

Möglichkeiten und Grenzen von Evaluation in der Verwaltungspolitik

Klaus-Peter Jacoby, Universität des Saarlandes

Zusammenfassung einer Podiumsdiskussion, die im Rahmen der Jahrestagung der Deutschen Gesellschaft für Evaluation (DeGEval) am 5.10.2001 an der Deutschen Hochschule für Verwaltungswissenschaften in Speyer stattfand.¹

Podiumsteilnehmer:

Prof. Dr. *Rudolf Fisch* (Rektor der Deutschen Hochschule für Verwaltungswissenschaften in Speyer), Moderation

Peter Seyfried (Ministerialdirigent, Innenministerium Baden-Württemberg)

Dr. *Rudolf Dieckmann* (Präsident a.D., Rechnungshof der Freien und Hansestadt Hamburg)

Prof. *Gerhard Banner* (Vorstand a.D. der Kommunalen Gemeinschaftsstelle für Verwaltungsvereinfachung – KGSt)

PD Dr. *Jörg Bogumil* (Fern-Universität Hagen)

Prof. Dr. *Hermann Hill* (Deutsche Hochschule für Verwaltungswissenschaften Speyer)

Prof. Dr. *Hellmut Wollmann* (Humboldt-Universität zu Berlin)

In einer ersten Runde der Podiumsdiskussion hatten die Teilnehmer Gelegenheit, in einleitenden Statements aus ihrer jeweils spezifischen Perspektive (öffentliche Verwaltung, Wissenschaft etc.) die zentralen Themen zur Diskussion zu stellen und ihre jeweiligen Position in allgemeiner Form abzustecken.

Herr *Seyfried* (Innenministerium Baden-Württemberg) nähert sich dem Themenfeld der Evaluation im Zusammenhang mit der Verwaltungsreform im Land Baden-Württemberg, die er im dortigen Innenministerium verantwortlich mitgestaltet. Sein Vortrag konzentrierte sich auf drei Thesen:

¹ Die Zusammenfassung basiert auf der wörtlichen Transkription eines von der Deutschen Gesellschaft für Evaluation zur Verfügung gestellten Tonbandmitschnittes. Dieser wurde von dem Autor im Rahmen des vorliegenden Artikels redaktionell bearbeitet.

- Evaluation gehöre zu den „Kunstregeln der Reform“, spiele sich aber überwiegend unterhalb der Wahrnehmungsschwelle der Verwaltungspolitik ab
- Evaluation in der öffentlichen Verwaltung begegne außerordentlichen methodischen Schwierigkeiten. Insbesondere im Hinblick auf betriebswirtschaftliche Kriterien fehle eine gesicherte Basis, die sich erst mit der flächendeckenden Einführung des Neuen Steuerungsmodells herausbilden werde.
- Schließlich fehlten meist die für die Durchführung methodisch anspruchsvoller Evaluationen erforderlichen Ressourcen.

Bei allen genannten Punkten handle es sich kontextunabhängig um grundsätzliche Problemfelder innerhalb öffentlicher Verwaltungen. Zudem seien insbesondere in der Bundesrepublik Deutschland die spezifischen Gegebenheiten eines föderalen Systems zu beachten. Infolge der ausgeprägten Autonomie, die die verschiedenen Politik- und Verwaltungsebenen im bundesstaatlich-dezentralen System der Bundesrepublik nicht zuletzt in ihrer Verwaltungspolitik besitzen, entwickeln sich unterschiedliche Verwaltungstraditionen, die einerseits die Entwicklung einheitlicher Instrumente verhindern und andererseits bemerkenswerte verwaltungspolitische Eigeninitiativen auch und gerade der Bundesländer eröffnen.

Im Hinblick auf die Landesverwaltung in Baden-Württemberg sieht Herr Seyfried in jüngerer Zeit bemerkenswerte Fortschritte. Seit 1988 existiert eine Stabsstelle für Verwaltungsreform im Innenministerium, die Reformprogramme nicht nur formuliere, sondern auch, zumindest mit einfachen Instrumenten und Verfahren, evaluiere. In allen Einzelprojekten dieses Reformprogramms seien Berichtschleifen vorgesehen, die zu Meilensteinen oder am Ende des Projekts zu einem Bericht über die beobachtbaren Erfolge verpflichten. Besonders hervorzuheben sei auch die Durchführung von Pilotprojekten im Personal- und Organisationsentwicklungsbereich, deren Pilotcharakter die Durchführung von Evaluationen zwingend erfordere. Trotz solcher positiver Beispiele für interne Evaluierung weist Herr Seyfried aber auch auf das weitgehende Fehlen externer Evaluierungen hin, das in erster Linie auf die knapp bemessenen Budgets zurückzuführen sei. Kompensiert werde dieses Defizit aber teilweise durch eine Reihe externer Kooperationen (so z.B. mit der Verwaltungshochschule Speyer, mit andern Bundesländern, dem Bund), durch die die methodische Kompetenz des Ministeriums gestärkt werde.

Gegenüber der eher pessimistischen Einschätzung seines Vorredners zur Möglichkeit externer Evaluierung äußerte sich Herr *Dieckmann* (Hamburger Rechnungshof) verhalten optimistisch. Zwar seien die Möglichkeiten der externen Evaluation noch nicht ausgeschöpft, jedoch sei gegenüber den 1970er und 1980er Jahren eine deutlich positive Entwicklung zu konstatieren. So sei das in § 7 der Hamburger Landeshaushaltsordnung vorgeschriebene Wirtschaftlichkeitsprinzip seit 1996 dahingehend konkretisiert worden, dass neben den Instrumenten Nutzen-Kosten-Untersuchungen (Abs. 2) und Kosten- und Leistungsrechnung (Abs. 4) vor der Durchführung von Maßnahmen mit finanzieller Bedeutung grundsätzlich deren Zielsetzung zu bestimmen sei. Während und nach ihrer Durchführung sind diese Maßnahmen auf Zielerreichung, Wirksamkeit und Wirtschaftlichkeit zu überprüfen (Erfolgskontrolle). Auslöser für diese Gesetzesinitiative waren vom Rechnungshof festgestellte Defizite an Erfolgskontrollen und von ihm selbst vorgenommene Wirkungsuntersuchungen, vor allem im Umwelt- und Wirtschaftsbereich. Sie wurden

vom Parlament mit einer Initiative gegenüber der Regierung aufgegriffen und unterstützt. Die Verwaltung erstellte einen Leitfaden Erfolgskontrolle (1991) und Hinweise auf die Planung und Durchführung von Erfolgskontrollen im Anwendungsbereich (1997).

Gleichwohl sei zu berücksichtigen, dass gesetzgeberischen Maßnahmen nicht unbedingt die erwünschten Wirkungen auf der Verwaltungsebene folgen müssen. In Hamburg habe es für die allgemein positive Entwicklung der letzten Jahre klar zu benennende Voraussetzungen gegeben, die auch für die weiterführende Etablierung einer Evaluierungstradition maßgeblich seien:

- die Entwicklung von Zielbildern für öffentliche Unternehmen, sowie über diese Zielbilder die Schaffung einer konkreten Messlatte für die Tätigkeit der Unternehmen;
- das Vorliegen messbarer Zielgrößen und messbarer Daten für die Wirkungen öffentlicher Maßnahmen;
- die unterstützende Wirkung des Neuen Steuerungsmodell mit seiner Produkt- und Kennzahlenbildung und seiner Forderung, Ziele zu setzen, die mit dem tatsächlichen Erfolg verglichen werden können;
- der Problemdruck einer Haushaltskrise, die notwendig die stärkere Bereitschaft weckte, Programme und Maßnahmen auf ihren Erfolg, ihre Wirtschaftlichkeit und Wirksamkeit zu bewerten;
- Ansprüche des Parlaments, das Evaluation als notwendiges Mittel zur Erfüllung seiner Kontrollfunktion einforderte. Hierzu gehöre neben einer Kontrolle des Aufgabenvollzugs durch die Exekutive auch eine, von Regierung und Parlament gleichermaßen voranzutreibende Aufgaben(zweck)kritik.

In Hamburg seien inzwischen sowohl verwaltungsinterne Erfolgskontrolle (interne Beteiligungsverwaltung, Abteilungen für Revision und Controlling) als auch externe Erfolgskontrolle in zunehmenden Maße institutionalisiert. Der Rechnungshof habe hier bei der Wahrnehmung seiner Rolle einen Schwenk von der titelbezogenen Ausgabenkontrolle und vom Aufgabenvollzug hin zu einer politischeren Aufgabekontrolle vollzogen.

Sehe man sich beispielhaft die Jahresberichte des Hamburger Rechnungshofs der letzten Jahre an, dann finden sich bis zu 50 Prozent der Prüfungen, die unter Erfolgskontrolle einzuordnen seien. Zu den Prüfungsthemen gehören u.a.

- Bewertung der Ziele und Umsetzung der einzelnen Elemente der Haushaltskonsolidierung (2000),
- Bewertung der Ziele und Umsetzung der einzelnen Elemente des Neuen Steuerungsmodells nach 5 Jahren Reformprozeß (2000).

Vielfach fehle den Parlamenten in dem Gefüge des Neuen Steuerungsmodells allerdings noch die Orientierung, insbesondere im Hinblick auf die Erfordernisse von Projektbeschreibungen und Kennzahlen für die Erfolgskontrolle. Hier komme es deutlich zu einem Wandel der Tätigkeit des Rechnungshofes, hin zu einer stärker beratenden Funktion. Mit dem Ziel, die Selbstkontrolle innerhalb der Verwaltung gegenüber der externen Fremdkontrolle in den Vordergrund zu stellen, seien in erster Linie Lernprozesse zu unterstützen. Allerdings befänden sich auch die Rechnungshöfe selbst noch in einem Lernprozess und seien mit einer integralen Evaluie-

rung der Verwaltungsreformen überfordert. Maßnahmen der Erfolgskontrolle seien daher in erster Linie projekt- oder sachgebietspezifisch.

Vor dem Hintergrund von Erfahrungen bei der Umsetzung von Verwaltungsreformen in verschiedenen deutschen Gemeinden kommt Herr *Banner* (Vorstand a.D. der KGSt) zu einer kritischeren Einschätzung der aktuellen Lage. Zwar sieht auch er die von seinen Vorrednern beschriebenen Fortschritte auf vielen Ebenen. Er geht allerdings von einem engeren und strengeren Begriff der Evaluierung aus. Er beziehe und beschränke das Konzept der Evaluierung auf *extern und unabhängig* durchgeführte Untersuchungen – in Abgrenzung von internen Verfahren wie Controlling, Kosten- und Leistungsrechnung, Berichtswesen etc. Lediglich im Bereich der interkommunalen Leistungsvergleiche gebe es Grenzfälle, die im Falle einer öffentlich transparenten Behandlung der Ergebnisse dem Anspruch von Evaluierung (im engeren und strengeren Sinne) gerecht würden. Ein Ausbau der externen Evaluierung sei dringend notwendig, wobei neben dem Steuerungspotential der Ergebnisse vor allem auch die Vorauswirkungen von Evaluationen auf die verbesserte Durchführung von Verwaltungsreformen relevant seien.

Die mangelhafte Institutionalisierung der externen Evaluierung führt er auf deren „Hybridnatur“ zurück. Sei sie auf der einen Seite ein sozusagen „naturwissenschaftliches“ Analyseverfahren, um Transparenz zu schaffen, so sei sie auf der anderen Seite vor allem auch ein politisches Instrument. Der Verweis auf knappe Budgets sei dabei in den seltensten Fällen hinreichend, sondern werde gewöhnlich induziert durch die Frage: Wer ist interessiert an Evaluation? Pointiert merkte Herr *Banner* an, ein aktives Interesse zeige vor allem, wer sich bereits vorher positiver Ergebnisse sicher sein und sich somit auf ein positives Echo der Medien verlassen könne. In allen anderen Fällen aber habe Evaluation mit Zurückhaltung und Widerstand zu rechnen. Auf einen einfachen Nenner gebracht, stehe in der politischen Logik das „gut aussehen“ vor dem „gut sein“. Das Steuerungsziel von Evaluationen trete dabei in den Hintergrund und es komme notwendig dazu, dass jedem Berichtswesen ein schönender Faktor innewohne.

Darüber hinaus seien vielfach wesentliche Voraussetzungen für die Durchführung von Evaluationen nicht gegeben. Aussagekräftige Evaluationen könnten in der Regel nur durchgeführt werden, wenn die Ziele einer Reform klar und differenziert genug seien. In der Realität aber würden politische Ziele nur sehr allgemein formuliert und selten strikt durch zeitliche und inhaltliche Meilensteine gestützt. Ein wichtiger Grund für diese Passivität der Verwaltungsführung im Hinblick auf die Verwaltungsmodernisierung liege darin, dass man dem politischen Gegner nur ungern eine Messlatte zur Verfügung stelle, die zum Nachweis des vermeintlichen eigenen Versagens benutzt und instrumentalisiert werden könne. Aus diesem Grunde seien flexible, weiche und unverbindliche Zielformulierungen politisch zumeist opportuner.

Zusammenfassend kam Herr *Banner* zu einer skeptischen Einschätzung im Hinblick auf die Zukunft extern und unabhängig durchgeführter Evaluierungen. Die fragmentierte Staatsarchitektur, die föderal stark zersplitterten Zuständigkeiten seien kein gutes Pflaster für Verwaltungspolitik und eine zielorientierte Verwaltungsreform. Zugespitzt lautete seine These: „Es gibt in der Bundesrepublik keine Verwaltungspolitik“, die die unterschiedlichen Verwaltungsebenen zusammenführe, dies insbesondere im internationalen Vergleich mit Staaten wie England oder Neu-

seeland. Schließlich spiele hier auch die traditionelle Grundlage der Verwaltungsphilosophie eine Rolle: Leistung der öffentlichen Verwaltung sei jahrzehntlang rechtliche Richtigkeit gewesen. Alle anderen Leistungskriterien wie etwa Wirtschaftlichkeit, Bürgernähe, inhaltliche Qualität, träten erst seit kurzem ins Bewusstsein.

Mit gebotener Vorsicht seien jedoch auch Faktoren auszumachen, die zu einem Ausbau der Evaluation in der öffentlichen Verwaltung führen könnten. So sei es denkbar, dass gerade die oft als Gegenargument genannten knappen Kassen die Rolle eines positiven Katalysators übernähmen, da sie den Zwang zu einer stärker zielgerichteten Mittelallokation erhöhten. Die methodischen Konzepte und Instrumente lägen zudem bereit, und auf internationaler Ebene gebe es zunehmend positive Beispiele, die den Handlungsdruck auch auf die deutschen Verwaltungen verstärkten. Die Zukunft der Evaluation hänge schließlich von Effekten der Gewöhnung und Routinisierung ab, von dem damit verbundenen Nachlassen des medialen und politischen Druck auf die, die nach einer Evaluation nicht so gut aussähen. Es werde dann eine gewisse Sachlichkeit und Gewöhnung an Evaluierung an Stelle einer heute noch üblichen Schwarz-Weiß-Malerei treten.

Eine anhand unterschiedlicher Formen der Evaluierung differenzierte Einschätzung der Situation in deutschen Verwaltungen gab Herr *Bogumil* (Fern-Universität Hagen). Dabei bezog er sich auf die Begriffe der *internen Evaluation*, auf *benchmarking* sowie auf *externe Evaluation*, bei denen hinsichtlich ihrer Möglichkeiten und Grenzen klare Unterschiede festzustellen sind.

- *Interne Evaluation*, beispielsweise in Form von Controlling, stelle einen Kernbestandteil des Neuen Steuerungsmodells dar. Insbesondere zur Sicherung der Funktionsfähigkeit dezentraler Strukturen sei interne Evaluation zur Steuerung unverzichtbar. Zwar habe es im Hinblick auf Instrumente wie Produktbeschreibung, Kosten- und Leistungsrechnung, verschiedene Formen des Berichtswesens etc. in den letzten Jahren spürbare Impulse gegeben, allgemein könne aber von einem funktionsfähigen Controlling in deutschen Kommunalverwaltungen noch nicht die Rede sein. In einigen Fällen hätten Umstrukturierungen der Verwaltung sogar zu einer Abnahme der strategischen Steuerungsfähigkeit in den Kommunen geführt. Auch die Verfügbarkeit sachgebietspezifischer Produktkataloge, Kosten- und Leistungsrechnungen etc. gewährleiste nicht zwangsläufig die Verwirklichung zentraler Interessen. Hierfür seien teilweise auch methodische Probleme ausschlaggebend. So seien zwar klar umrissene Einzelfragen durch quantitative Indikatoren darstellbar, als unlösbar stelle es sich dagegen häufig dar, qualitative Prozesse auf der Makro-Ebene nachzuzeichnen, die aber zum Teil gerade für eine strategische Steuerung entscheidend seien.
- *Benchmarking*, d.h. der Leistungsvergleich zwischen einzelnen Kommunen, spiele in der Modernisierungsdiskussion die wichtige Rolle eines *Leistungsverstärkers* auf dem Weg in das Neue Steuerungsmodell. Positiv hervorzuheben sei, dass sich bereits eine Vielzahl von Initiativen herausgebildet hätten (so z.B. der Speyer-Preis, das IKO-Netz der KGSt und andere deutsche wie auch internationale Netzwerke). Die Bereitschaft zum interkommunalen Vergleich habe in letzter Zeit deutlich zugenommen, gleichzeitig aber fehle es bislang an Transparenz und Öffentlichkeit, mithin an eigentlichem Wettbewerbsdruck. Bleibe

- die Rezeption der Ergebnisse aber verwaltungsintern, werde auch das vermeintliche benchmarking wieder auf eine Form von Selbstkontrolle reduziert. Aus diesem Grunde seien in Zukunft verstärkt Wege zu suchen, Leistungsvergleiche zu institutionalisieren und Transparenz und Öffentlichkeit zu gewährleisten. Dies laufe, zugespitzt, auf ein Pendant zur Stiftung Warentest für den öffentlichen Bereich hinaus, das keine Möglichkeit des Ausweichens zulasse und damit einen tatsächlichen Wettbewerbsdruck mit sich bringe.
- In Übereinstimmung mit seinen Vorrednern stellte Herr Bogumil fest, dass es sich bei der *externen Evaluierung* um das am wenigsten entwickelte Aufgabefeld handele. In Übereinstimmung mit Herrn Banner führte er dies vor allem auf die Logik eines im wesentlichen konkurrenzdemokratisch organisierten Systems zurück. Es gebe einen klaren Zielkonflikt zwischen dem Interesse der Verwaltungsspitzen an einer internen Verbesserung ihrer Ergebnisse und dem Zwang, diese mit den verschiedenen Fraktionen diskutieren zu müssen, da letztere ein natürliches Interesse an der politischen Verwertung negativer Ergebnisse hätten. Finanzielle Engpässe seien gegenüber dem mangelnden Interesse an der Durchführung von Evaluationen nachrangig. Externe Evaluation lasse sich daher nur durch Maßnahmen von außen, so z.B. durch gesetzgeberische Initiativen, dauerhaft stärken. Möglicherweise könne den Rechnungshöfen hier eine besondere Rolle zufallen, aber auch diese müsse in eindeutige Rechtsformen gegossen werden, um Überlastungen zu vermeiden. Unterstützung könne schließlich auch in der Universitätslandschaft mobilisiert werden, wenn gleich hier trotz einzelner renommierter Standorte in Deutschland noch ein vergleichsweise geringer Stand der Institutionalisierung von Verwaltungswissenschaft zu konstatieren sei.

An den von Herr Bogumil angesprochenen Qualitätswettbewerb von Kommunen knüpfte Herr *Hill* (Deutsche Hochschule für Verwaltungswissenschaften Speyer) in seinem Statement an. Hierbei schilderte er den Ablauf des von Herrn Klages und ihm initiierten Speyerer Qualitätswettbewerbs. An erster Stelle stehe hierbei – im Vorfeld der offiziellen Bewerbung – eine Selbstbewertung der Verwaltungen nach einem vorgegebenen Bewertungsschema. In einem zweiten Schritt erfolge eine Bewertung durch ein Fachgutachtergremium, zunächst anhand der schriftlichen Unterlagen, dann in einer „Begutachtung vor Ort“. Abschließend erfolge die Prämierung erfolgreicher Kommunen in einem Wettbewerb. Um den Wirkungshorizont über die Wettbewerbslaufzeit hinaus zu erweitern, habe man Wege der Institutionalisierung entwickelt, und zwar in der Form von Innovationsringen, die weniger dem klassischen benchmarking, als vielmehr der gegenseitigen Unterstützung in Lernprozessen, sozusagen einem „benchlearning“ dienten. In jüngerer Zeit widme man sich außerdem der Überführung der bisherigen nationalen Netzwerkaktivitäten in internationale, vor allem europäische Initiativen (so z.B. das CAT „Common Assessment Trainwork“).

In seinem Statement unterstrich Herr *Wollmann* (Humboldt-Universität Berlin), daß die externe Evaluierung von Verwaltungspolitik in der Bundesrepublik im internationalen Vergleich auffallend schwach entwickelt sei. Er verwies beispielhaft auf sektorale ebenso wie umfassendere Evaluierungen in Australien seit den 80er Jahren, ungeachtet der dortigen, ebenfalls föderalen Strukturen. Im Unterschied zu

Deutschland sei hier von der zentralen Ebene die Initiative ergriffen worden, ein umfangreiches Evaluierungsverfahren in Gang zu setzen. Ein anderes Beispiel lieferten die USA als ebenfalls föderal verfasstes Land, mit bemerkenswert breiten Evaluierungsaktivitäten. Auch in Schweden seien externe Evaluierungsverfahren in beachtlichem Umfang institutionalisiert.

Allerdings habe die Verwaltungsreformpolitik in den 1990er Jahren zweifellos eine neue Dynamik erfahren, insbesondere im Hinblick auf die Berücksichtigung von Verfahren der internen Evaluierung, wie Formen des Controlling und des Berichtswesens, in der Verwaltungspolitik. Trotz positiver Ansätze im Hinblick auf die Schaffung interner Transparenz in deutschen Kommunen könne allerdings von einem Controlling im fortgeschrittenen Sinne nach wie vor nicht gesprochen werden. Aufgrund des Fehlens externer Evaluierungen und damit verbunden des Einsatzes geeigneter Evaluierungsinstrumente seien schließlich auch die 1994 versuchten und ihrer Idee nach zu befürwortenden Experimentierklauseln für Kommunen auf einen engen Wirkungshorizont beschränkt geblieben.

Hervorzuheben bleibe hierbei, in Übereinstimmung mit den Vorrednern, der geringe Umfang externer Evaluierung. Auch Herr Wollmann konstatiert, dass dies ursächlich eher auf das Desinteresse der politischer Verwaltung als auf budgetäre Engpässe zurückzuführen sei. Ernsthafte Anstöße seitens der Bereiche angewandter Politik und Verwaltungsforschung seien auf die institutionelle Entwicklung der Verwaltungsreform, also eher auf die Implementationsforschung, begrenzt geblieben. Eine unzureichende Berücksichtigung hätten dabei evaluative Komponenten, d.h. Fragen nach den weiterreichenden und „letzlichen“ Wirkungen des Verwaltungshandelns, gefunden. Die genannten Defizite seien um so erstaunlicher, rufe man sich die Vorreiterrolle der deutschen Evaluationsforschung in den 60er und 70er Jahren in Erinnerung.

Als mögliche Ursachen der aktuellen Situation stellte Herr Wollmann folgende Aspekte zur Diskussion:

- die bisher verhältnismäßig kurze Laufzeit vieler aktueller Reformprojekte. So träten gerade in den nächsten Jahren in mehreren Bundesländern die seit den 80er Jahren in Gang gekommenen Verwaltungsreformen erst in eine evaluierbare Phase, so dass hier möglicherweise Raum für Impulse zu verstärkter externer Evaluierung gegeben sei;
- die hochgradige Fragmentierung der Verwaltungslandschaft, in der Bund, Länder und Kommunen weitgehend eigenständig agieren. Dies gelte auch im Hinblick auf die Initiierung von Evaluationen, bei der über die Verwaltungsebenen hinweg klare Grenzen existierten, die der Verfassungslogik entsprächen;
- das bislang unzureichende Interesse der Parlamente an der Verwaltungspolitik;
- die bislang unzureichend definierte Rolle der Rechnungshöfe im Hinblick auf eine Evaluierung von Verwaltungspolitik;
- die spezifischen methodischen und konzeptionellen Probleme, die mit der Durchführung externer Evaluationen verbunden sind und durch die vielfach eher zurückhaltende Einstellung der politischen und administrativen Akteure verstärkt würden. Wenig entwickelt seien außerdem Formen der Evaluation, die im Rahmen einer konstruktiven Interaktion, eines Dialogs mit der Verwaltung Veränderungen zu erfassen suchen. Der französische Politikwissenschaftler und

Evaluierungsexperte Jean-Claude Thoenig hat unlängst den in diesem Zusammenhang interessanten Vorschlag einer methodisch „verschlankten“ Variante von Evaluierung (er spricht von „Quasi-Evaluierung“) gemacht, die sich – gewissermaßen als „Evaluierung light“ – mit der eher oberflächlichen Erfassung von Veränderungen und Wirkungen begnügt und geeignet ist, als sozusagen „vertrauensbildende“ Maßnahme das auf Seite der politischen Verwaltung verbreitete Misstrauen abzubauen, abzuschwächen und in einen wechselseitigen Lernprozess einzutreten, an dessen Ende durchaus unabhängige und wissenschaftlich solide Evaluierungen stehen könnten.

Diskussion

Die Stellungnahmen der Podiumsteilnehmer zusammenfassend stellt Herr *Fisch* (Deutsche Hochschule für Verwaltungswissenschaften Speyer) fest, dass Evaluation in der Verwaltung die unterschiedlichsten Facetten haben könne, von der Qualitätssicherung bis hin zur Kontrolle. Grundsätzlich herrsche Konsens darüber, dass es aus einer Steuerungsperspektive unbedingt zur Aufgabe einer Verwaltungsführung gehöre, zu überprüfen, ob die initiierten Maßnahmen auch die erwünschten Erfolge brächten. Von der Umsetzung dieses Anspruchs sei die Verwaltungskultur aber noch weit entfernt. Dem stehe in der Tat häufig die Politisierung von Evaluationsergebnissen entgegen.

Beispielhaft für die Problematik der Umsetzung von Evaluationsergebnissen führt er eine Schweizer Initiative an, wo eigens eine parlamentarische Stelle zur Evaluierung politischer Programme bzw. auch im voraus zur Folgenabschätzung eingerichtet worden sei. Die Rezeption der hier erzeugten Daten habe sich aber als gering erwiesen. Ein fundamentales Interesse an diesen Informationen fehle. Grundsätzlich handele es sich also bei der Evaluation politischer Programme um ein schwieriges Gebiet, in dem noch viel Vertrauensarbeit mit Vertretern der öffentlichen Verwaltung zu leisten sei, um den Zugang externer Gutachter und die Rezeption und Umsetzung von Evaluationsergebnissen zu ermöglichen.

Herr *Seyfried* (Innenministerium Baden-Württemberg) hält dem entgegen, dass er Evaluation gerade *nicht* für ein politisches Instrument halte. Seine Diagnose bestehe im Gegenteil darin, dass sich Evaluation in der Verwaltung *unterhalb* der politischen Wahrnehmungsschwelle bewege. Überdies fehle es weitgehend an öffentlichen Diskussionen darüber, wie gut oder wie schlecht eine Verwaltung sei. Verwaltungspolitik habe nahezu keinen öffentlichen Aufmerksamkeitswert. Die Behandlung des Themas in den Medien beschränke sich gewöhnlich auf Bürokratiekritik oder Beamtschelte. Differenzierte sozialwissenschaftliche oder betriebswirtschaftliche Analysen aber würden selten aufgegriffen.

Zu einer differenzierteren Betrachtung des Gegenstand von Evaluationen riet Herr *Dieckmann* (Hamburger Rechnungshof). Seiner Ansicht nach laufe man fehl, wenn man von der *Evaluation der Verwaltungsreform* rede, sei es in Deutschland allgemein oder auch auf Landes- oder kommunaler Ebene. Auch hinsichtlich der Frage nach einem funktionierenden Controlling könne man die Diagnose nicht auf *die Verwaltung* im allgemeinen richten. Vielmehr müsse man den Untersuchungs-

gegenstand aufgabenspezifisch differenzieren. Wohl könne man fragen, ob etwa die Stadtreinigung, die Jugendarbeit oder die Altlastensanierung über ein funktionsfähiges Controlling verfüge. Jeder dieser Bereiche sei aber derart unterschiedlich, dass eine integrierte Betrachtung sich nur schwer realisieren lasse. Verwaltungsmodernisierung könne also nur als aufgabenspezifische Verwaltungsmodernisierung begriffen werden, wobei natürlich bereichsübergreifende gegenseitige Lernprozesse nicht auszuschließen seien und gefördert werden müssten. Unabhängig von der Untersuchungsebene, Verwaltung im allgemeinen oder aufgabenspezifische Tätigkeitsfelder, stelle sich bei der Evaluierung von Reformwirkungen aber immer ein grundsätzliches Problem, nämlich die lange Laufzeit von Reformen und die von Sektor zu Sektor völlig unterschiedlichen Rahmenbedingungen, die eine vergleichend bewertende Erfassung von Reformwirkungen praktisch verhindere.

In einem weiteren Diskussionsbeitrag mahnte Herr *Stockmann* (Universität des Saarlandes) eine Klärung der Begrifflichkeit und Terminologie an. Es würden die Begriffe Kontrolle und Evaluierung, die nur wenig miteinander zu tun hätten, vielfach vermengt. Der Umstand, dass sich begriffliche Unklarheit und Vermischung auch auf die Praxis übertrage, erkläre zum Teil die mit Politikevaluation verbundenen Schwierigkeiten. Immer dann, wenn Evaluation hauptsächlich als Kontrolle verstanden würde, könne man zwangsläufig nur mit einer geringen Akzeptanz rechnen. Verstünde man Evaluation aber als ein Instrument zur Gestaltung von Prozessen, als Element eines Qualitätsmanagements, und gewährleiste man eine kontinuierliche Partizipation der Evaluierten, dann erhöhten sich auch die Chancen, Akzeptanz für die Durchführung von Evaluationen zu erzeugen. Vor allem die Chance zur Umsetzung von Evaluationsergebnissen hänge entscheidend von dieser Akzeptanz ab. Aus eben diesem Grunde seien die in der Diskussion als möglicherweise beratende Instanz angeführten Rechnungshöfe zur Durchführung von Evaluationen kaum geeignet.

In Anlehnung an den Bericht des Bundesrechnungshofes von 1998 liege ein weiterer Grund für die Evaluationsproblematik auf Verwaltungsebene in dem nur gering ausgeprägten Verständnis der Verwaltungsbeamten dafür, was Evaluation eigentlich sei. Zwar könnten nicht unbedingt eigene Durchführungskompetenzen erwartet werden, aber auch bei der externen Vergabe von Evaluationsaufträgen, fehle ein Kenntnis dessen, was eigentlich von den Auftragnehmern erwartet werden dürfe. An dieser Stelle sei eine verstärkte Zusammenarbeit mit Aus- und Weiterbildungsinstitutionen oder anderen externen Einrichtungen anzuraten, um die Kompetenzdefizite schrittweise zu beseitigen.

Abschließend äußert sich Herr *Stockmann* positiv zu dem von Herrn *Bogumil* aufgeworfenen Gedanken einer Stiftung *Warentest* für externe Wirkungskontrollen. Es handele sich um einen Vorschlag, der bereits vor mehreren Jahren sektorspezifisch, in der Entwicklungszusammenarbeit, diskutiert worden sei. Hier habe ein Abgeordneter des Ausschusses für wirtschaftliche Zusammenarbeit die Gründung eines unabhängigen Instituts für Wirkungsuntersuchungen in der Entwicklungszusammenarbeit vorgeschlagen. Der Diskussionsansatz sei selbstverständlich auch auf andere Politikbereiche übertragbar.

Mit Blick auf die verwaltungsinternen Widerstände gegenüber Evaluation machte Herr *Wittmann* (Universität Mannheim) auf eine Situation aufmerksam, die seiner Ansicht nach ein Paradoxon darstelle. Es sei allgemein akzeptiert, dass Poli-

tiker in einem gewissen Sinne regelmäßig evaluiert würden, und zwar im Sinne der Kundenzufriedenheit, sprich Wählerzufriedenheit. Es sei also eine interessante Frage, warum Politiker, die selbst beständig einem Evaluationsprozess ausgesetzt seien, keine verstärkte Initiative dazu ergriffen, dass auch die Verwaltung systematischer evaluiert werde.

Als weiteren Punkt beleuchtete Wittmann die Möglichkeit der Institutionalisierung von Evaluationskompetenzen im Tätigkeitsbereich der Rechnungshöfe. So sei im amerikanischen Pendant der General-Accounting-Office in den USA zeitweise eine hervorragende Abteilung aufgebaut worden, die Programme Evaluation Division, von der jeder Abgeordnete des Kongresses die bestmöglichen verfügbaren Daten zur Evaluation politischer Programme einfordern konnte. Es sei zu fragen, inwieweit der Evaluationsgedanke auch in deutsche Rechnungshöfe auf eine ähnliche Weise hereingetragen werden könne.

Auf diese letzte Frage entgegnete Herr *Dieckmann*, er halte es durchaus für möglich, eine Evaluationskultur auch aktiv zu erzeugen. Die stärkere Position der Evaluationsforschung in den USA hänge nicht zuletzt damit zusammen, dass Evaluation dort seit den 60er Jahren durch gesetzgeberische Maßnahmen in vielen Formen institutionalisiert worden sei. Auch die Einrichtung verantwortlicher Abteilungen innerhalb oder außerhalb der Verwaltung könne ein wirksames Instrument der Institutionalisierung darstellen. Der Rechnungshof wende sich hierbei – hierin sei Herrn Stockmann zuzustimmen – gegen die Nutzung von Evaluation als Kontrolle. Solange man aber eine machtpolitische Instrumentalisierung nicht unterbinde, werde sie immer auch als Kontrolle empfunden. Möglicherweise könne sich hier eine stärkere Institutionalisierung von Evaluation an unabhängiger Stelle tatsächlich als Ausweg erweisen.

Auch Herr *Hill* befürwortete die Einrichtung einer unabhängigen Evaluationsinstanz, deren Dienste und Informationen den Parlamenten und Parlamentariern zur Verfügung stünden. Dies könne aber nicht von den Rechnungshöfen verlangt werden, da dies weder ihrem Auftrag, den geltenden verwaltungsrechtlichen Regelungen noch ihren budgetären Möglichkeiten entspräche. Die Einrichtung einer eigenen, unabhängigen Instanz sei dagegen weiterzuverfolgen. Grundsätzlich sei aus Sicht der Verwaltung aber entscheidend, dass nicht die Kontrollfunktion, sondern die Beratungsfunktion im Vordergrund stehen müsse. Für die Verwaltung sei nicht die Fehlersuche interessant, sondern die kontinuierliche Begleitung und Unterstützung von Lernprozessen.

Herr *Dieckmann* ging noch einmal auf eine Präzisierung der Rolle der Rechnungshöfe ein, die sie aus seiner Sicht hinsichtlich der Institutionalisierung von Evaluation spielen können. Seiner Ansicht nach geht die Diskussion um Evaluation auf der einen und Kontrolle auf der anderen Seite fehl. Nach seinem Verständnis habe der Rechnungshof, auch und gerade als Kontrollinstanz, vor allem beratende Aufgaben, und dies würde auch von den Verwaltungen so empfunden. Aus diesem Grunde ergäben sich nur in den seltensten Fällen Widerstände gegen evaluierende Schritte des Rechnungshofes.

Trotz seines beratenden Evaluationsverständnisses könne der Rechnungshof aber keinesfalls die Rolle einer unabhängigen Instanz im Dienste der Parlamente ausfüllen, wie sie zuvor diskutiert worden sei. Die Unabhängigkeit der Rechnungshöfe bestehe, ihrem Auftrag entsprechend, gerade darin, dass sie sich selbst ihre

Aufgaben stellen und nicht in Funktion der durch das Parlament oder die Regierung gestellten Ansprüche tätig würden. Eine Instrumentalisierung der Rechnungshöfe über das Element Gutachten müsse in jedem Falle vermieden werden.

Herr *Stockmann* unterstützte einen weit gefassten Evaluationsbegriff, der sich auf unterschiedliche Instrumente und Verfahren beziehen könne, betonte aber noch einmal die besondere Wichtigkeit *externer*, unabhängiger Evaluationen. Die externe Durchführung stehe dabei keineswegs im Widerspruch zu einem formativen Charakter. Im Gegenteil habe man mit partizipativen Vorgehensweisen, eingesetzt von externen Gutachtern, die besten Erfahrungen bei der Evaluation von Reformprozessen wie auch Reformwirkungen. Um derartige Evaluationsprozesse in Gang zu setzen, sei der Aufbau eines Vertrauensverhältnisses zwischen Gutachter und Verwalter von zentraler Bedeutung. Aus diesem Grunde seien auch die Rechnungshöfe nicht in der Lage, die Rolle einer unabhängigen Instanz zu übernehmen, wie sie mehrfach in der Diskussion vorgeschlagen wurde. Die Funktionen externer Kontrolle und beratende Prozessbegleitung seien eben doch nicht widerspruchsfrei miteinander zu vereinbaren. Im Gegensatz dazu hätten Gutachter aus dem universitären Umfeld deutlich bessere Möglichkeiten, ihre unabhängige Position glaubhaft darzustellen.

Herr *König* (Deutsche Hochschule für Verwaltungswissenschaften Speyer), wandte sich gegen die Einschätzung, dass Deutschland im internationalen Vergleich im Hinblick auf die Evaluation von Verwaltungspolitik rückständig sei. Die Situation unterschiede sich in anderen Ländern nur geringfügig. Selbst in dem als positives Beispiel genannten Australien gebe es zwar eine umfangreichere Evaluationstätigkeit, die aber, wie dies Herr *Banner* auch für Deutschland dargestellt habe, über geschönte Selbsttestate kaum hinauskomme. Kritische Ansätze kämen, wenn überhaupt, grundsätzlich von Instanzen außerhalb der Verwaltung. Aus verständlichen Gründen aber stelle sich die Verwaltung nach außen hin nur ungern zur Diskussion. Die von wenigen Gutachtern und Evaluationsforschern beherrschte Kunst bestehe darin, mit den Mitgliedern der Verwaltung konstruktiv auch über Misserfolge sprechen zu können. Hier sei die hinsichtlich der Evaluation als besonders fortschrittlich geltende Entwicklungszusammenarbeit in einer Sondersituation, da man sich Misserfolge nicht zwangsläufig selbst zuschreiben müsse, sondern Verantwortung auf vermeintliche long-distance-Verantwortliche abwälzen könne. Obwohl mittlerweile auch andere Sektoren, wie etwa die Verkehrspolitik oder die Verteidigungspolitik, Fortschritte gemacht hätten, sei die Situation dort grundsätzlich schwieriger. In jedem Fall bilde die Verwaltungspolitik im Hinblick auf den Stand der Evaluationstradition das Schlusslicht. Dies aber sei kein spezifisch deutsches Problem, auch kein spezifisches Problem eines föderalen Staates, sondern ein für die meisten Länder geltendes Phänomen, das in erster Linie mit dem für Politiker wenig attraktiven Charakter der Verwaltungspolitik zu tun habe.

Den spezifischen Charakter der Verwaltungspolitik betonte auch Herr *Wollmann*. Es handele sich hier gewissermaßen um ein „Reservat der Exekutive“, auf das die Parlamente keinen oder nur geringen Einfluss nähmen. Demgegenüber sei daran zu erinnern, dass „substantielle“ Politiken, so etwa Sozialpolitik, Umweltpolitik, Verkehrspolitik, ihre jeweilige gesellschaftliche Klientel haben. Das Vorhandensein von Interessengruppen erhöhe auch das Interesse seitens der Politiker. Eine solche gesellschaftliche Klientel fehle der Verwaltungspolitik jedoch, weshalb sich

für Politiker ein Engagement in diesem Politikfeld politisch, vor allem mit Blick auf die Gewinnung von Wählerstimmen, kaum „lohne“.

Die Plenumsveranstaltung endete mit einem Schlusswort des Vorsitzenden der Deutschen Gesellschaft für Evaluation (DeGEval), Prof. *Daniel*, der damit auch die Jahrestagung 2001 abschloss.