine Regierungsreform-Kommission und etwas später das Parlament, Evaluationen als regelmäßigen Bestandteil des Politikzyklus zu verankern. Die bürgerliche Regierung veranlaßte 1981, die Evaluation verfahrensmäßig als Bestandteil des jährlichen Budgetprozesses zu institutionalisieren. Wie in Großbritannien, so war auch hier das Grundmotiv, die Staatsaufgaben zu begrenzen und, wenn möglich, eine lange Tradition wohlfahrtsstaatlicher Interventionspolitik zu revidieren.

Die allgemeine philosophische Verbindung zu Großbritannien ist sogar noch deutlicher am Beispiel Norwegen zu erkennen, wo die Evaluationsfunktion 1985 ebenfalls als Teil einer Reform des Budgetsystems verankert wurde, nachdem 1982 eine allgemeine Produktivitätskampagne vorausgegangen war. Hinzu kommt in Norwegen wiederum, daß der Rechnungshof, der wie in den USA ausschließlich dem Parlament berichtet, die traditionellen haushaltsrechtlichen Richtigkeitsprüfungen um output-orientierte Kontrollen ergänzt hat. Als weiteren Beleg für die Entstehung einer zweiten Welle von Evaluationsbemühungen sei auf Dänemark hingewiesen, wo sich der Grundgedanke ebenfalls erst in den 80er Jahren ausbreitete. Auch Frankreich kann als Beispiel eines Landes gelten, das sich bis vor kurzem stark auf traditionelle rechtliche und budgetäre Kontrollinstrumente stützte, 1982 aber unter der sozialistischen

Präsidentschaft die Notwendigkeit der Evaluation betonte und 1983 eine Bestandsaufnahme durchgeführt hat. Selbst die Schweizer Bundesregierung hat 1987 einen Fond von mehreren Millionen Franken eingerichtet, um die Evaluationsforschung zu befördern.

Dieser grobe Überblick über die Entwicklung in einigen Ländern dürfte drei Trends erkennen lassen:

- In den 80er Jahren entsteht eine zweite Evaluationsbewegung, die sehr stark an Fragen des Output des Regierungshandelns, möglicherweise weniger an den realen Auswirkungen, den Impacts, orientiert ist. Selbst die meisten derjenigen Länder, die die Evaluation schon in den 60er Jahren etabliert hatten, nahmen einen neuen Anlauf mit diesem Verfahren.
- Zweitens läßt sich feststellen, daß die erste Evaluationsbewegung in den 60er und Anfang der 70er Jahre primär am Prozeß der Planung und der Programmentwicklung orientiert war und damit letztlich auf die Bedürfnisse der programmverantwortlichen Fachpolitiker in Regierung und Parlament zugeschnitten war, während die zweite Welle Ende der 70er Jahre eher zentralen politischen Bemühungen entspringt, die Berechtigung laufender Reformprogramme zu überprüfen, und in erster Linie am Budgetprozeß ausgerichtet ist.
- In jedem Fall gewinnt in diesem Zusammenhang die externe Kontrollinstanz, der Rechnungshof, an Bedeutung. Damit hat sich der Schwerpunkt der Evaluation in einer Reihe von Ländern in den 80er Jahren von einer vormals primär regierungs-internen Angelegenheit auf die Parlamente verschoben. Die Parlamente gewinnen dabei umso mehr an Bedeutung, wie ihnen die nationale Rechnungsprüfungsbehörde ausschließlich oder doch primär zuarbeitet.

2. Ursachen der Entwicklung

Wie läßt sich diese Entwicklung, wie lassen sich diese drei Trends erklären? Welches sind die Konstellationen, aus denen der unterschiedliche Reifegrad und die Entwicklungsrichtungen nationaler Evaluationssysteme erwachsen sind?

2.1 Veränderung der fiskalischen Situation

Das Entwicklungsmuster der Frühstarter: USA, Schweden, Bundesrepublik und Kanada dürfte zunächst einmal *ceteris paribus* in der volkswirtschaftlichen Situation und folglich den fiskalischen Bedingungen seine Ursache haben, unter denen die Regierungen wirken konnten. Die erste Gruppe von Ländern erhielt den Impuls zur Programmevaluation in einer Situation blühender Volkswirtschaften und wachsender Budgets, die die Regierungen in die Lage versetzten, kostspielige Sozialprogramme sowie Reformen im Bildungs- und Gesundheitswesen in Angriff zu nehmen. In diesem Kontext sind auch formale Planungssysteme entstanden, sei es, daß sie sich wie in der Bundesrepublik auf die mittelfristige Finanzplanung beschränkten, sei es, daß sie sogar versuchten, die Budgetseite mit der Programmseite zu integrieren, wie es in den USA, Schweden und Kanada der Fall war. In jedem Falle wurde die Evaluation entweder als logisch konsequenter Abschluß des Planungssystems betrachtet oder aber dadurch stimuliert, daß die neuen Interventionsprogramme neue Informationsbedürfnisse auslösten. Die Unsicherheit, mit der die z.T. innovativen oder zumindest mit innovativem Anspruch verfolgten Programme operierten, ließ das Bedürfnis nach Wirkungskontrollen entstehen. Die Evaluation hatte dabei primär die Funktion, bestehende und neue Programme effektiver zu gestalten. Nur gelegentlich waren Evaluationen erforderlich, um diese Programme gegen politische Kritik mit dem Nachweis ihrer Effektivität zu legitimieren.

Die Auswirkungen der völlig geänderten ökonomischen Situation Ende der 70er Jahre waren fast die gleichen: die Notwendigkeit, mit knappen Ressourcen zu haushalten, induzierte einen Innovationsstreß in verschiedenen Regierungen und führte zur Entwicklung von Techniken der Haushaltskürzung, des Cutback-Management. Entsprechend ist das Verfahren, über das während der zweiten Welle der Evaluationsbemühungen die Institutionalisierung betrieben wurde, typischerweise auch der Budgetprozeß gewesen: Produktivitätskampagnen oder Financial Management-Initiativen in Großbritannien, Kanada, Norwegen und den Niederlanden führten entweder zur Reaktivierung bestehender Evaluationsversuche wie in Kanada und Großbritannien, oder dieser neue Managerialismus institutionalisierte das Konzept der Evaluation erstmals, um die Ressourcenallokation innerhalb des Budgets zu rationalisieren wie in Norwegen und den Niederlanden. Typischerweise sind auch die Träger dieser zweiten Bewegung nicht die Programm-Manager in den Ressorts, sondern eher die Finanzminister und Rechnungshöfe als traditionelle Wächter des Budgets. Damit geht allerdings einher, daß die Erwartungen an die Evaluation und ihre Funktion innerhalb des Politikzyklus sich leicht verlagert haben. Anstatt Programme effektiver zu gestalten, besteht das Anliegen nun eher darin, ineffektive Programme zu erkennen und so zu legitimierbaren Budgetkürzungen zu gelangen.

2.2 Politische Konstellationen

Natürlich ist es nicht die ökonomische Situation per se, die politischen und organisatorischen Wandel hervorruft, sondern letztlich sind es Politiker und Regierungsparteien, die den Kurs des Staatsschiffes verändern. In den ökonomisch blühenden 60er Jahren regierten in den Ländern, in denen frühzeitig die Evaluation institutionalisiert worden war,

Reformparteien wie die Sozialdemokraten in der Bundesrepublik und in Schweden, die Demokraten in den USA und die Liberalen in Kanada; sie initiierten die Reform- und Interventionsprogramme und hatten offenbar auch eine starke Affinität zu den Sozialwissenschaften, deren methodisches Wissen für die Durchführung von Evaluationen benötigt wurde. Ein weiterer Beleg für diesen politischen Zusammenhang dürfte das Aufgreifen des Evaluationsgedankens in Frankreich nach dem Antritt des sozialistischen Präsidenten im Jahre 1981 sein.

Eher konservative Regierungen in Großbritannien zwischen 1970 und 74 sowie seit 1979, in Kanada 1979 und wiederum seit 1984, in Norwegen 1981–1986, in Dänemark seit 1982, in Schweden zwischen 1976 und 1982, in den Niederlanden fast ununterbrochen seit 1977 – diese konservativen Regierungen sahen sich nicht nur mit einer veränderten ökonomischen Situation konfrontiert, sondern benutzten auch das z.T. schon verfügbare Evaluationsinstrumentarium, um das Budget neu zuzuschneiden und dabei gelegentlich auch Reformprogramme der Vorgängerregierungen zu stützen, die sie von Anfang an mit Argwohn betrachtet hatten. Wo analytische Kapazitäten vorhanden waren, konnten die ins Amt kommenden konservativen Regierungen diese benutzen, um existierende Programme „analytisch zu delegitimieren“, wie Wagner und Wollman (1986) es ausgedrückt haben. Die Ausnahme von der Regel sind die konservativen Regierungen in Washington (seit 1980) und Bonn (seit 1982), die die Evaluation nicht intensiviert haben, sondern auf dem schon bestehenden Niveau weiterbetrieben, ja sie in den USA sogar etwas zurückgenommen haben.

Zu den politischen Rahmenbedingungen gehört auch der internationale neo-konservative Konsensus, daß der Staat sich aus der Gesellschaft zurückzuziehen habe, zumindest aber nicht weiter

expandieren solle und daß manch ein Programm seine Zielsetzung überschossen habe. In Verbindung mit Privatisierungsprogrammen dienten Deregulierung und Entbürokratisierung sowie das Zurückfahren der Planungssysteme, die zur Entwicklung und Steuerung von Reformpolitik gedient hatten, dazu, den Kurs zu ändern. Eine zweite Strategie bestand darin, das traditionelle Budgetverfahren zu verbessern, indem neue Kontrolltechniken z.T. aus dem Privatsektor übernommen wurden, die zumindest im Ansatz der Evaluationsmethodik sehr nahe kommen. Schließlich ließen sich modernisierte ökonomische Kontrolltechniken auch gut in die überkommenen Verfahren der Haushaltskontrolle einbinden.

2.3 Verfassungskontext

Eine Erklärung der unterschiedlichen Entwicklungswege, die die verschiedenen Zentralregierungen gegangen sind, kann sich nicht nur auf den Zusammenhang von wachsender Wirtschaft, Interventionsprogrammen und reformistischen Regierungen oder schrumpfender Wirtschaft, Sparpolitik und konservativen Regierungen stützen. Ein solches Modell wäre zu einfach. Es würde beispielsweise nicht erklären, warum in den USA unter der republikanischen Präsidentschaft Nixons eine Weiterentwicklung der Evaluation eintrat, noch ließe sich daraus erklären, warum sich hier in Bonn nach dem Regierungswechsel von 1982 so gut wie nichts am Stand der Evaluation änderte. Herausragender Bedeutung scheint drittens die Verfassungskonstellation zu haben, in der Exekutive und Legislative zueinander stehen. Meines Erachtens spielt bei der Institutionalisierung der Evaluation eine große Rolle, ob wir es mit einem System der Parteienregierung oder mit einem Präsidialsystem zu tun haben. In einem System des Party Government, in dem sich üblicherweise Mehrheitsfraktionen und Regierung einerseits sowie

Oppositionsfractionen andererseits gegenüberstehen, entwickelt das Parlament als Ganzes kaum das Bedürfnis, sich unabhängig von der Regierung eine Informationsbasis aufzubauen, in der die Evaluation ihren Platz hat. Zwar haben wir gesehen, daß in Kanada, Norwegen und Großbritannien sehr wohl die Parlamente Initiativen zur Stärkung der Evaluation entwickelt haben, in diesen Fällen haben die Parlamente aber lediglich bekräftigt, was die Regierung ohnehin wünschte. Insofern sind schon die Verhältnisse in der Bundesrepublik gegen Ende der großen Koalition, als an verschiedenen Punkten parlamentarische Evaluationsinitiativen zum Tragen kamen, charakteristisch. Zudem muß man feststellen, daß es im allgemeinen zwischen Regierungs- und Oppositionsparteien kaum unterschiedliche Auffassungen über die grundsätzliche Nützlichkeit der Evaluation gibt.

Der Fall der USA verweist auf die eminente Bedeutung, die das Präsidialsystem für die amerikanische Evaluationsentwicklung und wohl auch für die Situation in Frankreich hat. In beiden Ländern ist es möglich, daß sich parlamentarische Mehrheiten gegen einen Präsidenten bilden, die der Exekutive Evaluationspflichten aufzwingen können oder ihre informationelle Unabhängigkeit betonen. So ist zu verstehen, daß der amerikanische Kongreß sich einen gewaltigen Mitarbeiterstab als Gegenbürokratie aufgebaut hat und zu einem starken Pfeiler im amerikanischen Evaluationssystem wurde, obwohl die Exekutive von sich aus auf dem Gebiet der Evaluation frühzeitig aktiv geworden war. Parlamentarisches Interesse vermag dann die Evaluationskapazitäten der Exekutive sogar zu bewahren, wenn die Administration selbst nicht allzu enthusiastisch auf diesem Gebiet ist, geht es doch immer um die Erhaltung des informationellen Machtgleichgewichts.

Ein weiterer, hiermit zusammenhängender konstitutioneller Faktor, dem man einige Bedeutung

beimessen sollte, ist die unterschiedliche Zuordnung der zentralen Rechnungshöfe. In den meisten Ländern, so den USA, Kanada, Norwegen, Niederlande, Dänemark sind diese externen Kontrollinstanzen der Legislative zugeordnet. Während der deutsche Bundesrechnungshof historisch der Exekutive berichtete und erst seit 1969 seine Funktion auch und verstärkt am Bundestag und am Rechnungsprüfungsausschuß ausgerichtet hat, hat etwa der französische Cours des Comptes, obwohl ebenfalls dem Parlament berichtend, bis heute eine eindeutig exekutive Orientierung bewahrt, was mit dem Karrieresystem seiner Mitglieder zusammenhängt. In Schweden schließlich ist der Rechnungshof ausschließlich ein Instrument der Zentralregierung, und die parlamentarische Rechnungsprüfung verfügt hier nur über einen kleinen Mitarbeiterstab, um ad hoc Untersuchungen anzustellen. Man könnte generalisieren, daß die Parlamente die Kontrolle durch die Rechnungshöfe umso stärker als Machtbasis sogar in einem System der Parteienregierung nutzen, je unabhängiger die Finanzkontrolle von der Exekutive institutionalisiert ist oder je stärker sie sogar dem Parlament zugeordnet ist. In den USA konnte diese Anbindung des Rechnungshofes in Verbindung mit dem Präsidialsystem natürlich die Evaluationsfunktion im Kongreß nur noch weiter stärken.

2.4 Verfügbarkeit sozialwissenschaftlichen Wissens

Evaluationsmethoden und Programmforschung im allgemeinen sind nirgendwo das Kind der Verwaltung gewesen, sondern haben sich im Bereich der Sozial- und Wirtschaftswissenschaften entwickelt. Es spricht alles dafür, daß der Entwicklungsstand der Evaluationskultur in den betreffenden Regierungssystemen ganz entscheidend auch davon abhängt, ob die Wissenschaft eine Tradition in angewandter Forschung hat und sich von selbst mit der Evaluation befaßt und die Rolle

des Evaluators und allgemeiner: des Programmforschers professionalisiert hat. Unzweifelhaft reflektiert die führende Rolle der USA bei der Institutionalisierung der Programmevaluation als Verfahren und bei der Schaffung von Evaluationseinheiten im Regierungssystem die lange Tradition angewandter quantitativer Sozialforschung. Das Angebot von Seiten der Wissenschaft ist in den meisten anderen hier betrachteten Ländern nicht so breit, sei es, daß es keine ausgeprägte Tradition angewandter Forschung in den Universitäten gibt, sei es, daß die Professionalisierung der Sozialwissenschaften wie in der Bundesrepublik erst spät eingesetzt hat.

Die unterschiedlichen Ausgangsbedingungen in diesem Bereich dürften auch indirekt die Bereitschaft der Verwaltungen beeinflußt haben, den Gedanken der Evaluation und die methodischen Fertigkeiten für ihre Durchführung aufzunehmen, denn vom Vorhandensein einer längeren Tradition angewandter ökonomischer und Sozialforschung dürfte es auch abhängen, ob in nennenswerter Zahl derart vorgebildete Personen zumindest in administrative Stabspositionen gelangen. Eben dies war der Fall in den USA, Kanada und in Schweden: in Stockholm konnten schon in den 60er Jahren vor allem Ökonomen und Verwaltungswissenschaftler Führungspositionen erreichen und bildeten im übrigen das Rekrutierungsfeld für die Rechnungsprüfung, als die fiskalischen Wirksamkeitskontrollen eingeführt wurden.

Wenn der öffentliche Dienst – und die Politiker – hingegen primär ein juristisches Training wie in Frankreich, Dänemark und der Bundesrepublik oder gar eine Ausbildung in klassischen Studien (wie in Großbritannien) durchlaufen haben, dürfte der Regierungsapparat weniger rezeptiv für Sozialwissenschaftler und sozialwissenschaftliches Wissen sein.

Darüber hinaus ist sicherlich auch das Vorhandensein unabhängiger, nicht universitärer Forschungsinstitute mit Evaluationskapazitäten eher begünstigt worden in Ländern, die über eine Tradition angewandter Sozialforschung verfügen. Daß sich mit der Aufnahme von Evaluationsaktivitäten im Regierungssystem in den meisten Ländern derartige unabhängige Forschungsinstitute entwickelt haben oder zumindest gestärkt wurden, dürfte plausibel sein. Ohne Zweifel gehört dieser wissenschaftsimmanente Faktor in das Ursachenbündel, um Unterschiede sowohl im Angebot als auch, vermittelt über die Affinität der Verwaltung zu den Sozialwissenschaften, Unterschiede in der Nachfrage nach Evaluationen zu erklären.

2.5 Retardierende Faktoren

Es ist gleichermaßen interessant, sich die Frage zu stellen, welche Faktoren umgekehrt dazu beigetragen haben, daß der Gedanke der Evaluation und die Nachfrage nach dieser Form von Programmforschung sich erst spät entwickelt haben. Dazu gehören sicherlich funktionale Äquivalente wie das Vorhandensein einer guten amtlichen Statistik. Für Kontinentaleuropa und Skandinavien mit ihrer langen Staatstradition läßt sich sagen, daß die amtliche Statistik seit langem zum Rückgrat der öffentlichen Verwaltung gehört hat, so daß man sich hierauf bei ökonomischen und politischen Analysen üblicherweise bis zu einem gewissen Grade stützen kann. Dies wird besonders deutlich am Beispiel der Niederlande mit ihrer langen Tradition öffentlicher Planung-, die in Tinbergens Konzeption der 30er Jahre wurzelt. Andererseits ist es bekanntlich das Fehlen zuverlässiger amtlicher Statistiken in den USA gewesen, das den Aufstieg angewandter Forschung erforderlich machte und begünstigte.

Ein weiteres Äquivalent, das für lange Zeit den wachsenden Evaluationsbedarf befriedigt haben dürfte, ist das Vorhandensein eines Beiratssystems für die Regierungen. Dies wird besonders deutlich am Beispiel Schwedens, aber auch Dänemarks und der Niederlande. Ein derartiges Beiratssystem scheint, wenn schon nicht umfassende Evaluationen, so doch zumindest Vollzugsprobleme dokumentieren zu helfen, und daran sind Verwaltungsleute ja nicht selten am meisten interessiert.

In diesem Zusammenhang sei auch die Verwaltungsgerichtsbarkeit oder die Tradition des Ombudsman in Skandinavien oder aber der investigative Journalismus, wie wir ihn in den USA und Schweden finden, als ein funktionales Äquivalent erwähnt, das in Verbindung mit weiteren Randbedingungen, aber keineswegs natürlich ausschließlich, eine verzögerte Entwicklung erklären könnte.

Überlegenswert wäre auch, ob sich die Evaluation unter den Bedingungen föderaler Systeme wie in den USA, der Bundesrepublik und Kanada tendenziell eher entwickelt als im unitarischen Staatsaufbau, wo die Zentralregierung tendenziell eher Vollzugsinformationen abrufen kann. Auch der Umfang der Staatsaufgaben und in föderalen Systemen: der Umfang zentralstaatlicher Aufgaben dürfte eine Rolle bei der Bestimmung des Evaluationsbedarfs spielen, wie dies schließlich auch die Größe des Landes und die administrative Distanz tun.

Wir haben es offenbar mit einem multifaktorellen Verursachungszusammenhang zu tun, aus dem sich die von Land zu Land unterschiedlichen Entwicklungsmuster von Evaluationssystemen erklären lassen. Das Hauptgewicht dürfte dabei allerdings auf den fiskalischen, den politischen, den konstitutionellen und den wissenschaftsimmanenten Ursachen liegen.

3. Promotoren der Institutionalisierung

Ich möchte der Frage nach der Initiierung von Evaluationsbemühungen und ihrer Verfestigung bis zur Institutionalisierung noch einmal nachgehen, indem ich die bisher verfolgte, vom Kontext der Regierungsapparate ausgehende historische Betrachtungsweise verlasse und untersuche, welche Akteure im politisch-administrativen System sich als zentrale Promotoren der Entwicklung ausmachen lassen und welche Unterschiede in dieser Hinsicht zwischen den betrachteten Ländern bestehen. Denn selbst wenn der Bedarf an Evaluationen in einem System anerkannt ist, bedeutet dies keineswegs, daß er auch schon befriedigt wird. Es ist geradezu eine anthropologische Konstante, daß Menschen es nicht lieben, in irgendeiner Form überwacht zu werden, und daher der Institutionalisierung von Kontrollinstrumenten widerstehen werden, weil sie negative Sanktionen insbesondere für ihre Karrieren erwarten, wenn Fehler dokumentiert werden. Hinzu kommt, daß externe Kontrollen definitionsgemäß die vorhandene Machtbalance zwischen Organisationen stören und daher politische Widerstände zu erwarten sind.

3.1 Parlamente

Parlamente haben die Exekutive in einer Reihe von Ländern gedrängt, die Programmevaluation einzuführen. In der Bundesrepublik waren es parlamentarische Initiativen im Bereich der Entwicklungshilfe und der Technologiepolitik in den 60er Jahren, die sich an die Exekutive richteten, ohne daß der Bundestag allerdings selbst seinen wissenschaftlichen Dienst auf die Evaluationsfunktion ausgerichtet hätte. Das kanadische Parlament griff 1977 die Initiative des Rechnungshofs auf und ermächtigte ihn, die Evaluationsbemühungen der Exekutive zu überwachen. Auch in Großbritannien stärkte, wie gesagt, das Parlament

seit 1979 das System der Untersuchungsausschüsse. In allen diesen Fällen ging es darum, die Exekutive dazu zu bewegen, Evaluationen zu produzieren. In den USA hingegen institutionalisierte der Kongreß nicht nur Evaluationspflichten für die Administration, sondern entwickelte seine eigenen Evaluationskapazitäten. Auch dort, wo der Ansatz für die Verankerung der Evaluation primär budgetärer Natur war, wie in den Niederlanden und Norwegen, waren es die Parlamente, die im Zuge einer Haushaltsreform die Evaluationsfunktion – allerdings in voller Übereinstimmung mit den betreffenden Regierungen – betont haben.

3.2 Rechnungshöfe

In den meisten Ländern, insbesondere in denen, die der Evaluationsbewegung erst auf der zweiten Welle gefolgt sind, spielten auch die Rechnungshöfe eine aktive Rolle. Dies gilt für Schweden bereits im Jahre 1967, für den amerikanischen Rechnungshof sei die Etablierung des Evaluationsamtes im Jahre 1980 in Erinnerung gerufen, und der kanadische Rechnungshof veranlaßte das Parlament, die Evaluationsfunktion des Rechnungshofs seit 1976 zu verstärken. Für Großbritannien gilt ebenfalls, daß mit der Etablierung des National Audit Office im Jahre 1983 die Finanzkontrolle in Richtung Evaluation ausgebaut wurde. In den Niederlanden und in der Bundesrepublik (Rürup/Färber 1985, 190 f.) haben die Rechnungshöfe jedoch, wie wir gesehen haben, eher eine reaktive Rolle gespielt. Obwohl sie berechtigt und verpflichtet sind, die Regierungsaktivitäten auf Effizienz und Effektivität zu kontrollieren und den Stand der Evaluation in der Exekutive kritisch zu prüfen, darf man wohl sagen, daß das traditionelle Kontrollverständnis, das sich in erster Linie auf haushaltsrechtliche Konformität und bestenfalls Sparsamkeit richtete, immer noch dominant ist. Nichtsdestoweniger scheint ihr

Rollenverständnis im Wandel begriffen zu sein, und zwar in dem Maße, wie andere Akteure des politischen Systems Steigerung der Wirtschaftlichkeit als politisches Ziel verfolgen.

3.3 Kabinett

Was läßt sich zur Funktion der Kabinette bei der Institutionalisierung der Programmevaluation sagen? Einige Kabinette entwickelten und verfolgten vor allem eine allgemeine Produktivitätssteigerung im öffentlichen Sektor als programmatischen Punkt, so die breite holländische Initiative von 1981, die norwegische Produktivitätskampagne 1982 oder Großbritanniens Financial Management Initiative von 1983. Diese Fälle gehören der zweiten Welle an, die u.a. vom fiskalischen Zwang motiviert wurde. Die Frühstarter der 60er Jahre hingegen zeichnen sich dadurch aus, daß zentrale politische Initiativen, Evaluationskapazitäten aufzubauen, fehlten, weil das vorrangige Interesse der Programmentwicklung im Rahmen der Reformpolitik galt. Die Evaluation entwickelte sich hier z.T. aus dezentralen Initiativen der Programmverantwortlichen in den Fachministerien. Während in jenen Jahren die Programmevaluation als unpolitische technische Angelegenheit angesehen wurde, wurde sie in den 80er Jahren in Zusammenhang mit budgetpolitischen Zielsetzungen ein Anliegen hoher politischer Priorität und zog mithin die Aufmerksamkeit einiger Kabinette auf sich.

3.4 Zentrale Regierungseinheiten

Das bedeutet natürlich nicht, daß man erst während der zweiten Welle zentrale Regierungseinheiten vorfindet, die sich mit der Evaluation befassen. Das amerikanische Office of Management and Budget hatte bereits 1970 die Rolle übernommen, die Ressorts mit den Evaluationstechniken bekannt zu

machen, soweit dies nicht ohnehin schon der Fall war. Erinnert sei auch an den britischen Central Policy Review Staff und das Verfahren der Policy Analysis and Review sowie die spezielle Effizienz-Einheit unter Derek Rayner nach 1979. Kanada ist wohl das hervorstechendste Beispiel für eine relativ spät gebildete zentrale Einheit mit dem 1978 eingerichteten Office of Comptroller General. Auch diese Einheit verfolgt die Aufgabe, die Praxis in den Ressorts voranzutreiben. Man könnte hinsichtlich der Funktion derartiger zentraler Einheiten, die regierungsweit operieren, verallgemeinernd sagen, daß ihre Aufgabe in erster Linie darin bestand und besteht, die Institutionalisierung voranzutreiben, Richtlinien zu entwickeln und die Ressorts praktisch zu unterstützen. Ihre Aufgabe besteht jedoch nirgends – und das gilt natürlich auch für die Bundesregierung, hier das Bundeskanzleramt – darin, Programme in den verantwortlichen Ressorts selbst zu evaluieren. Offenbar hat man in allen Ländern erkannt, daß ein derartiger Eingriff in die z.T. formal bestehende Ressortautonomie das Konfliktniveau insbesondere in Koalitionsregierungen hinauftreiben würde.

Zu den zentralen Regierungseinheiten könnte man natürlich auch dasjenige Amt oder Ministerium rechnen, das mit der Aufstellung des Regierungshaushalts befaßt ist. Die organisatorische Position des amerikanischen Office of Management and Budget, das dem Präsidenten unmittelbar zugeordnet ist, bietet natürlich eine besonders günstige Ausgangsbasis, um die Evaluation in den Ressorts zu kultivieren. Abgesehen von der formalen hierarchischen Position dieser Behörde ist von wesentlicher Bedeutung, wie stark Evaluationsergebnisse bei der Haushaltsaufstellung Berücksichtigung finden. Finanzministerien können im Prinzip denselben Einfluß auf die Evaluation in den Ressorts gewinnen, wenn deren Ergebnisse verfahrensmäßig in den Budgetprozeß einfließen. Dies gilt z.B. für das holländische Finanzmi-

nisterium, das Vertreter in Stabsgruppen der Ministerien entsendet, um an den Berichten für die Programmrevision mitzuwirken.

Es scheint ein Vorteil der jüngeren neo-konservativen Management-Orientierung gegenüber der älteren Programm-Orientierung der Evaluation zu sein, daß genau dieser Zusammenhang zwischen Evaluation und Budgetierung hergestellt wird und damit die Evaluation in den Fachressorts verstärkt wird. Allerdings bleibt abzuwarten, welche Auswirkungen diese Verbindung von Evaluation und Budget auf den Inhalt und die methodische Qualität von Evaluationen hat. Der Zeitdruck der Budgetaufstellung und die Kostenorientierung könnten dazu führen, daß das methodische Anspruchsniveau an die Evaluationen abnimmt, ja daß möglicherweise im Budgetprozeß legitimatorische Motive die überhandnehmen und methodisch schwer überprüfbare Erfolgsberichte produziert werden.

Anmerkungen

Der Vortrag basiert auf meiner umfassenderen Untersuchung „Genesis and Structure of Evaluation Efforts in Comparative Perspective“, die erscheinen wird in: Ray Rist (ed.), Program Evaluation and the Management of Government: Patterns and Prospects across eight Nations

Hans-Ulrich Derlien, Notwendigkeit und Grenzen der Erfolgskontrolle öffentlicher Interventionsprogramme, in: Eberhard Wille (Hg.), Konkrete Probleme öffentlicher Planung, Frankfurt u.a. 1986, 291–308

Bert Rürup/Gisela Färber, Programmbudgets der „zweiten Generation“, in: Die öffentliche Verwaltung 33 (1980), 661–672

Bert Rürup/Gisela Färber, Kontrollorientierte Ansätze einer Budgetreform, in: Die Verwaltung 18 (1985), 173–200

Peter Wagner/Helmut Wollmann, Fluctuations in the development of evaluation research: do „regime shifts“ matter?, in: International Social Science Journal 1986, 205–218

Herr Furmaniak

Gibt es Fragen zum Referat von Professor Derlien?

Herr Lorenzen

Sie haben die Wünsche geschildert in vielen Ländern, Evaluation durchzuführen. Können Sie ein bißchen ergänzen, wie weit dieser Wunsch auch durch das Können realisiert werden konnte. Wieweit also die Enttäuschung über das, was man eigentlich gerne gewollt hätte, aber dann doch nicht leisten konnte, hier und da zu einem gewissen Rückschlag geführt hat.

Herr Derlien

Eine gute Frage, Herr Lorenzen. Sie haben sofort den Schwachpunkt meiner Ausführungen entdeckt. Ich stütze mich bei dem, was ich gesagt habe, auf Länderberichte, die im Rahmen des Internationalen Instituts für Verwaltungswissenschaften erarbeitet worden sind in einer Arbeitsgruppe, der Ray Rist als amerikanischer und ich als deutscher Repräsentat angehören. Nun ist es so, daß wir die Qualität des methodischen Geschäfts in den Ländern ganz gut beurteilen können, die der ersten Welle angehören. Da liegen ja inzwischen 15 Jahre Erfahrung vor. Das trifft etwa auf die USA und die Bundesrepublik zu. Es ist aber außerordentlich schwierig, schon etwas über die Qualität der Berichte und die Präsentation der Untersuchungsergebnisse auszusagen in den Ländern, die erst Anfang der 80er Jahre angefangen haben; etwa in Holland ist es schwer auszumachen und in Großbritannien ebenfalls. Ich muß da leider passen. Ich habe, glaube ich, auch angedeutet, daß ich mir hier im einen oder anderen Fall gelegentlich die Frage stelle, ob, gerade auch wegen der engen Verzahnung mit dem Budgetprozeß, das Ganze nicht stärker in Richtung

auf eine Kostenträgerrechnung hinausläuft, gerade wenn man also im britischen System nach value for money schreit. Da liegt es ja nahe, daß man die Kosten pro Dienstleistungseinheit oder pro Verwaltungsakt zu ermitteln versucht. Aber wie gesagt, da die Entwicklung noch im Laufen begriffen ist, kann sowohl das eine wie das andere geschehen. Ich kann das nicht abschließend beantworten.

Herr Ernst

Ich wollte nur Ihre Aufmerksamkeit lenken auf eine Entwicklung, die sich in der europäischen Gemeinschaft vollzieht. Ich weiß nicht, wie weit die Entwicklung genau ist, möchte aber darauf hinweisen, daß die europäische Gemeinschaft eine task force haben, um die Förderung kleiner und mittlerer Unternehmen durch EG-Vorhaben im Hinblick auf Erfolg zu evaluieren.

Herr Derlien

Das ist interessant, Herr Ernst, ich darf das vielleicht so aufnehmen, daß selbst das britische System, das ich etwas der mikroökonomisch-betriebswirtschaftlichen Betrachtung verdächtige, möglicherweise, wie dieser EG-Versuch zeigt, doch auf die Linie dessen, was Evaluationsstandard ist, einschwenkt.

Herr Furmaniak

Ich glaube, wir müssen jetzt, das ist der technische Zwang durch Maschinen, die Diskussion abbrechen, da die Satellitenübertragung beginnt.