

Evaluation von NPM-Pilotprojekten¹

WERNER BUSSMANN

New Public Management (NPM) ist ein von Wettbewerbs- und privaten Managementprinzipien inspiriertes Modell der Führung öffentlicher Verwaltungen. Die helvetische NPM-Variante ist gekennzeichnet durch die Zuteilung umfassender operativer Verantwortung an Verwaltungseinheiten (sog. "agencies", Steuerung über Leistungsaufträge) und die Ersetzung der bisherigen inputorientierten Ressourcenzuteilung durch sog. Globalbudgets (vgl. Beitrag Haldemann in diesem Heft). In seiner Begrifflichkeit und in seinem Anspruch ("wirkungsorientierte Verwaltungsführung") ist das NPM darauf ausgerichtet, Wirkungsziele festzulegen und diese regelmässig zu überprüfen. Diesem Anspruch hat NPM "in eigener Sache" freilich noch nicht nachgelebt: Die Evaluation der Wirksamkeit des NPM steht noch in den Anfängen. In Form von 8 Thesen werden die Herausforderungen, möglichen Fallgruben und Lösungswege bei der Evaluation von NPM-Projekten skizziert.

These 1: Evaluation von Institutionen gehört zum Schwierigsten

Im Laufe der vergangenen Jahre hat das NPM zumindest in der Rhetorik einen eindeutigen Siegeszug angetreten. Es hat sich nach den Worten seiner wichtigsten deutschsprachigen Protagonisten weltweit als überlegenes Modell der Verwaltungsführung herausgestellt. In mehreren Kantonen (BE, LU, BS, ZH, SH, SO, FR, GR) sind inzwischen NPM-Pilotprojekte oder ihnen ähnliche Projekte eingeführt worden.

¹ Beim vorliegenden Beitrag handelt es sich um den überarbeiteten Text eines Referats am IDHEAP- Seminar "Verwaltungskontrolle in der Demokratie", Institut Kurt-Bösch, Sitten vom 26. Juni 1997

Der rhetorische Siegeszug des NPM in der Schweiz erweckt gewisse Bedenken angesichts der Tatsache, dass die Wirkungen des NPM bisher nirgends seriös evaluiert worden sind, wie Christopher Pollitt (1995) vor einiger Zeit dargelegt hat. In der Zwischenzeit sind ebenfalls keine umfassenden Evaluationen vorgenommen worden. Dies ist nicht erstaunlich, denn die Evaluation von NPM-Projekten gehört zum Schwierigsten der Evaluation überhaupt.

Die Probleme der Evaluation von NPM hängen mit dessen sehr weitgehendem Anspruch zusammen. NPM möchte nicht einfach ein neues technizistisches Modell der Verwaltungsführung sein. Es möchte vielmehr eine tiefgreifende Veränderung der Verwaltungsführung auslösen. Prozesse sollen im Sinne der Wirkungssteuerung geändert, die Motivation in der Verwaltung soll im Sinne einer Vertrauens- und Vereinbarungskultur verbessert und auch die Verwaltungsstrukturen sollen sich durch flachere Hierarchien auszeichnen (vgl. Beitrag Haldemann in diesem Heft). Beizufügen ist, dass all diese Veränderungen natürlich nicht Selbstzweck sind, sondern letztlich den Kunden der Verwaltung, ja generell den Bürgerinnen und Bürgern dienen sollen.

Ferner ist darauf hinzuweisen, dass NPM-Projekte eine grosse Zahl von Betroffenen, im Evaluationsjargon "stakeholders", aufweisen. Es sind dies namentlich die NPM-Projektleitung, die Amtsleiterinnen und Amtsleiter sowie die Mitarbeiterinnen und Mitarbeiter der Pilotdienststellen, die Querschnittsämter (Finanzen, Personal, Liegenschaften), die Departementsleitung, die Gesamtregierung, Parlamentarierinnen und Parlamentarier, Bürgerinnen und Bürger, Interessengruppen, politische Parteien u.a.m.

These 2: Bei der Evaluation sind NPM-Projekte, Personen und substantielle Politiken einzubeziehen

Die Vielzahl der mit der Evaluation verbundenen Ansprüche sowie die zahlreichen Stakeholders bringen es mit sich, dass die Evaluation von NPM-Projekten letztlich drei Aspekte aufweist, die in der Evaluationsforschung unter unterschiedlichen Titeln thematisiert werden:

- Es geht erstens um die Evaluation der mit NPM eingeführten neuen Prozesse. Hierbei ist zu untersuchen, ob und inwieweit sich die Verwaltungsabläufe verändern, ob ein Kulturwandel in der Verwaltung stattfindet, ob sich das Arbeitsklima verbessert oder zumindest nicht verschlechtert und ob keine unerwarteten Nebenwirkungen (z.B. eine Zunahme der Korruption als Folge des Verzichts auf Inputsteuerung) eintreten.
- Bei dieser Evaluation kann zweitens die personelle Dimension nicht völlig ausgeklammert werden. Inwieweit sich die Reform bewährt, hängt stark davon ab, wie sich die Amts- und Departementsleitung dazu einstellt, ob das Verhältnis zur NPM-Projektleitung durch Vertrauen oder Misstrauen geprägt ist und ob beispielsweise das Präsidium wichtiger parlamentarischer Kommissionen (v.a. Geschäftsprüfungs- und Finanzkommissionen) den Neuerungen positiv oder skeptisch gegenübersteht. Evaluation von NPM-Reformen ist häufig auch eine Evaluation von Personen. Diese werden versuchen, die Evaluation zu ihren Gunsten einzuspannen, worauf wir zurückkommen werden.
- Als drittes ist zu untersuchen, ob sich durch NPM messbare Vorteile für die Bürgerinnen und Bürger (z.B. infolge kostengünstigerer Leistungserbringung und damit indirekt geringerer Steuern) und für die Kundinnen und Kunden (z.B. in Form qualitativ besserer Dienstleistungen) ergeben. Denn dies sollte ja das wichtigste Ziel der Reform sein: bessere und kostengünstigere Leistungen zu erbringen.

Die Evaluationsforschung als Wissenschaftszweig, ja als Transdisziplin (Scriven 1991, S. 1-43 und S. 363) hat sich bisher mit substantieller Politik (seit den sechziger Jahren v.a. mit amerikanischen Sozial- und Bildungsprogrammen, später auch in anderen Ländern und mit anderen Politikbereichen) befasst. In der Schweiz ging und geht es beispielsweise um konkrete Entwicklungshilfeprojekte, um Drogenpräventionsprogramme oder um den Erfolg der Wohnbauförderungspolitik. Evaluationen sog. institutioneller Politiken, welche die Wirkungsweise von Institutionen (z.B. Parlamenten, demokratischen Rechten, Mechanismen des Verwaltungshandelns) verändern, sind aber bisher wenig durchgeführt worden (vgl. Bussmann, Klöti und Knoepfel 1997: 59-62). In der Schweiz gab es allenfalls Evaluationen von Schulversuchen, Modellversuchen im Strafvollzug und von einzelnen Institutionen (BUWAL:

KNOEPFEL und ZIMMERMANN 1992. Amt für Umweltschutz SG: KNOEPFEL, BAITSCH und EBERLE 1995).

Institutionelle Politiken dürften indessen in nächster Zeit sehr viel bedeutsamer werden. Im Bereich staatlicher Politik ist ein eigentlicher Paradigmenwechsel auszumachen. Während der vergangenen drei, vier Jahrzehnte hat nämlich der Staat neue Instrumente und Kooperationsformen angewandt, um die Gesellschaft zu verändern. Campbell (1972) hat dieses Paradigma unter dem Titel "The experimenting society" (es war in der Tat aber der experimentierende Staat) exemplarisch beschrieben. Nun, scheint es, wollen Wirtschaft und Gesellschaft neue Modelle der Verwaltungsführung ausprobieren: Das Modell des Experimentierens wird somit auf den Staat selbst angewandt. Aus diesem Grund wird, so die Vermutung, die Evaluation institutioneller Politik in Zukunft stark an Gewicht gewinnen.

Man kann sich nun fragen, ob die Evaluation ihrem neuen Gegenstand überhaupt gewachsen ist. Ist die Evaluation institutioneller Politik nicht sehr viel schwieriger als diejenige substantieller Politik? Wie erwähnt sind Prozesse, Personen und Politikergebnisse gleichzeitig einzubeziehen. Das macht solche Evaluationen vermutlich schwieriger als sog. Black-box-Evaluationen substantieller Politik, also Evaluationen, welche nur messen, ob bestimmte Massnahmen gewisse Resultate verursacht haben, ohne nach den genauen Gründen für den Erfolg oder Misserfolg zu fragen. Soweit substantielle Evaluationen auch die Wirkungsmechanismen (unter Einschluss institutioneller Vorkehrungen) einbeziehen, stellen sich ihnen indessen ähnliche Schwierigkeiten und Herausforderungen wie der Evaluation institutioneller Politik. Holkeri und Summa, welche sich mit der Evaluation finnischer NPM-Projekte befassen, weisen indessen noch auf ein weiteres Problem hin, die oft diffusen Ziele von NPM-Vorhaben und die Vielzahl parallel laufender Reformen. Beides kann zu Problemen der Wirkungserfassung führen. Verschiedene NPM-Projekte, beispielsweise in den Kantonen Solothurn, Bern und Genf, sind mit noch anderen Zielen, insbesondere finanzieller Art, verbunden. Diese prägen das Umfeld der NPM-Versuche sehr stark. Die beiden Autoren (HOLKERI und SUMMA 1996, S. 23) kommen deshalb zum Schluss: "...it seems evaluations of public management reform programmes present specific problems, which may not be as prominent in policy evaluation in general. Public management reform programmes tend to be series of

policy statements, decisions and projects that accumulate over a long period of time rather than coherent programmes with a simple set of objectives".

These 3: Die Vergleichsbasis ist entscheidend für das Ergebnis einer Evaluation von NPM-Reformen

Im Zentrum der Evaluation als Wirkungsforschung steht die Isolierung von Projekten, Massnahmen oder Politiken. Hintergrundfolie bildet die Frage, was eigentlich wäre, wenn die betreffende Massnahme oder Politik nicht in Kraft wäre. Die klassische Evaluationsforschung interessiert sich somit primär für konterfaktische Fragen². Etwas salopp ausgedrückt geht es um den Vergleich "policy on - policy off."

Ein solcher Vergleich wird am ehesten beim sog. experimentellen Design ermöglicht. Die Massnahmen oder das Projekt wird bei einer zufällig ausgewählten Gruppe getestet; eine andere, ebenfalls zufällig ausgewählte Kontrollgruppe dient dem Vergleich. Ebenfalls zur Ermittlung der Wirkungen eingesetzt werden kann ausserdem der Längsschnitt-Vergleich. Er ist vor allem nützlich in Fällen, bei denen eine Massnahme im Zeitablauf variiert wird, beispielsweise beim Start eines Projekts, beim Inkrafttreten eines neuen Erlasses oder bei deren Beendigung bzw. Aufhebung (namentlich im Rahmen von Deregulierungsvorhaben).

An sich wären die Voraussetzungen für die Evaluation von NPM-Projekten im vorerwähnten klassischen Sinn durchaus gegeben. NPM-Projekte werden nicht flächendeckend eingesetzt, sondern sind vielmehr beschränkt auf sog. Pilotämter. Sie beginnen ausserdem zu einem klar bestimmten Zeitpunkt. Die Möglichkeiten sowohl des Quervergleichs wie des Längsschnittvergleichs sind somit intakt. Leider gibt es aber ein paar unangenehme Probleme.

² Solche Fragen spielen neuerdings auch in der Geschichtsforschung eine wichtige Rolle. Um den Beitrag der Handelsbeziehungen mit Nazi-Deutschland zur territorialen Unversehrtheit der Schweiz zu ermitteln, muss die Hypothese eines Abbruchs oder einer Reduktion dieser Handelsbeziehungen unter Einschluss der möglichen Reaktionen Deutschlands gedanklich durchgespielt werden.

- Erstens werden die NPM-Pilotprojekte nicht nach dem Zufall auf die Amtsstellen verteilt. Das Motto heisst vielmehr: "Freiwillige vor!" Es spielen somit Prozesse der Selbstauswahl. Für solche gibt es in der Evaluationsforschung den schönen Begriff "creaming" ("abrahmen"). Selbstauswahl bevorzugt häufig initiative Personen oder Amtsstellen. Dadurch kann der Quervergleich verfälscht werden.
- Zweitens wäre zu untersuchen, von welcher Ausgangsbasis die betreffenden Dienststellen starten. Handelt es sich um Dienststellen, die bereits vorgängig sehr viele Reformen unternommen haben, dann wird im Rahmen des Quervergleichs ein Teil des Erfolgs zu Unrecht dem NPM zugerechnet. Handelt es sich umgekehrt um Dienststellen mit einem grossen Reformbedarf, dann kann das Gegenteil eintreten. Im letzteren Fall würde die NPM-Reform im Vorher-nachher-Vergleich hingegen möglicherweise zu gut abschneiden (Start von sehr tiefem Ausgangspunkt).

Hinter der Wahl der Vergleichsebene und der Vergleichsobjekte steckt demzufolge viel evaluativer Sprengstoff. Es fragt sich letztlich, was dem NPM-Pilotprojekt gegenübergestellt werden soll: Eine Verwaltung im traditionellen, ja gar negativen Sinne oder eine Verwaltung, welche neue Managementmodelle und Kooperationsformen (z.B. Vereinbarungen, kundenorientierte Organisation) bereits einsetzt. Es besteht häufig eine Tendenz, dem Modell der NPM-Verwaltung eine traditionelle Bürokratie, wie sie Max Weber beschrieben hat, gegenüberzustellen. Viele Dienststellen haben aber bereits Elemente eingebaut, welche den Elementen der Ökonomik, der Responsivität, der Transparenz sowie der Notwendigkeit einer sinnvollen organisatorischen Ausdifferenzierung Rechnung tragen (vgl. Bussmann 1996/97).

Zusammenfassend: Ein Quervergleich von NPM-Projekten mit herkömmlicher Verwaltung wird darauf achten müssen, dass nicht "archaische", sondern zeitgemässe Formen der Verwaltung als Vergleichsbasis genommen werden. Da NPM-Projekte Energien bündeln und ihre Protagonisten beflügeln ("Hawthorne-Effekt"), ist darauf zu achten, dass auch Verwaltungsmodelle, die auf der Inputsteuerung und einer engen Verbindung von Regierungsspitze und operationeller Verwaltung beruhen, mit gewissem Elan verteidigt werden (die Evaluationsforschung be-

zeichnet dies als "John Henry-Effekt"). Bei einem Längsschnitt-Vergleich ist darauf zu achten, von welchem Ausgangspunkt der Versuch gestartet wurde. Bei der Planung von NPM-Projekten ist zudem mehr als bisher der Erfassung des Ist-Zustands vor der Reform Beachtung zu schenken. Evaluationsanliegen sind deshalb bereits bei der Planung von NPM-Reformen ausreichend einzubeziehen³, was bisher noch nicht geschehen ist.

These 4: Evaluation von NPM-Projekten kann von schönfärberischen Alibiübungen bis zu ernsthaften Wirkungsanalysen gehen

Evaluation hat heute viele Facetten. Die vorgehend dargestellte Evaluationsform hat einen bilanzierenden Charakter (summative Evaluation; Bilanzevaluation) und erstreckt sich auch auf die Wirkungen ausserhalb der Verwaltung. An sich können nur solche Evaluationen den Anspruch des NPM, eine in allen Aspekten überlegene Form moderner Verwaltungsführung zu sein, ausreichend objektiv auf ihren Wirklichkeitsgehalt überprüfen.

Das heisst nicht, dass bei NPM-Projekten nicht auch noch andere Evaluationsbedürfnisse bestehen. Von Seiten der NPM-Verantwortlichen sowie der beteiligten Amtsleiterinnen und Amtsleiter besteht namentlich ein Bedarf nach kontinuierlicher Begleitung und Beratung. Auch dafür wird häufig der Begriff der Evaluation verwendet (formative evaluation/development evaluation/Verbesserungsevaluation, vgl. Bussmann, Klöti und Knoepfel 1997: 264). Ziel solcher Evaluationen ist es, Probleme aufzudecken und Verbesserungspotentiale aufzuzeigen. Es geht nicht um einen Vergleich "policy on - policy off", sondern um das Finden von Lösungen, die noch besser sind ("policy on - improved policy"). Obwohl eine solche Form von Evaluation intern, z.B. in Form von Erfahrungsaustausch-Gruppen oder Qualitätszirkeln vorgenommen werden kann, lohnt sich häufig der Beizug externer Berater. Diese haben eine unvoreingenommene Sicht der Dinge, kennen möglicherweise vergleich-

³ So schreiben HOLKERI und SUMMA 1996, S. 22: "...a prerequisite for good evaluation is that the need to evaluate is taken into account in planning the reforms."

bare Situationen aufgrund von früheren Forschungs- oder Beratungsaufträgen und sind zudem interessenmässig weniger belastet als amtsinterne Stellen. Ihre Argumente haben möglicherweise bei allen Beteiligten eine grössere Autorität.

Zwischen den beiden extremen Positionen, die aufgezeigt wurden, gibt es allerlei Zwischenschattierungen, beispielsweise Evaluationen bilanzierenden Charakters, welche sich aber nur auf die verwaltungsinternen Prozesse beziehen und welche die gesellschaftlichen Wirkungen des NPM ausblenden. Es gibt auch Evaluationen, welche nur die Hauptkunden der Verwaltung einbeziehen, aber die Auswirkungen auf die "unwichtigeren" Kunden (das sind meistens Gruppen, die im politischen Konzert weniger laute Stimmen haben), vernachlässigen.

In der Anfangsphase des NPM, in der wir heute stehen, stehen zudem noch andere Evaluationstypen im Vordergrund. Es geht nämlich vorerst einmal darum abzuklären, ob sich die Vorstellungen der NPM-Theoretiker überhaupt in die Praxis umsetzen lassen. Es ist beispielsweise abzuklären, ob die Amtsstellen wirklich über vermehrte Selbständigkeit in operativen Belangen verfügen und ob sie diesen Freiraum auch ausnützen. Es geht auch darum zu überprüfen, ob und inwieweit Leistungsvereinbarungen eingehalten werden. Einbezogen werden können zudem die Meinungen der wichtigsten Kunden über das neue Funktionieren der Institution. Bei solchen Evaluationen handelt es sich freilich bloss um erste Anfänge einer Wirkungsüberprüfung des NPM.

Es ist in gewissem Sinne legitim, dass die an einem NPM-Projekt Beteiligten die Evaluation für ihre Interessen einspannen. Problematisch wird es allerdings, wenn mit Evaluationen Etikettenschwindel betrieben wird. Dies kann etwa in der Form erfolgen, dass NPM-Projektverantwortliche eine Verbesserungsevaluation in Auftrag geben, ein intensives Coaching erhalten und am Schluss die Resultate für eine Bilanz des NPM-Versuchs verkaufen. Ein Etikettenschwindel liegt m.E. auch dann vor, wenn eine rein betriebswirtschaftliche und damit verwaltungsinterne Analyse herangezogen wird, um die "Wirksamkeit" des NPM-Versuchs zu belegen. Es entspricht einem Gebot der Objektivität, bei der allfälligen Präsentation von Studien in der Öffentlichkeit deren Hauptzweck (Begleitung/Verbesserung/Entwicklung, Bilanz/Rechenschaft) und deren inhaltliche

Reichweite (verwaltungsinterne Prozesse bis zu den Outputs, Impacts und Outcomes⁴) klar darzulegen.

These 5: Eine zu grosse Nähe von Evaluatoren zu NPM-Verantwortlichen kann die Glaubwürdigkeit der Evaluation gefährden

Das Nebeneinander verschiedener Evaluationstypen eröffnet somit durchaus ein gewisses Stör-, wenn nicht gar Missbrauchspotential. Auch Bilanzevaluationen können in ihrer Übungsanlage gefährdet sein, wenn sie - vor allem bei längerer Dauer - für Verbesserungen am NPM-Projekt genutzt werden. Diese Situation ist in der pädagogischen und psychologischen Evaluation durchaus bekannt (Behandlungs-, Beobachtungs-, Testeffekte usw.). Ein schönes Beispiel dafür ist die Evaluation der Gesamtschule Dulliken:

1970-1978 wurde die Gesamtschule Dulliken als Schulversuch für eine integrierte Oberstufe eingerichtet. Der Versuch wurde von Psychologischen Institut der Universität Bern eingehend evaluiert. Trotz der positiven Resultate der Evaluation wurde der Schulversuch dann aber abgebrochen. Der Grund lag in der intensiven wissenschaftlichen Begleitung. Daraus ergaben sich Verbesserungen des Schulversuchs. Jenzer (1981: 71) äussert sich über den zweideutigen Beitrag der Evaluation wie folgt: "Dank ihrer Mithilfe hat der Schulversuch ... ein gewisses Raffinement erreicht, welches von anderen Schulzentren nicht einfach übernommen werden kann. ... Die Evaluation hat mit ihren 'Fallstudien' wesentlich dazu beigetragen, den 'Sonderfall' von Dulliken zu konstituieren, der, unter anderem weil er ein 'Fall' ist, für eine Generalisierung im ganzen Kanton nicht in Frage kommt (Jenzer 1981: 71)".

Der Fall "Dulliken" zeigt sehr schön, dass an eine Evaluation unterschiedliche Erwartungen gerichtet werden. Die Politiker wünschten eine

⁴ Als Outputs werden hier die Produkte der Verwaltung, als Impacts die Auswirkungen auf das Verhalten der Zielgruppen/anvisierten Kunden und als Outcomes die gesellschaftlichen Auswirkungen bezeichnet (vgl. BUSSMANN, KLÖTI und KNOEPFEL 1997, S. 103-115).

Bilanzevaluation, die Lehrerschaft der Gesamtschule eine Verbesserungsevaluation. Das Psychologische Institut der Universität Bern verpflichtete sich für ersteres. Durch die intensiven Kontakte mit der Gesamtschule (es wurden mehrere Lizentiatsarbeiten und Dissertationen erstellt) ergab sich für die Forscherinnen und Forscher dann aber auch eine Gelegenheit, aus dem Elfenbeinturm herauszutreten und die Praxisrelevanz ihrer wissenschaftlichen Arbeit unter Beweis zu stellen.

Vor diesen Verlockungen werden auch Evaluationen des NPM nicht gefeit sein. Die Interessen der NPM-Verantwortlichen und das Anliegen von Wissenschaftlern und Evaluatoren, praxisrelevante Empfehlungen abzugeben und zu administrativem Lernen beizutragen, werden starke Anreize bilden, auf die Anliegen der NPM-Projektleiter und Mitarbeiter einzugehen. Dies kann namentlich bei Quervergleichen zu grossen Verzerrungen führen. Eine mögliche Abhilfe gegen diese Gefahr wäre entweder eine sehr distanzierte Evaluation oder möglicherweise die ebenso intensive Beratung eines "Verwaltungszwillings" ohne NPM.

Letztlich besteht ein inhaltlicher Gegensatz zwischen der Bilanz- und der Verbesserungsfunktion. Beide Funktionen lassen sich in derselben Evaluation nicht erfüllen. Wenn es wichtig ist, die NPM-Reform zu beraten, dann sollte diese Funktion von jener der kritischen (Bilanz)-Evaluation getrennt werden. Die Aufgabe der Überwachung und Verbesserung kann beispielsweise einer internen Steuerungsgruppe übertragen werden, welche über ausreichenden Evaluations-Sachverstand verfügt. Diese kann auch ein Monitoring/Controlling⁵ sicherstellen (vgl. Bussmann 1995). Die kritische Bilanzierung ist hingegen einem externen Evaluationsteam anzuvertrauen. Beide Gruppe arbeiten mit Vorteil zusammen, da das Wissen der Steuerungsgruppe auch für die externe Evaluation bedeutsam ist (Holkeri und Summa 1996: 22). Dabei ist jedoch darauf zu achten, dass die Unabhängigkeit und Objektivität der externen Evaluation sichergestellt bleiben.

⁵ Das Controlling ist paradoxerweise unter diesem Namen nur im deutschsprachigen Raum bekannt; im angelsächsischen Sprachraum gibt es unter anderem Namen (evaluation, performance review u.a.m.) z.T. ähnliche Ansätze.

These 6: Die Evaluation wird vorderhand keinen Entscheid erlauben, wie erfolgreich NPM ist

Die NPM-Übungsanlage ist darauf ausgerichtet, mit den bestehenden Pilotprojekten ausreichend Erfahrung zu sammeln, um danach darüber zu entscheiden, ob NPM flächendeckend eingesetzt werden kann. Dies ist zumindest die Hoffnung der NPM-Protagonisten, und viele NPM-Verantwortliche werden durch diese Idee beflügelt. Die Erfahrungen mit andern Experimenten, beispielsweise im Schul-, Strafvollzugs- und Heimbereich, lassen indessen die Prognose als realistisch erscheinen, dass dies kaum der Fall sein wird.

Die vorangegangenen Ausführungen haben gezeigt, dass bei der Evaluation von NPM-Versuchen Fallstricke in ausreichender Zahl vorhanden sind. Die methodischen Schwächen von solchen Evaluationen sind bei der Vergabe von solchen Aufträgen noch zu wenig bewusst. Sie werden ohne Zweifel aber dann thematisiert werden, wenn es darum geht, über die Verallgemeinerung des NPM-Modells zu entscheiden. Politiker werden sich gestützt auf Wissenschaftler unterschiedlicher Disziplinen und unterschiedlicher politischer Ausrichtung ausgiebig über methodische Fragen auslassen.

Ausserdem ist durchaus möglich, dass die allfälligen positiven Wirkungen des NPM unter den anderen Faktoren, die die Leistungen der Verwaltung beeinflussen, beispielsweise der Persönlichkeitsfaktor, das Politikfeld, die föderalistische Zusammenarbeit, die konstruktive oder ablehnende Haltung des Parlaments u.v.a.m., nur beschränktes Gewicht haben und auch im Falle methodisch sauberer Bilanzevaluationen im allgemeinen Rauschen untergehen. Wenn wiederum die Schulversuche als Analogie herangezogen werden können, dann geben die damit gemachten Erfahrungen Anlass zur Vermutung, dass die Unterschiede zwischen Modellen im Vergleich zur Leistungsvarianz innerhalb der betreffenden Modelle gar nicht dermassen ins Gewicht fallen. Es ist jedenfalls eine Grunderfahrung der quantitativ orientierten, bilanzierenden Evaluationsforschung, dass es häufig äusserst schwierig ist und in vielen Fällen auch eine recht grosse Zahl von Versuchseinheiten braucht, um Wirkungen statistisch mit ausreichender Zuverlässigkeit nachzuweisen.

These 7: Entscheidend ist, wo NPM gut ist und nicht ob es gut ist

Möglicherweise sind sowohl die NPM-Reformen wie auch allfällige Evaluationen noch vom Glauben an das ideale Modell der Verwaltungsführung inspiriert. Es ist vermutlich an der Zeit, diesen Glauben über Bord zu werfen, weil er zu unfruchtbaren Grabenkämpfen führt. In der Personalführung jedenfalls hat die frühere Auseinandersetzung über harte und weiche Führungsstile seit längerem einer differenzierten Sicht Platz gemacht: der situativen Führungslehre. Sie gibt beiden Führungsstilen gleichermaßen Platz.

Auch in der Verwaltungslehre gibt es Ansätze zu einer differenzierten Sicht der Modelle der Verwaltungsführung. Zu erwähnen ist beispielsweise das 4-Kreise-Modell der Bundesverwaltung, das von vier unterschiedlichen Gegebenheiten für die Führung ausgeht, je nach der Nähe zur politischen Entscheiden auf der einen bzw. der Nähe zu funktionierenden Märkten auf der andern Seite⁶.

Viele Divergenzen über den Einsatz des NPM entspringen unterschiedlichen Perspektiven über das relevante administrative Tätigkeitsfeld. Prof. Matthias Finger, ein vehementer Befürworter des NPM, orientiert beispielsweise seine Vorstellungen primär an halbstaatlichen Organisationen wie der Telecom und der Post, die mehr und mehr in Wettbewerbsmärkten arbeiten. Prof. Peter Knoepfel, ein ebenso vehementer Gegner des NPM, orientiert sich umgekehrt an den "politiques publiques", somit an Politiken wie der Umwelt- oder Raumplanungspolitik, die durch hoheitliches, z.T. auch kooperatives Handeln, jedenfalls aber (bis heute) durch die Abwesenheit von Märkten gekennzeichnet sind. Es ist durchaus möglich, dass in den jeweiligen Gebieten ganz unterschiedliche Voraussetzungen für den Einsatz von NPM-Modellen bestehen.

⁶ Die 4 Kreise sind: 1. Kernverwaltung (Ministerialfunktion, insbesondere Querschnittämter), 2. Bundesämter mit Leistungsauftrag und Globalbudget (sog. FLAG-Ämter), 3. Öffentliche Anstalten und Unternehmen (verwaltungsexterne Leistungserbringung im Bundesbesitz unter öffentlichem Recht wie Post, SBB), 4. Gemischtwirtschaftliche und private Unternehmungen (Leistungserbringung unter privatem Recht, Unternehmen mit öffentlicher Beteiligung wie Telecom, Swisscontrol).

Damit könnte sich die Auseinandersetzung darüber, *ob* NPM ein geeignetes Modell der Verwaltungsführung ist, auf die Frage verlagern, *wo* NPM geeignet ist. Ziel müsste sein, ein besseres Verständnis darüber zu entwickeln, in welcher Situation welches Verwaltungsmodell angebracht ist (sog. contingency approach oder situativer Ansatz).

Das Erarbeiten eines situativen Ansatzes der Verwaltungsführung ist allerdings eine viel anspruchsvollere Angelegenheit als die Beantwortung der Frage, ob NPM geeignet ist. Auf letztere Fragen werden Evaluationen vorderhand widersprüchliche Antworten liefern. Je nach dem Kontext, in welchem NPM angewendet wird (Politikbereich, staatliche Ebene u.a.m.) werden nämlich die Ergebnisse unterschiedlich ausfallen. Die Suche nach den massgebenden Variablen für den Erfolg oder Misserfolg des NPM ist jedenfalls eine Herausforderung, die sowohl sehr viel empirische Arbeit wie auch theoretische Weiterentwicklung erfordert. Nur durch gemeinsame Anstrengungen der Verwaltungswissenschaft und durch internationale Vergleiche (Holkeri und Summa 1996: 23) werden sich Fortschritte erzielen lassen. Es ist deshalb nötig, dass in der Schweiz die empirische Arbeit in Form von Evaluationen mit der Theorieentwicklung vernetzt werden⁷. Ein Vorbild könnte hier die finnische Verwaltungswissenschaft sein, die ein koordiniertes Evaluationsprogramm für NPM-Reformen erarbeitet hat und einen regen internationalen Kontakt pflegt (vgl. etwa NASCHOLD 1995; HOLKERI und SUMMA 1996).

These 8: Die Verwaltung der Zukunft wird nach unterschiedlichen Modellen geführt werden

Die "Einheit der Verwaltung" war im Staats- und Verwaltungsrecht des letzten und des beginnenden 20. Jahrhunderts eine ausserordentlich wichtige Maxime. Sie galt primär für das Auftreten und die Entscheide (Verfügungen) der Verwaltung gegenüber der Gesellschaft, prägte aber auch ihren inneren Aufbau. Ersteres ist im Zuge der Segmentierung der Verwaltung nach Politikfeldern (wie Umweltschutz), ja Politik-Teilfeldern (wie Fischerei, Gewässerschutz, Wassernutzung, um ein für Bewil-

⁷ So hat die Schweizerische Evaluationsgesellschaft eine Arbeitsgruppe "Verwaltungsreform" eingesetzt, in die Evaluation von NPM-Reformen behandelt wird.

ligungsverfahren bedeutsames Beispiel zu wählen) weitgehend illusorisch geworden; letzteres dürfte im Gefolge der aufkommenden Experimente möglicherweise ebenfalls bald der Vergangenheit angehören.

Denn die Erfahrungen mit Experimenten der vergangenen Jahrzehnte, namentlich im Schulbereich, zeigen, dass Versuchsprojekte ein zähes Leben haben. Weder werden sie abgebrochen, wenn sie nicht ganz so erfolgreich wie erwartet sind, noch werden sie flächendeckend umgesetzt, wenn sie bisherigen Modellen überlegen sind. Herkömmliche Verwaltungsformen und neue Modelle werden deshalb über geraume Zeit koexistieren (und die Überlegungen unter These 7 lassen dies auch durchaus als sinnvoll erscheinen) bzw. die Verwaltungsformen werden sich mit der zu erwartenden Persistenz von Reformen möglicherweise noch multiplizieren. Damit werden sich der Verwaltungswissenschaft ein reiches Betätigungsfeld und der herkömmlichen Finanz- und Verwaltungskontrolle nicht zu unterschätzende Herausforderungen stellen.

* * *

Bibliographie

BUSSMANN Werner, Instrumente der Erfolgskontrolle, in: *VOP*, Bd. 6 (1995), S. 345-351.

BUSSMANN Werner, Konturen eines neuen Paradigma der Verwaltungswissenschaft, in: *Verwaltung & Management*, H. 6 (1996), S. 336-339 und H. 1 (1997), S. 25-20.

BUSSMANN Werner, KLÖTI Ulrich, KNOEPFEL Peter, *Einführung in die Politikevaluation*, Basel 1997.

CAMPBELL Donald T., Reforms as Experiments, in: Weiss Carol H., *Evaluating Action Programs, Readings in Social Action and Education*, Boston 1972, p. 186-223.

HALDEMANN Theo, *New Public Management. Ein neues Konzept für die Verwaltungsführung des Bundes?*, Bern 1995.

HOLKERI Katju, SUMMA Hilka, *Evaluation of Public Management Reforms in Finland. From ad hoc Studies to a Programmatic Approach.*

- Paper presented for OECD meeting on „Contemporary Developments in Performance Management“ (1996), H. 4-5.
- JENZER Carlo, Die Evaluation des Schulversuchs „Gesamtschule Dulliken“, in: Meylan Jean-Pierre (Hg.), *Evaluation von Innovationen im Bereich der Primarschule*, Bern/Stuttgart 1983, S. 61-71.
- KNOEPFEL Peter, Plädoyer für ein tatsächlich wirkungsorientiertes Public Management, *Swiss Political Science Review* 2 (1) 1996, S. 151-164.
- KNOEPFEL Peter, Total Quality Management et fédéralisme. Le point de vue de l'analyse des politiques publiques/Total Quality Management und Föderalismus. Betrachtungen aus der Sicht der Politikanalyse, *Cahiers de l'IDHEAP*, No. 159, Chavannes-près-Renens 1996.
- KNOEPFEL Peter, ZIMMERMANN Willi (unter Mitarbeit von Sailer G., Matafora, E.), Evaluation des BUWAL. Expertenbericht zur Evaluation der Luftreinhaltung, des ländlichen Gewässerschutzes und der UVP des Bundes. Schlussbericht, Bern 1992. (zitiert in: BBl 1992 III; 1502 und FF 1992 III: 1424 gemäss Bericht der Geschäftsprüfungskommission des Nationalrates vom 18. Mai 1992).
- KNOEPFEL Peter, BAITSCH Christof, EBERLE Armin, Überprüfung der Aufbauorganisation des Amtes für Umweltschutz des Kantons St. Gallen, *IDHEAP*, Chavannes-près-Renens/St. Gallen 1995.
- MEYLAN Jean-Pierre (Hg.), *Innovation und Evaluation von Schulversuchen*, Bern/Stuttgart 1981.
- MEYLAN Jean-Pierre (Hg.), *Evaluation und Innovationen im Bereich der Primarschulen*, Bern/Stuttgart 1983.
- NASCHOLD Frieder, Ergebnissteuerung, Wettbewerb, Qualitätspolitik. Entwicklungspfade des öffentlichen Sektors in Europa, Berlin 1995.
- POLLITT Christopher, Justification by Works or by Faith? Evaluating the New Public Management, *Evaluation*, 1/2 (1995), p. 133-154.
- SCRIVEN Michael, *Evaluation Thesaurus*, 4th Ed. Sage, Newbury Park 1991.