

INTOSAI



Leitfaden

Programmevaluierung

Leitfaden Programmevaluierung

Juli 2010

INTOSAI, Internationale Organisation der Obersten Rechnungskontrollbehörden
Arbeitsgruppe Programmevaluierung

Vorsitz
Didier Migaud
Präsident des französischen Rechnungshofs
Vorsitzender der Arbeitsgruppe Programmevaluierung
Cour des Comptes
13 Rue Cambon
F-75100 Paris
tel: ++33 (1) 42 98 95 74
fax: ++33 (1) 42 98 96 02
presidence@ccomptes.fr

Mitglieder

Belgien • Chile • Costa Rica • Deutschland • El Salvador • Finnland • Gabun
• Georgien • Libyen • Litauen • Marokko • Mexiko • Österreich • Pakistan •
Papua-Neuguinea • Polen • Schweiz • Südkorea • USA

Niederlande und Schweden, die der Arbeitsgruppe nicht mehr angehören, haben
aktiv deren Tätigkeit in den vergangenen Jahren unterstützt.

Vorwort



**Vorwort von Didier Migaud, Präsident des französischen Rechnungshofes,
Vorsitzender der INTOSAI-Arbeitsgruppe Programmevaluierung**

Im Rahmen des 14. INTOSAI-Kongresses, der 1992 in Washington D.C. stattfand, wurde die Arbeitsgruppe Programmevaluierung eingesetzt und deren Vorsitz dem französischen Rechnungshof übertragen. Programmevaluierung erschien damals als vielversprechende Initiative zur Stärkung und Verbesserung der Prüfungstätigkeit der staatlichen Finanzkontrolle.

Dank der aktiven Mitarbeit ihrer 20 Mitglieder und der Beiträge der Rechnungshöfe, die die Fragebögen beantwortet haben, konnte unsere Arbeitsgruppe ihre Kenntnisse im Bereich der Evaluierung vertiefen: Wie ist Evaluierung zu definieren und abzugrenzen gegenüber der Wirtschaftlichkeitsprüfung? Was sind die wichtigsten Schritte und Vorgaben für ihre Durchführung? Jeder INTOSAI-Kongress, bei dem die Arbeitsgruppe über ihre Tätigkeiten berichtet hat, hat deren Mandat aufs Neue bestätigt: den Einstieg in Programmevaluierungen für alle ORKB zu erleichtern und die Fachbegriffe, Methoden und Verfahren zu erläutern.

Die Evaluierung unterscheidet sich stark von unserer sonstigen Prüfungstätigkeit, einschließlich der Wirtschaftlichkeitsprüfung, die der Evaluierung wohl am nächsten kommt. Sie ergänzt nicht nur die Wirtschaftlichkeitsprüfung, sondern geht über sie hinaus als einziges Instrument, mit dem komplexe öffentliche Maßnahmen erfasst, erklärt, deren Erfolg bemessen, deren Ziele erkannt und deren Umsetzung nachvollzogen werden können.

Daher zwingt uns diese neue Aufgabe der Evaluierungstätigkeit nicht nur in organisatorischer Hinsicht zum Wandel, sondern auch im Hinblick auf unsere angewandten Methoden und Verfahren und sogar unsere interne Organisationskultur. Die Evaluierung komplexer öffentlicher Maßnahmen erfordert strikte Methoden, die Anhörung der verschiedenen Stakeholder und eine zunehmende Inanspruchnahme pluralistischen Sachverständs. Das ist der Preis, zu dem wir den Erwartungen unserer Mitbürger entsprechen, die sich öffentliche Maßnahmen nicht nur transparenter sondern auch wirkungsvoller wünschen.

Dieser Leitfaden, der mit der wertvollen Unterstützung des amerikanischen Rechnungshofes erstellt wurde, basiert auf der gesamten Tätigkeit der Arbeitsgruppe, die diese seit ihrer Gründung durchgeführt hat. Es handelt sich weder um eine Normensammlung noch um eine praktische Anleitung, auch wenn hier zahlreiche konkrete Fallbeispiele der ORKB vorgestellt werden. Der Leitfaden ist erarbeitet worden, um den Einstieg in die Programmevaluierung zu erleichtern, das Verfahren besser zu planen und die wichtigsten Inhalte zu bestimmen. So vorbereitet obliegt es den Rechnungshöfen selbst, einen eigenen Weg einzuschlagen und für jede Evaluierung die bestgeeigneten Methoden und Vorgehensweisen zu finden.

Ein offener Weg liegt vor uns: Es ist nun an uns, diesen zu beschreiten und so viele Partner wie möglich zu überzeugen, uns dabei zu begleiten. Wenn wir wollen, dass Evaluierung in der staatlichen Finanzkontrolle an Bedeutung gewinnt, müssen wir den Wissensaufbau und die Verbreitung optimierter Verfahren fördern. Dies ist das Ziel, das sich unsere Arbeitsgruppe für die kommenden Jahre gesetzt hat und das im Einklang mit der strategischen Planung der INTOSAI steht. Dabei zähle ich auf Sie alle, dabei mitzuhelfen, dieses Ziel zu erreichen!

Didier Migaud

I*nhalt*

Vorwort	iii
Inhalt	v
Einführung	ix
Kapitel 1: Aufgaben der Programmevaluierung	1
[1.1] Neue Perspektiven zur Wirtschaftlichkeit	1
[1.2] Überlegungen zu Strategien des Wandels	11
[1.3] Der Wert in der Evaluierung	15
Fußnoten.....	16
Kapitel 2: Definition der Programmevaluierung	20
[2.1] ORKB definieren ihre Funktionen „Prüfung der Rechnungsführung“ und „Wirtschaftlichkeitsprüfung“ unterschiedlich	20
[2.2] Einige Institutionen definieren Wirtschaftlichkeitsprüfung allgemein so, dass Programmevaluierung einbezogen ist.....	24
[2.3] Evaluierungsdefinitionen anderer internationaler Organisationen.....	27
[2.4] Operative Definition der Programmevaluierung für ORKB.....	30
Fußnoten.....	33

Kapitel 3: Planung der Programmevaluierung.....	38
[3.1] Zusammenarbeit mit den Stakeholdern der Evaluierung	40
[3.2] Festlegung eines Evaluierungskonzepts	49
[3.3] Verknüpfung des Konzepts mit Evaluierungsfragen	51
[3.4] Beurteilung, ob ein Team zur Ausführung von Evaluierungsplänen geeignet ist.....	54
Fußnoten.....	56
Kapitel 4: Durchführung der Programmevaluierung.....	60
[4.1] Einsatz der Evaluierungsmethodik	60
[4.1.a] Einige direkt verfügbare Methodenquellen	64
[4.1.b] Einige indirekt verfügbare Methodenquellen	69
[4.2] Setzung eines Rahmens für die Projektmanagementplanung	77
Fußnoten.....	83
Kapitel 5: Aufbau von Organisationskapazität für Programmevaluierung	88
[5.1] Evaluierungsteam	88
[5.2] Entwicklung organisatorischer Kompetenz	93
[5.3] Evaluierungsgemeinschaft.....	96
Fußnoten.....	102

Kapitel 6: Zukunftsperspektiven	105
[6.1] ORKB als Vermittlerinnen des Wandels	105
[6.2] Wege zu neuem Wissen.....	107
[6.3] Unterstützung einer Evaluierungskultur	108
Fußnoten.....	112
Anhang A: Nützliche Links.....	114
Anhang B: Literaturhinweise.....	121
Anhang C: Glossar	124
Verzeichnis der Veröffentlichungen der ORKB	142

Abkürzungen

ACF	Administration for Children and Families
AEA	American Evaluation Association
AfrEA	African Evaluation Association
ANAO	Australischer Rechnungshof
ASPE	Assistant Secretary for Planning and Evaluation
DAC	Development Assistance Committee (OECD)
Danida	Danish International Development Assistance
EEA	European Evaluation Association
EU	Europäische Union
GAO	Rechnungshof der USA
HHS	U.S. Department of Health and Human Services
INTOSAI	Internationale Organisation der Obersten Rechnungskontrollbehörden
IOCE	International Organisation for Cooperation in Evaluation
OECD	Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung
ORKB	Oberste Rechnungskontrollbehörde
SEVAL	Schweizer Evaluierungsgesellschaft
USAID	U.S. Agentur für internationale Entwicklung

Einführung

Dieser Leitfaden zeigt Wege auf, wie ORKB die Evaluierung „zu einem wesentlichen Bestandteil [ihrer] täglichen Arbeit . . . ethischen Grundsätze, internen Organisationskultur und Aufgabenerfüllung“¹ machen können. Bei der Abschlussitzung der 16. Jahreskonferenz der AEA 2001 wurde vorgeschlagen, statt der Vergabe von Evaluierungen, darauf aufmerksam zu machen, dass

Evaluierung zwar schwierig, aber möglicherweise wirklich nutzbringend ist, aber nur diejenigen daran mitwirken sollten, die es tatsächlich versuchen wollen . . . „Nicht jeder ist dazu in der Lage! Nur wer wirklich nach Qualität strebt, sollte sich daran versuchen. Wenden Sie umgekehrte Psychologie an!“²

Später in der Sitzung wurde gefolgert, dass

die Tatsache, dass die Evaluierungsdefinition nicht allgemeine Zustimmung findet, kein Problem ist. „Die Stärke unserer Fachrichtung ist ihre Vielfalt. Selbst wenn man sich auf den gemeinsamen Begriff Evaluierung verständigt, so gibt es gleichwohl viele verschiedene Herangehensweisen. Wir haben unseren Standpunkt klargemacht... Bewertungen, Sachverhaltsfeststellungen und Beibringung von Nachweisen ziehen sich durch alle Evaluierungsarten. Die Aufgabe besteht darin, diese auf unterschiedliche Fälle anzuwenden“³.

Am Ende der Sitzung schloss James R. Sanders die Konferenz mit folgenden Worten:

Die Einführung der Evaluierungstätigkeit bei Organisationen bzw. Behörden als fester Bestandteil der Aufgabenwahrnehmung sehe ich weiterhin als gemeinsames Ziel an. Evaluierung wird Teil der Kultur, eine gemeinsame Sprache Die Frage, was umfassend verbreitet und eingeführt wird, ist entscheidend. Meiner Meinung nach geht es darum, eine gemeinsame Denkweise zu fördern, kritische Fragen zu stellen, Antworten zu nutzen, um

einen Wandel herbeizuführen, weiterhin nach optimierten Verfahrensweisen zu streben . . . Als Prüfer sollte man immer auch die eigenen Evaluierungen auf den Prüfstand stellen Es ist Teil der Verbreitung und Einführung gleicher Vorgehensweisen — auch die eigene Aufgabenerfüllung zu hinterfragen⁴.

Die Einführung in den Leitfaden Programmevaluierung mit Anmerkungen über den Abschluss einer Evaluierungstagung mag dem Leser zunächst merkwürdig erscheinen. Und dennoch ist die Einführung gleicher Verfahren — „eine *Evaluierungskultur für staatliche Maßnahmen, Verwaltungen und Strategien aufzubauen*, [auch wenn] es sich dabei nicht immer um Evaluierung im engeren Sinne handelt“ — genau das, womit sich der vorliegende Leitfaden befasst.⁵

Rechnungshöfe verfügen bereits über eine große Anzahl an Berichten über Prüfungen der Rechnungsführung, Wirtschaftlichkeits-, Querschnitts-, und Maßnahmenprüfungen sowie sonstige Veröffentlichungen, von denen mehrere beim näheren Hinsehen einer Evaluierung nahe kommen, wenn sie auch nicht explizit genannt werden. So beschreibt beispielsweise der Rechnungshof von Wales in seinem strategischen Plan für den Zeitraum 2004–2009 ein Programm, das nur anhand von Evaluierungsmethoden durchführbar ist, wie es die sechs Kapitel des vorliegenden Leitfadens aufzeigen:

Der walisische Rechnungshof kann seinen Nutzen für die öffentliche Verwaltung über den seiner Vorgängerorganisationen hinaus verbessern, indem . . . er mit einem größeren Kreis von Stakeholdern system- und sektorübergreifend zusammenarbeitet und Einfluss ausübt.

Der walisische Rechnungshof verfügt über eine Reihe von Möglichkeiten, um sein Potential besser ausschöpfen zu können.

wie . . . Unterstützung und Respekt für unsere Rolle, Verbesserungen in der Verwaltung durch Stakeholder zu fördern, [und] . . . unsere Fähigkeit, einen Einblick in alle öffentlichen Dienstleistungen in Wales zu gewinnen . . .

Zur besseren Aufgabenplanung und -wahrnehmung hat der walisische Rechnungshof einen Matrixansatz verwendet, um Personal und Fachwissen

den entsprechenden Aufgaben zuzuordnen. Der Rechnungshof wird die vorhandenen Methoden und Verfahren stärker ausschöpfen, die vorhandenen Fachkenntnisse erweitern und die Matrixstruktur innerhalb des Rechnungshofes fördern, . . . Methoden entwickeln zur Evaluierung der Ergebnisse von Partnerschaften und Kooperationen, die Adressatenorientierung sicherstellen, die Rolle des gemeinnützigen, kommunalen und privaten Sektors prüfen, Systeme evaluieren . . .⁶

Welche Ansprüche stellt die Programmevaluierung an die Rechnungshöfe? Es ist nötig, Strategien des Wandels auszuarbeiten, um die Evaluierung in der Organisationskultur zu verankern, die Fähigkeit zu erwerben, z.B. Evaluierungen zu konzipieren, zu planen und durchzuführen und die Zusammenarbeit mit den jeweiligen Auftraggebern zu gestalten. Ebenso ist eine neue Art des Lernens erforderlich im Hinblick auf das Wissen, wie man nach Datenquellen und Analysemethoden sucht und deren Qualität gewährleistet. Eine umfassende Evaluierungskompetenz ist aufzubauen; das heißt, dass über den einzelnen Prüfer und das Evaluierungsteam hinaus der ganze Rechnungshof insgesamt seine Kompetenzen in dieser Hinsicht weiterentwickelt. Programmevaluierung verlangt neue Perspektiven in der internen Planung. Und schließlich erfordert sie eine Toleranz gegenüber Arbeits- und Denkweisen, die mit den Gepflogenheiten zu brechen scheinen. Sie erwartet, dass die ORKB in die Zukunft schauen und neue Wege im Lernen und ihrer Aufgabenwahrnehmung beschreiten.

Der Leitfaden behandelt die Grundlagen der Programmevaluierung, die zum Teil auch schon angewandt wurden — einige implizit, andere explizit. Viele Beispiele im Text — und auch in den neben stehenden Kästen — befürworten diese Grundkonzepte und erläutern diese. Alle wurden direkt von Rechnungshöfen, Fachverbänden im Bereich Evaluierung und anderen Fachgruppen übernommen. Die Fallbeispiele reichen von Bangladesch bis Neuseeland, von Frankreich bis zur Volksrepublik China und von Georgien bis Brasilien.

Kapitel 1 behandelt den Zusammenhang zwischen Wirtschaftlichkeitsprüfung und Programmevaluierung und zeigt einige Möglichkeiten auf, durch übergreifende Tätigkeiten Evaluierungskompetenz auszubauen. Der Aufwand bei Wirtschaftlichkeitsprüfungen kann ganz unterschiedlich sein. Im Rahmen der Wirtschaftlichkeitsprüfung können übliche Prüfungsrichtlinien oder eine Vielfalt

anderer, auch sozialwissenschaftlicher Methoden, angewandt werden. Und sie können in unterschiedlichem Umfang den Forschungs-, Analyse- oder Evaluierungsbereich einbeziehen. Rechnungshöfe brauchen keinen speziellen Auftrag, um Programmevaluierungen durchzuführen.

Kapitel 2 zeigt auf, wie ORKB weltweit die Bezeichnungen Prüfungen der Rechnungsführung, Wirtschaftlichkeitsprüfungen, Programmevaluierungen und verwandte Begriffe verwenden. Einige Definitionen sind in nationalen Gesetzen enthalten, andere aus der jeweiligen Prüfungspraxis abgeleitet. Eine Sichtung der Fachliteratur, in der untersucht wurde, wie Programmevaluierung an verschiedenen Orten und im Laufe der Zeit definiert wurde, ergab, dass es keine allgemeingültige Begriffsdefinition gibt, weder für die Gesamttätigkeit noch für Teilbereiche.

Am Ende von Kapitel 2 wird eine Arbeitsdefinition entwickelt — die eine ausreichende Grundlage für die Erarbeitung einer eigenen Definition bieten soll, um z. B. auch den Mitgliedern von AfrEA, die an einer praxisorientierten Definition für die Finanzkontrolle in Afrika mitarbeiten und „die Relevanz der U.S.-‘Programmevaluierungsnormen’ für die Evaluierungsarbeit vor Ort“ in ihren Evaluierungsrichtlinien 2000: Checkliste für Prüfer hinterfragen, entgegenzukommen.

Kapitel 3 und 4 fassen die Diskussion über konkrete Aufgaben der Programmevaluierung zusammen. Kapitel 3 stellt einige der Herausforderungen bei Planung und Konzeption vor und schildert die Beschränkungen, denen Prüfer bei der Berichterstattung über den Erfolg von Strategien, Maßnahmen u.ä. unterliegen.

Kapitel 4 nennt nützliche Methodenquellen für die Evaluierungstätigkeit. Evaluierungsgesellschaften und -verbände berufsständischer Prüfer liefern eine Fülle von „Checklisten“, „Handbüchern“, „Matrizen“ und verschiedenen anderen Datenquellen und Analysemethoden. Diese sind oft als Druckversion bzw. elektronisch verfügbar und dienen als Referenzmaterial bei der Wahl einer geeigneten Methodik für Evaluierungen.

Kapitel 5 schildert Besonderheiten der Programmevaluierung. Der Aufbau von Programmevaluierungskapazitäten erfordert die Entwicklung neuer Personalkompetenzen und Methoden innerhalb der ORKB. Das bedeutet auch, sich

mit der Gemeinschaft externer Evaluierungsexperten zusammenzutun, die in der Wissenschaft, als Berater und weltweit in Evaluierungsverbänden tätig sind.

Schließlich bietet Kapitel 6 eine Zukunftsperspektive: Auswirkungen auf strategische Beziehungen zwischen Finanzkontrolle und Evaluierungsgesellschaften, zwischen INTOSAI-Arbeitsgruppen bzw. zwischen öffentlichen und privaten Institutionen.

Das Dokument schließt mit drei Anhängen. Anhang A, „Nützliche Links“, zeigt internationale Organisationen auf, die Stipendien, Studienbeihilfen, Kursarbeiten, Workshops, Stellenangebote und andere Ressourcen für den Aufbau von entsprechender Fachkapazität bereitstellen. Anhang B ist ein Literaturverzeichnis und Anhang C ein Glossar der Evaluierungstermini, die im vorliegenden Dokument verwendet werden.

Die Schaffung einer Einheit für Programmevaluierung innerhalb einer bestehenden Organisation erfordert konzentrierten Einsatz und Langzeitverpflichtung für die zuständigen Führungskräfte und Prüfer — und einige risikobehaftete Entscheidungen. Die Förderung einer Evaluierungskultur verlangt einen längerfristigen Arbeitsaufwand, um einen formalen, geregelten Ablauf der Planung, Durchführung und Behandlung der Ergebnisse einzurichten und aus Analyse und Experiment Erfahrungen zu gewinnen. Was für die einzelne Organisation innerhalb des wachsenden Netzwerks der weltweiten Evaluierungsgesellschaften wirklich am besten ist, erfordert starkes Engagement.

In diesem Leitfaden wird keine einfache Formel für organisatorischen oder kulturellen Wandel geliefert. Das Dokument ist weder ein Lehr- noch ein Übungsbuch, sondern vermittelt einen Überblick über Konzepte, Fallbeispiele sowie Informationsquellen und stellt Methoden und Ansätze vor. Manchen Lesern mag vieles bereits bekannt sein, anderen wird der Leitfaden neue Ideen, Mittel und Hilfe für Wandel und Erweiterung der Prüfungstätigkeit liefern.

1 Siehe: James R. Sanders, “Mainstreaming Evaluation,” in: J. Jackson Barnette und James R. Sanders, Hg., *The Mainstreaming of Evaluation*, New Directions for Evaluation Nr. 99

(San Francisco, Calif.: Jossey-Bass, 2003), S. 3. Sanders was speaking at the American

Evaluation Association's 2001 conference: siehe American Evaluation Association, Evaluation 2001, 7.-10. November, St. Louis, Missouri. <http://www.eval.org/Training/conferencehistory.asp> (7. April 2009)).

- 2 David D. Williams und Mark L. Hawkes, zitieren M. Q. Patton, in: "Issues and Practices Related to Mainstreaming Evaluation: Where Do We Flow from Here?" in: Barnette and Sanders, *The Mainstreaming of Evaluation*, S. 77.
- 3 Williams und Hawkes, quoting Patton, in: Barnette und Sanders, *The Mainstreaming of Evaluation*, S. 77.
- 4 Williams und Hawkes, zitieren Sanders, in: Barnette und Sanders, *The Mainstreaming of Evaluation*, S. 79. Sanders war 26 Jahre als Beigeordneter Leiter des Evaluierungszentrums an der Western Michigan Universität bis zu seinem Ruhestand 2001 tätig, hatte den Vorsitz des Gemeinsamen Ausschusses für Standards for Educational Evaluation 1988 bis 1998 inne und wurde 2001 Präsident der AEA.
- 5 Dignan, "Mainstreaming Evaluation or Building Evaluation Capability? Three Key Elements," in: Barnette and Sanders, *The Mainstreaming of Evaluation*, S. 12.
- 6 Auditor General for Wales, *Making Public Money Count: Wales Audit Office Strategy 2006–2009* (Cardiff: Wales Audit Office, Mai 2006), S. 3–4 and 9–10. Wales Audit Office, Publications, Our Plans and Accounts. www.wao.gov.uk (7. April 2009).

Kapitel 1: Aufgaben der Programmevaluierung

DIE SECHS KAPITEL DES VORLIEGENDEN DOKUMENTS fordern den Leser auf, die Programmevaluierung auf neue Weise zu betrachten. Diese wird häufig unter die Kategorie „Wirtschaftlichkeitsprüfung“ subsumiert. Die kanadische ORKB gab im Frühling 2003 das Diskussionspapier „Ergebnisberichte: Leistungserwartungen und -resultate“ heraus, das anregt „die Berichterstattung über Wirtschaftlichkeit . . . durch die überzeugende Darstellung konkreter Fallstudien zu ergänzen . . .“¹ Im Folgenden werden der Nutzen der Programmevaluierung für die Berichterstattung und der mögliche Zusatznutzen für ORKB, die Evaluierungskompetenzen für die Qualitätssteigerung aufbauen, dargestellt.

[1.1] Neue Perspektiven zur Wirtschaftlichkeit

Mit Interesse an dieser neuen Perspektive vermerkte INTOSAI kürzlich, dass

es heute ein gemeinsames Ziel der meisten Staaten ist, die Qualität der öffentlichen Dienstleistungen zu verbessern, zumal die Erwartungen der Bevölkerung (häufig mit Bezug auf die aus dem privaten Sektor bezogenen Dienstleistungen) an diese Qualität weiterhin zunehmen. Um diese Art von Verbesserungen zu fördern, betreiben viele Staaten die Modernisierung der öffentlichen Verwaltung, damit diese leichter zugänglich und unbürokratischer wird, den Bürgern eine größere Palette an Leistungen anbietet, die zügig erbracht werden. Die Qualität der öffentlichen Dienstleistungen ist ein Thema von wachsender Bedeutung, bei dem

Parlament und Regierung in aller Welt von der staatlichen Finanzkontrolle einen Beitrag in der Berichterstattung über Wirtschaftlichkeitsprüfungen erwartet².

[1] WERT DER EVALUIERUNG:

[W]as bedeutet „gut“ überhaupt . . . ? Wie Wittgenstein feststellte, steht hinter „gut“ . . . eine ganze Reihe von Bedeutungen. Die gängigste davon ist folgende: „entspricht den Kriterien oder Maßstäben, die der Bewertung zugrunde gelegt werden.“

John R. Searle, *Speech Acts: An Essay in the Philosophy of Language* (New York: Cambridge University Press, 1969), S. 152

ORKB können für Erhebungszwecke als Einrichtungen angesehen werden, die üblicherweise Prüfungen der Rechnungsführung nach den Kriterien Rechenschaftspflicht, Recht- und Ordnungsmäßigkeit, Kontrolle und Korrektur (u. a. Brasilien und Griechenland) durchführen bzw. als Organisationen, die auch Wirtschaftlichkeitsprüfungen vornehmen und sich dabei außer auf Sparsamkeit und Wirtschaftlichkeit auch auf Wirksamkeit konzentrieren (z. B. Südkorea und Schweden) und als Gruppe, die häufig über die Wirtschaftlichkeitsprüfungen hinausgeht, sich politische Strategien vornimmt und öffentliche Maßnahmen bewertet (Frankreich und USA).³ Die Rolle der Programmevaluierung für ORKB wurde ausdrücklicher in einem Dokument aus dem Jahre 2004 unter folgendem Titel dargestellt: „Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen: Richtlinien und Grundsätze für die Wirtschaftlichkeitsprüfung, beruhend auf den INTOSAI-Richtlinien für die Finanzkontrolle sowie den Erfahrungen der Praxis“:

In den letzten Jahren war „Programmevaluierung“ ein wichtiges Diskussionsthema. Dabei ging es um die Bedeutung der Programmevaluierung für die Organe der staatlichen Finanzkontrolle. Die INTOSAI-Arbeitsgruppe für Programmevaluierung wurde eingesetzt, um die Erarbeitung von Grundsätzen und Arbeitshilfen zu fördern. Es wird allgemein anerkannt, dass Programmevaluierung ähnliche Ziele wie die Wirtschaftlichkeitsprüfung verfolgt, nämlich das Verhältnis zwischen Zielvorgaben und Mitteleinsatz einerseits sowie die Ergebnisse staatlicher

Maßnahmen andererseits zu analysieren. Die Aussage, dass Programmevaluierung eine wichtige Aufgabe für ORKB ist, die über die entsprechende Befugnis und Sachkompetenz verfügen, hat ebenfalls allgemeine Zustimmung gefunden.⁴

Offenbar berücksichtigen ORKB zunehmend die „Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen“ und gestalten ihre Arbeitsmethoden entsprechend um. Diesem Trend folgt auch die jordanische ORKB, die in ihrem Strategischen Plan 1998 - 2008 den „Verzicht auf vorgängige Prüfungen [und] . . . die verstärkte Durchführung von Wirtschaftlichkeitsprüfungen“ als Langzeitziele nennt und dazu anmerkt, dass es zu ihren Aufgaben gehört, „Methoden [vorzuschlagen], die zur sparsamen, wirtschaftlichen und wirksamen Mittelverwaltung beitragen“.⁵

Die Spanne der heutigen Prüfungstätigkeit kann von Evaluierungen, die ohne besondere Hervorhebung in Prüfungsberichte einfließen, bis zu Evaluierungen reichen, die in einer eigenen Kategorie aufgeführt werden und bei denen die Prüfer explizit als Programmevaluatoren bezeichnet werden. Viele, die Wirtschaftlichkeitsprüfungen durchführen und dabei herkömmlichen Prüfungsrichtlinien folgen, werden evaluierend tätig, indem sie umfangreichere Methoden einschließlich sozialwissenschaftlicher Methoden anwenden und z. B. Beziehungen zwischen Mitteleinsatz und Zielvorgaben eines Vorhabens bzw. Ziele und Ergebnisse einer politischen Maßnahme prüfen. Schauen wir uns z. B. einen Bericht vom 19. Februar 2009 aus den Niederlanden mit dem Titel „Vergleichende Studie über die Berichte von Aufsichtsorganen: Gute Beispiele und Ratschläge für bessere Berichte“ an. Die zuständigen Prüfer befassten sich offensichtlich sowohl mit den Auswirkungen von Maßnahmen als auch mit Fragen der Wirtschaftlichkeit. Sie haben fest gestellt, dass

„Good Governance“ in den letzten Jahren ganz oben auf der Tagesordnung sowohl im öffentlichen als auch privaten Sektor stand. Im öffentlichen Sektor hat es zur Entwicklung einer Vielzahl von entsprechenden Regelwerken geführt. Eine gute interne Aufsicht über das Management ist eines der Schlüsselprinzipien guter Unternehmens- bzw. Regierungsführung . . .

Aufsichtsorgane sind ein relativ neues Phänomen im öffentlichen Sektor und ihre endgültige Form muss sich in der Praxis noch herauschälen. Die Rechenschaftspflicht der Aufsichtsorgane für ihre eigene Leistung ist auch noch in Entwicklung. Es gibt große Unterschiede in der Art, in der sie Rechenschaft ablegen.

...

Die Bedeutung von Aufsichtsorganen und ihrer öffentlichen Rechenschaftspflicht gegenüber Politik und Gesellschaft hat den Rechnungshof dazu veranlasst, sich näher mit den von ihnen herausgegebenen Berichten zu befassen. Es wurden 84 Berichte (aus dem Berichtsjahr 2007) und Websites von Aufsichtsorganen von Sozial- und Bildungseinrichtungen bzw. Verwaltungsbehörden ausgewertet. Sie wurden mit den in den jeweiligen Sektoren anwendbaren Regeln guter Unternehmensführung und den maßgeblichen, vom Rechnungshof angewendeten Normen verglichen. Die Auswertung sollte Aufschluss über den Informationswert der Berichte der Aufsichtsorgane geben und zur künftigen Entwicklung dieser Organe als Elemente der Fachaufsicht beitragen.⁶

Kanada liefert ein anderes Beispiel einer Wirtschaftlichkeitsprüfung, die in Programmevaluierung übergeht. In „Reform der öffentlichen Verwaltung: Fortschritte, Rückschläge und Herausforderungen“ aus dem Jahre 2001 berichtet der kanadische Rechnungshof, dass

die Enquête-Kommission für Haushaltswesen und Rechenschaftspflicht (Lambert-Kommission) 1979 feststellte, dass – wenn es um die staatliche Aufgabenerfüllung gehe – die Personalwirtschaft in allen ihren Aspekten genauso wichtig, wenn nicht wichtiger sei, als die Bewirtschaftung der finanziellen Mittel, dass jedoch die parlamentarische Kontrolle der Personalwirtschaft eingeschränkter sei als diejenige der Bewirtschaftung der finanziellen Mittel durch den Rechnungsprüfungsausschuss. Dies ist auch heute noch so.⁷

Eine Prüfung des kanadischen Rechnungshofes

der Fortschritte und Herausforderungen der Exekutive bei Reformbestrebungen im öffentlichen Dienst im Verlauf des letzten Jahrzehnts, über die er dem Parlament gegenüber umfassend berichtet und eine Stellungnahme zu den Erfolgsvoraussetzungen abgegeben hat,⁸

stellt fest, dass trotz der bisherigen Reformanstrengungen

auf Gebieten wie Modernisierung und Verbesserung der öffentlichen Dienstleistungen sowie bessere Auswertung von Wirtschaftlichkeitsinformationen zur Stärkung von Maßnahmen, Erzielung gewünschter Resultate und Verbesserung der Rechenschaftslegung über die Wirtschaftlichkeit weiterhin erhebliche Anstrengungen erforderlich sind.⁹

Die Prüfung stützte sich auf „eine Unmenge an Dokumenten sowie der 1999 durchgeführten großen Mitarbeiterbefragung“ sowie auf Interviews

mit 13 Fachleuten in obersten bzw. Oberbehörden und einer Vielzahl anderer [einschließlich] . . . einem Dutzend Führungskräften in der nationalen Hauptstadtregion, . . . acht Regionalchefs, . . . und vier Persönlichkeiten, die das ganze Jahrzehnt lang oder einen Teil davon leitende Gewerkschaftsfunktionäre waren.

Um Erinnerungslücken der Befragten auszugleichen „wurden statistische Informationen analysiert, die das Ausmaß des Wandels anzeigen...“¹⁰

Selbst ein Projekt in Finnland, mit dem die Überwachung der Produktsicherheit geprüft wurde, kann als evaluierende Prüfung verstanden werden:

Die Produktsicherheitsüberwachung wurde nicht aktiv und sorgfältig wahrgenommen. . . . Ein Hinweis auf den passiven Überwachungsansatz ist, dass die lokalen Behörden sich der wesentlichen Grundsätze nicht bewusst waren, welche die Verbraucherschutzbehörde für die kommunale Produktsicherheitsüberwachung als wichtig ansieht... Der Umfang der Überwachung stand in keinem angemessenen Verhältnis zur Größe der

Gemeinde, ihrer Industriestruktur oder den Risiken. Einige kleine Landgemeinden investierten viel mehr in die Überwachung als manche großen Städte.

Eine direkte Evaluierung der Überwachungswirksamkeit kann auf Grundlage der Prüfung nicht vorgenommen werden. Die staatlichen Aufsichtsbehörden waren der Ansicht, dass die Wirksamkeit weitgehend von der Überwachung durch die lokalen Behörden abhängt . . . Die Verbraucherschutzbehörde nimmt an, dass manche Unfälle in Haushalt und Freizeit durch mangelhafte Produkte und Leistungen mitverursacht wurden und, dass selbst die beschränkte Überwachung in gewissem Maß Gesundheitsrisiken, Unfälle und sogar Todesfälle verhindert. Detaillierte Informationen zur Überwachung sind jedoch nicht verfügbar.¹¹

Ein letztes Beispiel stammt aus Brasilien, aber im Internet auf internationaler Ebene sind noch viel mehr Fallbeispiele zu finden und eventuell noch mehr in den noch nicht ins Netz gestellten ORKB-Archiven. In einem Bericht aus dem Jahre 2003 über eine Evaluierung von Maßnahmen zur Erkennung und Behandlung von Fehlsichtigkeit beschreibt der brasilianische Rechnungshof den Prüfungsanlass folgendermaßen:

In Anbetracht der Tatsache, dass Zugang zu Bildung ein verfassungsmäßiges Grundrecht ist, sollte evaluiert werden, inwieweit das Nationale Gesundheitsprogramm für Schulkinder dazu beigetragen hat, den Lernfortschritt von Schulkindern mit Fehlsichtigkeit zu verbessern. Dies wurde durch Analyse von fünf Fragestellungen durchgeführt:

- 1 Inwieweit trägt die Kampagne [„Schau mir ins Auge“] zur Verbesserung der Schulleistungen der Zielgruppe bei?
- 2 Ist das Kriterium für die Auswahl der von der Kampagne im Jahre 2001 erfassten Gemeinden dasjenige, das dem Gleichheitsgrundsatz am besten entspricht?
- 3 Hat sich die Zahl der Schüler im Laufe der Kampagne 2000 verringert?

-
- 4 Erfolgte die Auslieferung der Brillen an die Schüler in der Kampagne 2000 fristgerecht?
 - 5 Würden neue Partnerschaften den Zugang zu dem Programm verbessern?¹²

Wie im Kapitel 3 ausgeführt, handelt es sich um typische Fragen, die den Versuch dokumentieren, eher eine Evaluierung als eine Prüfung durchzuführen.

Neben dem Prüfungsbericht über Zweck, Umfang und Methodik sowie Ergebnisse der Evaluierung und den Empfehlungen wurden die Vorteile dargelegt, welche die Umsetzung der Empfehlungen den brasilianischen Kindern und Gemeinden sowohl quantitativ als auch qualitativ bringen würde.

Mehrere dieser Vorteile würden sich aus bedeutenden Programmänderungen ableiten — u.a. Änderung der Zeitpläne, Einführung eines IT-Systems, Informationsverbreitung an Bundesstaaten, Stadtgemeinden, Landgemeinden, Familien und Presse sowie Bildung von Partnerschaften im Gesundheitswesen.¹³

Die „INTOSAI-Leitlinien für die Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen“ heben hervor, dass

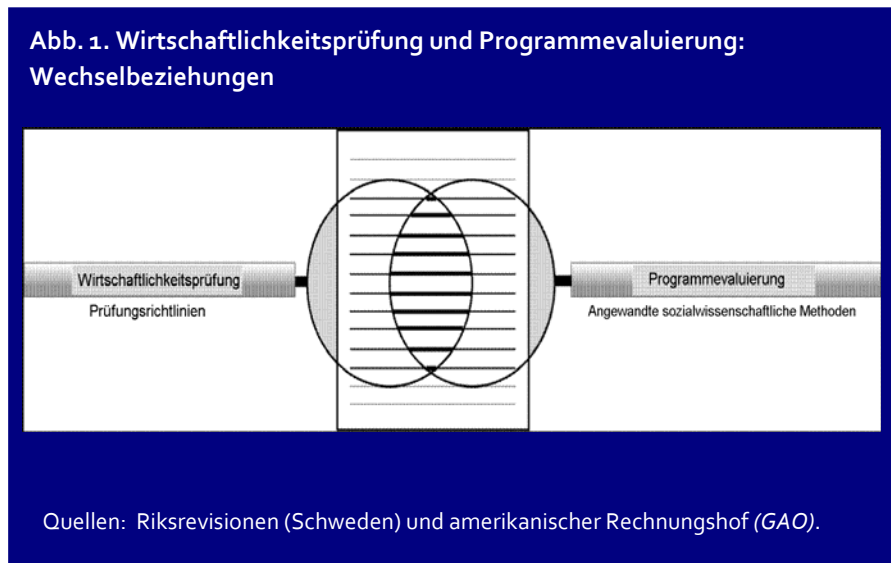
manche ORKB keine Wirtschaftlichkeitsprüfungen durchführen müssen oder sich selbst ein wenig eingeschränkt in ihrer Kapazität und Erfahrung bezüglich solcher Prüfungen erachten können. Andere ORKB können eine lange Tradition in der Durchführung sowohl von weitergehenden Wirtschaftlichkeitsprüfungen als auch von komplexen Programmevaluierungen haben.¹⁴

Wirtschaftlichkeitsprüfung und Programmevaluierung sind häufig miteinander verwoben. Der Arbeitsumfang bei Wirtschaftlichkeitsprüfung kann eng oder weit sein.

Abb. 1 zeigt ein Modell der vielen Möglichkeiten für ORKB evaluierend tätig zu werden. Bei Wirtschaftlichkeitsprüfungen können sie herkömmliche

Prüfungsrichtlinien (siehe linke Seite der Abbildung) oder eine Vielfalt anderer auch sozialwissenschaftlicher Methoden (siehe rechte Seite) anwenden. Und sie können verschiedene Forschungs-, Analyse- oder Evaluierungsniveaus einbeziehen.

ORKB brauchen keinen speziellen Auftrag für die Durchführung einer Programmevaluierung. Jede ORKB kann Evaluierungen gesondert vornehmen oder sie in eine Wirtschaftlichkeitsprüfung einbeziehen (Mitte von Abb. 1). Je nachdem welcher Kurs eingeschlagen wird, kann dies zu einem organisatorischen Wandel führen und die Institution sogar aufwerten..



Die Republik der Philippinen, die anmerkte, dass sie im Jahre 2002 den Namen ihres Rechnungshofs in „Verwaltungsbehörde“ geändert hat, verzeichnet unter den Funktionen dieses Amtes, dass es u. a. „ressortübergreifende Wirtschaftlichkeitsprüfungen“ durchführt.

Die philippinische ORKB stellt im Jahre 2009 auf ihrer Website fest, dass eine ressortübergreifende Prüfung

sich auf die Prüfung von Vorhaben oder Maßnahmen bezieht, die von mehr als einer Regierungsstelle umgesetzt werden und Folgendes liefern soll:

- einen Gesamtüberblick über die Umsetzung der verschiedenen Programmsegmente, der ggf. die Identifizierung von verbesserungswürdigen Bereichen ermöglicht, . . .
- [eine] Grundlage für Prüfer, um zu erkennen, dass festgestellte Mängel nicht von einer einzigen Stelle verursacht wurden, sondern möglicherweise durch die Art der Zusammenarbeit der in die Maßnahme einbezogenen Stellen; . . .
- [eine] Gelegenheit, um gegebenenfalls Programmänderungen vorzunehmen.¹⁵

Die „INTOSAI-Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen“ berücksichtigen diese Arten von Vorhaben, indem sie zwei Ansätze für die Wirtschaftlichkeitsprüfung anführen:

Beim ergebnisorientierten Ansatz geht es hauptsächlich um Fragen wie: „Wie hoch ist die Wirtschaftlichkeit bzw. welche Resultate wurden erzielt, wurden die Anforderungen erfüllt bzw. die Ziele erreicht?“ Bei diesem Ansatz werden die Wirtschaftlichkeit im weitesten Sinne (Sparsamkeit, Wirtschaftlichkeit und Wirksamkeit), vorgegebene Normen (Zwecke, Ziele, Vorschriften usw.) bzw. Prüfungskriterien (die vor Prüfungsbeginn zu definieren sind) untersucht. . . .

Beim problemorientierten Ansatz . . . geht es vornehmlich um Problemfeststellung und Problemanalyse, normalerweise ohne Bezug auf vordefinierte Prüfungskriterien . . . Eine Hauptaufgabe . . . besteht darin, das Vorhandensein bestimmter Probleme zu prüfen und die Gründe dafür unter verschiedenen Perspektiven zu analysieren (Probleme in Bezug auf Sparsamkeit, Wirtschaftlichkeit und Wirksamkeit von staatlichen Unternehmen oder Maßnahmen). Beim problemorientierten Ansatz geht es um Fragen wie: „Bestehen die angegebenen Probleme wirklich und wenn ja,

[2] VERSCHIEDENE PRÜFUNGS- UND EVALUIERUNGSARTEN DER ORKB:

Ist die Wirtschaftlichkeit im weitesten Sinne der Maßstab für die Bewertung, können sowohl auf die Abschlussprüfung als auch auf die Wirtschaftlichkeitsprüfung spezialisierte Prüfer zum Einsatz kommen. In einigen Ländern kann eine individuelle Wirtschaftlichkeitsprüfung viele unterschiedliche Untersuchungsarten und sogar mehrere Programmevaluierungen einschließen. In diesem Sinne kann die Programmevaluierung als eine der vielen möglichen Instrumente betrachtet werden, die bei der Wirtschaftlichkeitsprüfung genutzt werden. Programmevaluierungen können in die Kategorie „Wirtschaftlichkeitsprüfung“ eingeordnet werden. Es handelt sich um eine Prüfungsform, die zunehmend an Bedeutung gewinnt.

INTOSAI Auditing Standards Committee, Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen: Richtlinien und Grundsätze für die Wirtschaftlichkeitsprüfung, beruhend auf den INTOSAI-Richtlinien für die Finanzkontrolle sowie den Erfahrungen der Praxis (Stockholm: Juli 2004), S. 25.

wie können sie nachvollzogen werden und was sind die Gründe?“¹⁶

Abb. 1 rückt beide Ansätze ins Blickfeld. Der ergebnisorientierte Ansatz wird oft durch die Brille der Wirtschaftlichkeitsprüfung gesehen, während der problemorientierte Ansatz in Gebiete der Programmevaluierung übergreifen kann.

[1.2] Überlegungen zu Strategien des Wandels

In der Fachliteratur wird behauptet, dass ein Wandel in der Art der Aufgabenerfüllung einer ORKB häufig einen Wandel der Organisationskultur mit sich bringt. In einer Forschungsarbeit wurde beispielsweise festgestellt, dass ein solcher

Wandel zwangsläufig langfristig angelegt ist, d. h. mindestens 5 bis 10 Jahre bis zu seiner vollständigen Umsetzung erfordert. Dabei wird angenommen, dass für einen erfolgreichen Kulturwandel zwei Dinge von ausschlaggebender Bedeutung sind:

- Die Leitung muss den Wandel uneingeschränkt befürworten und dieses mit Worten und Taten zum Ausdruck bringen.
- Es müssen entsprechende Fortbildungen angeboten werden, die bei den Mitarbeiterinnen und Mitarbeitern diejenigen Kenntnisse und Fähigkeiten fördern und entwickeln, die mit den gewünschten Werten und Überzeugungen im Zusammenhang stehen.¹⁷

In der Untersuchung wurde festgestellt, dass die Maßnahmen, die zur Vermittlung von Werten und Überzeugungen ergriffen werden, für die Wandlungsfähigkeit einer Organisation äußerst wichtig sind: Ausarbeitung eines entsprechenden Grundsatzpapiers und Verteilung an das Personal; ein mit den Werten und Überzeugungen in Einklang stehender Führungsstil, Belohnungen, Anreize und Beförderungen, interne Informationsveranstaltungen sowie die Anpassung der Organisationsstruktur an Wertvorstellungen und Überzeugungen.¹⁸

Studien über den Wandel in der staatlichen Finanzkontrolle, insbesondere die Modernisierung von Prüfungsmethoden, kamen zu ähnlichen Schlussfolgerungen. In Thailand z. B., wo „eine neue Haushaltssystematik den Rechnungshof [veranlasste], mehr Gewicht auf Wirtschaftlichkeitsprüfungen zu legen“, stellte eine Studie mit dem Titel „Strategie zur Verbesserung der Planung von Wirtschaftlichkeitsprüfungen“ fest, dass

die neue Organisationsstruktur zu einem bedeutsamen Wandel innerhalb der ORKB führt: die Wegversetzung qualifizierter Mitarbeiter aus ihren Fachgebieten, Personalabbau usw. [und dass]

die große Nachfrage nach Wirtschaftlichkeitsprüfungen zu zahlreichen Problemen führte:

- Schwierigkeiten bei Auswahl und Festlegung des Prüfungsumfangs,
- unzureichende Prüfungsnachweise zur Stützung der Aussagen der Prüfungsberichte,
- Schwierigkeiten bei der Einigung über Fragen wie Ressourcen, Budgets und Zeitpläne,
- knappe Personal-, Sach- und Finanzmittel, so dass nicht alle geplanten Prüfungen vollständig durchgeführt werden können.¹⁹

Die Studie empfiehlt als beste Strategie

zur Verbesserung des Planungsprozesses sowie zur Unterstützung der Aufgabenerfüllung des thailändischen Rechnungshofes:

- 3.1 die Erlangung eines klaren Verständnisses des Planungsprozesses und der Rollen der wichtigsten Akteure
- 3.2 die Einschaltung externer Sachverständiger,
- 3.3 die Nutzung einer risikobasierten Langzeitplanung,
- 3.4 die Erstellung von Planungsrichtlinien für Wirtschaftlichkeitsprüfungen,
- 3.5 die Erarbeitung von Überwachungsinstrumenten,
- 3.6 die Verbesserung der Kommunikation,
- 3.7 die Förderung von stetigem Lernen.²⁰

[3] DAS GEMEINWOHL — KONTROLLE GEGEN EVALUIERUNG

[K]ontrolle ist in der Verwaltung von großer Bedeutung; ohne sie könnten wir Abweichungen von geplanten Ergebnissen nicht verhindern.

Kontrollieren heißt . . . vor allem abwenden, aufdecken, korrigieren, anleiten . . .

Den [Beamten der] Dienststellen und Institutionen obliegt die Verwaltungsfunktion, wogegen der Rechnungshof als Partner fungiert, der das höchste Ziel jedes Verwaltungsbeamten erreichen möchte: die Förderung des Gemeinwohls.

. . .

Die Wirtschaftlichkeitsprüfung ist das Prüfungsverfahren, das in den vergangenen Jahrzehnten am meisten weiterentwickelt wurde. Diese Prüfungsart spielt bei der Modernisierung der öffentlichen Einrichtungen in vielen Ländern eine wichtige Rolle. Der Zweck der Programmevaluierung besteht darin, die Wirksamkeit von staatlichen Maßnahmen und Vorhaben zu prüfen. Diese Prüfungsart zielt darauf ab festzustellen, inwieweit die umgesetzten Maßnahmen die von der Verwaltung beabsichtigten Wirkungen erreichten.

Der brasilianische Rechnungshof führte Wirtschaftlichkeitsprüfungen auf mehreren Gebieten durch wie z. B. Landwirtschaft, Bildungswesen, Sozialhilfe und Gesundheitswesen. Maßnahmen in diesen Bereichen werden als vorrangig angesehen . . . und haben zum Ziel, die Armut zu bekämpfen sowie die Bürgerrechte zu stärken und die gesellschaftliche Eingliederung zu fördern.

Mit den vom Rechnungshof 2002 durchgeführten Evaluierungen staatlicher Maßnahmen wurde geprüft wie die Regierung Ausgabemittel von 5,7 Mrd. \$ zugunsten der mehr als 80 Millionen Unterstützungsempfänger investiert.

Brasilianischer Rechnungshof, Jahresbericht Haushaltsjahr 2002 (Brasilia: Tribunal de Contas da União, Planning and Management Secretariat, 2003), S. 3 und 36–37. <http://www.tcu.gov.br> (17. April 2009).

Der estnische Rechnungshof hat im Jahre 2002 auf seiner Website ebenfalls eine Studie mit dem Titel „Wirksamkeit von Wirtschaftlichkeitsprüfungen: Der Fall Estland“ veröffentlicht, deren Ausgangspunkt folgender ist:

Wirtschaftlichkeitsprüfungen sind schon längst fester Bestandteil der Tätigkeit vieler Rechnungshöfe. Ihre Verbreitung löste Diskussionen über die Unabhängigkeit der Prüfer und ganz allgemein über die staatliche Finanzkontrolle aus. Immer häufiger, besonders nach Umsetzung umfangreicher Reformprogramme der öffentlichen Verwaltung weltweit, wurden die Kontrollbehörden gezwungen, ihre eigene Tätigkeit zu rechtfertigen und ihre Ergebnisse und Leistungen darzustellen. Die Verfahren zur Beurteilung der Wirksamkeit der Wirtschaftlichkeitsprüfung haben deshalb an Bedeutung gewonnen.²¹

Ähnlich wie bei der Sachkompetenz für die Programmevaluierung sei dies zu einem nicht unerheblichen Teil der Tatsache zu verdanken, dass

eine Wirtschaftlichkeitsprüfung eine Kombination aus herkömmlicher Prüfung und Unternehmensberatung ist und dabei Kenntnisse und Techniken aus vielen Fachgebieten einbezieht Sie profitiert von der Unabhängigkeit und Objektivität der Prüfer und ihrer Fähigkeit zu einer sachgerechten Berichterstattung. Diese Qualitäten werden durch spezialisierte Analysensysteme und Umsetzungs Kompetenzen ergänzt, die bei Unternehmensberatern zur Verfügung stehen.

...

Es gibt keine standardisierten Methoden oder Verfahren für die angemessene Durchführung einer Wirtschaftlichkeitsprüfung. Um erfolgreich zu sein, sind geeignete Methoden und Techniken unter Berücksichtigung der verfügbaren Mittel, des Prüfungsansatzes und des Prüfungsgegenstands zu wählen.²²

Die Auswertung solcher Ergebnisse, Studien sowie Prüfungsempfehlungen kann hilfreich für Rechnungshöfe sein, die die Wirtschaftlichkeitsprüfung um

Evaluierungskompetenzen ergänzen wollen. Ein Bericht schildert wie Dienststellen ihre Programmevaluierungskapazität aufgebaut haben:

Zuerst unterstützte die Leitung der Dienststelle verstärkt die Erfüllung der Rechenschaftspflicht und die Verbesserung der Wirtschaftlichkeit der Aufgabenerfüllung. Damit sollte eine Evaluierungskultur institutionalisiert werden. In einem zweiten Schritt verbesserten sie administrative Systeme oder griffen auf spezielle Datensammlungen zurück, um die Datenqualität zu verbessern. Drittens ermittelten sie mit Hilfe externer Quellen oder durch entsprechende Personalentwicklung den Sachverstand, der zur Sicherstellung der Zuverlässigkeit von Analysen und Schlussfolgerungen erforderlich war. Schließlich suchten die die Zusammenarbeit mit Programmpartnern und Interessengruppen oder vermittelten ihre Anliegen an diese und bemühten sich um deren Unterstützung und Mitwirkung, um Evaluierungsressourcen und Sachverstand noch wirksamer einsetzen zu können.²³

[1.3] Der Wert in der Evaluierung

Wie Kapitel 2 zeigt, definieren ORKB die Prüfung der Rechnungsführung, Wirtschaftlichkeitsprüfung und Programmevaluierung unterschiedlich. Manche Definitionen sind vom Gesetz vorgegeben, andere stammen aus unterschiedlicher Praxis. Die Fachliteratur, die einen Überblick über die historische Entwicklung der Definitionen von Programmevaluierung gibt, kommt zu dem Schluss, dass es keine allgemein gültige Definition gibt. So stellt z. B. im Jahre 2002 die „Kanadische Evaluierungsgesellschaft, Projekt zur Unterstützung von anwaltlicher und berufsständischer Entwicklungsevaluierung (Nutzen, Outputs und Kenntniselemente)“ fest, dass

- es keinen einzigen [Evaluierungs-] Output oder -vorgang gibt, der auf alle Situationen und Zwecke anwendbar ist. Die Fachliteratur liefert kaum Anhaltspunkte,

-
- [Programm-] Outputs und Nutzen ebenso eng miteinander verknüpft sind wie Outputs und Fachwissen bzw. Können.²⁴

Die am Ende von Kapitel 2 vorgeschlagene Arbeitsdefinition der Programmevaluierung liefert eine ausreichende Grundlage, von der jede ORKB ihre eigene Definition ableiten kann.

-
- 1 John Mayne, *Office of the Auditor General of Canada, Reporting on Outcomes: Setting Performance Expectations and Telling Performance Stories* (Ottawa, Ontario: 14. Mai 2003), S. 13 (Hervorhebung im Original).
www.oag-bvg.gc.ca/internet/English/meth_lp_e_907.html (7. April 2009).
 - 2 INTOSAI-Komitee für die Richtlinien für die Finanzkontrolle, Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen *Richtlinien und Grundsätze für die Wirtschaftlichkeitsprüfung, beruhend auf den INTOSAI-Richtlinien für die Finanzkontrolle sowie den Erfahrungen der Praxis* (Stockholm, Schweden: Juli 2004), S. 21. <http://intosai.connexcc-hosting.net/blueline/upload/aimplgperfaudg.pdf> (7. April 2009).
 - 3 Bzgl. der ersten Gruppe, siehe z.B., Brazilian Court of Audit, *TCU Annual Report, Fiscal Year 2002* (Brasilia: Tribunal de Contas da União, Planning and Management Secretariat, 2003), S. 5. www.tcu.gov.br (17. April 2009), und Greece, Supreme Court of Audit, Athens. <http://www.elsyn.gr/elsyn/root.jsp> (17. April 2009). Bzgl. der zweiten Gruppe, siehe z.B., Board of Audit and Inspection of Korea, Responsibilities and Functions, Board of Audit and Inspection Act, Seoul, Korea. <http://www.bai.go.kr> (17. April 2009), and Riksrevisionen, *The Swedish National Audit Office: Booklet 2008* (Stockholm: n.d.), S. 2. www.riksrevisionen.se/templib/pages/NormalWithIntrosPage___881.aspx (17. April 2009). Bzgl. der dritten Gruppe siehe z. B., France, Cour des Comptes, Juridictions financiers, Publications, Rapports publics thématiques, *La formation professionnelle tout au long de la vie* (Paris: Oktober 2008). www.ccomptes.fr/fr/JF/Accueil.html (17. April 2009), and GAO, *Technology Assessment: Using Biometrics for Border Security*, GAO-03-174 (Washington, D.C.: 14. Nov. 2002). www.gao.gov (17. April 2009).

2009).

- 4 INTOSAI-Komitee für die Richtlinien für die Finanzkontrolle, Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen, S. 23.
- 5 ARABOSAI, ISC members, Jordan, Strategic Plan. www.arabosai.org/en/index.asp (17. April 2009).
- 6 Netherlands Court of Audit, Publications, Audits, Comparative Study of the Reports of Supervisory Boards: Good Examples and Tips for Better Reports (The Hague, The Netherlands: 19. Feb. 2009). www.rekenkamer.nl (17. April 2009).
- 7 L. Denis Desautels, Auditor General of Canada, *Public Service Management Reform: Progress, Setbacks and Challenges* (Ottawa, Ontario: Office of the Auditor General of Canada, Feb. 2001), S. 28. Siehe: Office of the Auditor General of Canada, Reports and Publications, Audit Methodology, Studies and Tools. www.oag-bvg.gc.ca/internet/English/admin_e_41.html (17. April 2009).
- 8 L. Denis Desautels, Auditor General of Canada, *Public Service Management Reform: Progress, Setbacks and Challenges*, S. 25.
- 9 L. Denis Desautels, Auditor General of Canada, *Public Service Management Reform: Progress, Setbacks and Challenges*, S. 25.
- 10 John Edwards, "Looking Back from 2000 at Public Service 2000," S. 7, in: L. Denis Desautels, Auditor General of Canada, *Public Service Management Reform: Progress, Setbacks and Challenges*.
- 11 National Audit Office of Finland, Publications, Performance Audit Reports, Product Safety Monitoring, abstract 69/2004 (Helsinki: 2004), S. 2–3. www.vtv.fi (17. April 2009).
- 12 Brazil, Brazilian Court of Audit, Publicações institucionais, Em inglês, Executive Summaries 6, TCU Evaluation of the Actions for Detection and Correction of Visual Impairment (Avaliação do TCU sobre as ações de detecção e correção de problemas visuais) (Brasília: Tribunal de Contas da União, Government Programs Control and Evaluation Secretariat, 2003),

executive summary, S. 8–9. www.tcu.gov.br (17. April 2009).

- 13 Brazil, Brazilian Court of Audit, *TCU Evaluation of the Actions for Detection and Correction of Visual Impairment*, S. 17.
- 14 INTOSAI-Komitee für die Richtlinien für die Finanzkontrolle, Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen, S. 25.
- 15 Republic of the Philippines, Commission on Audit, Government-wide and Sectoral Performance Audit Reports, Quezon City. www.coa.gov.ph (17. April 2009).
- 16 INTOSAI-Komitee für die Richtlinien für die Finanzkontrolle, Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen, S. 26–27.
- 17 GAO, *Organizational Culture: Techniques Companies Use to Perpetuate or Change Beliefs and Values*, GAO/NSIAD-92-105 (Washington, D.C.: 27. Feb. 1992), S. 2. www.gao.gov (17. April 2009).
- 18 Die Liste ist in Anlehnung an GAO, *Organizational Culture*, Tabelle 1, S. 3. entstanden.
- 19 Jarawan Ruangswadipong, *Strategy to Improve Planning for Performance Audits*, Thailand CCAF- FCVI Fellow 2000–01 Report, S. 5. www.ccaf-fcvi.com/fellows/reports/Thailand2000.pdf (17. April 2009).
- 20 Jarawan Ruangswadipong, *Strategy to Improve Planning for Performance Audits*, S. 5.
- 21 Jane Etverk, "Measuring Performance Audit Effectiveness: The Case of Estonia," M.A. thesis, Department of Public Administration, University of Tartu, Tartu, Estonia, 27. Juni 2002, S. 5. National Audit Office of Estonia, About NAO, Documents, Other Documents, Tallinn, Estonia. www.riigikontroll.ee/?lang=en (17. April 2009).
- 22 Jane Etverk, "Measuring Performance Audit Effectiveness," S. 18.
- 23 GAO, *Program Evaluation: An Evaluation Culture and Collaborative*

Partnerships Help Build Agency Capacity, GAO-03-454 (Washington, D.C.: 2. Mai 2003), S. 14. www.gao.gov (17. April 2009).

- 24 Rochelle Zorzi, Martha McGuire, und Burt Perrin, *Canadian Evaluation Society Project in Support of Advocacy and Professional Development: Evaluation Benefits, Outputs, and Knowledge Elements* (Okt. 2002), Anh. C, *Literature Review*, S. 8.
<http://consultation.evaluationcanada.ca/englishhome.htm> (17. April 2009).

Kapitel 2: Definition der Programmevaluierung

Das vorliegende Kapitel empfiehlt, eine mehr oder weniger einheitliche Definition des Begriffs Programmevaluierung zu verwenden. Zuerst wird erläutert, dass ORKB die Prüfung der Rechnungsführung und die Wirtschaftlichkeitsprüfung aufgrund der Unterschiede in den nationalen Gesetzen, aber auch aufgrund ihrer eigenen Art der Umsetzung des Rahmenkonzepts der INTOSAI-Prüfungsnormen unterschiedlich definieren. Danach wird ausgeführt, dass einige Rechnungshöfe in ihre Definitionen des Begriffs „Programmprüfung“ ein Programmevaluierungskonzept aufgenommen haben, ohne Programmevaluierung ausdrücklich zu definieren. Nachfolgend werden einige Definitionen des Begriffs „Evaluierung“ angeführt, der im Allgemeinen von anderen internationalen Maßnahmen angewendet wird. Schließlich bietet das Kapitel eine Arbeitsdefinition der Programmevaluierung an, die dazu beitragen kann, dass ORKB die Prüfung der Rechnungsführung und die Wirtschaftlichkeitsprüfung systematisch um die Programmevaluierung ergänzen.

[2.1] ORKB definieren ihre Funktionen „Prüfung der Rechnungsführung“ und „Wirtschaftlichkeitsprüfung“ unterschiedlich

Jede ORKB definiert ihre Funktion „Prüfung der Rechnungsführung“ unabhängig im Rahmen des geltenden nationalen Rechts bzw. der INTOSAI-Prüfungsnormen um.¹ Betrachten wir nur zwei Beispiele — Kolumbien und China. In Kolumbien,

[besagt die Staatsverfassung von 1991, Art. 267, dass „die staatliche Finanzkontrolle die öffentliche Aufgabe ist, die der Rechnungshof der Republik wahrzunehmen hat. Dieser überwacht die Haushaltsführung der öffentlichen Verwaltung sowie von Einzelpersonen oder Unternehmen, die öffentliche Finanzmittel oder Vermögenswerte bewirtschaften.“]²

Laut dem chinesischen Rechnungshofgesetz (Art. 1), verabschiedet am 31. August 1994 und novelliert am 28. Februar 2006, ist

dieses Gesetz verfassungskonform mit dem Ziel formuliert, die Staatsaufsicht durch Rechnungsprüfung zu stärken, die Einhaltung des Haushaltsrechts und der Wirtschaftsordnung des Landes zu sichern, die Wirtschaftlichkeit der Verwendung staatlicher Mittel zu verbessern, eine gute Regierungsführung zu fördern und die gesunde Entwicklung der nationalen Wirtschaft und Gesellschaft zu gewährleisten.³

Einige ORKB erweitern die „Prüfung der Rechnungsführung“ um die „Wirtschaftlichkeitsprüfung“, die sie ebenfalls unterschiedlich definieren. Dies tun sie weitgehend in Übereinstimmung mit dem Wortlaut der INTOSAI-Prüfungsnorm 1.0.40 zur Wirtschaftlichkeitsprüfung.⁴ Betrachten wir wieder zwei Beispiele — Australien und Niederlande. Die Studie zur „Wirtschaftlichkeitsprüfung beim australischen Rechnungshof“ hebt Folgendes heraus:

Eine Wirtschaftlichkeitsprüfung ist eine unabhängige, objektive und systematische Prüfung der Maßnahmen, Ressourcen, Informationssysteme, Wirtschaftlichkeitsmaßstäbe im Bereich der öffentlichen Hand sowie der Einhaltung der Rechtsvorschriften und politischen Vorgaben.

Wirtschaftlichkeitsprüfungen spielen eine bedeutende Rolle bei der Verbesserung der öffentlichen Verwaltung.

Wirtschaftlichkeitsprüfungen umfassen die Evaluierung der Umsetzung konkreter staatlicher Maßnahmen, Strategien, Projekte und Aktivitäten. Mit ihnen wird auch geprüft, wie gut die innere Verwaltung unter folgenden Kriterien arbeitet:

-
- Sparsamkeit (Kostenminimierung);
 - Wirtschaftlichkeit (Maximierung des Output/Input-Verhältnisses);
 - Wirksamkeit (inwieweit angestrebte Ergebnisse erzielt werden) und
 - Einhaltung der rechtlichen und politischen Vorgaben.

Es gibt vier Arten der Wirtschaftlichkeitsprüfung:

- Prüfung einer Maßnahme oder einer Tätigkeit einer einzelnen öffentlichen Stelle;
- Prüfung von Sicherheitsvorkehrungen bzw. Geheimschutzmaßnahmen;
- Querschnittsprüfungen und
- Folgeprüfungen.

...

Mit dieser Prüfung lassen sich auch verfahrensmäßige Verbesserungsmöglichkeiten erkennen, die in entsprechenden Leitfäden festgehalten werden können. Mit diesen lassen sich die neuen Erkenntnisse im gesamten Bereich der öffentlichen Hand verbreiten und damit Verbesserungen des Verwaltungshandelns in allen Verwaltungszweigen erreichen.⁵

Eine weitere Veröffentlichung des australischen Rechnungshofes aus dem Jahre 2004 zu Verbesserungsmöglichkeiten bei der Prüfung der jährlichen Leistungsberichterstattung fügt hinzu, dass

Prüfungen der jährlichen Ergebnisberichte nur Teilaspekte der Wirtschaftlichkeit des Verwaltungshandelns abdecken und in der Regel qualitative Resultate nicht hinreichend erfassen können, besonders in Bezug auf Ergebniswirksamkeit Werden Evaluierungen vorgenommen und die daraus gewonnenen Erkenntnisse gut genutzt, um die Wirtschaftlichkeit des Verwaltungshandelns in Jahresberichten zu erläutern, lässt sich eine entscheidende Lücke in der aktuellen Wirtschaftlichkeitsinformation füllen.

Evaluierungen stellen sowohl die Frage, ob eine Dienststelle „das Richtige“ unternimmt und, ob sie dieses „richtig“ tut. Seit zwei Jahrzehnten haben das Finanzministerium und der Rechnungshof die Durchführung von Evaluierungen gefördert. Vom Finanzministerium entwickelte Arbeitshilfen liefern nützliche Hinweise für die Einbeziehung der Evaluierung in ein System des Leistungsmanagements.⁶

Im niederländischen Handbuch für Wirtschaftlichkeitsprüfungen heißt es:

In der [niederländischen] Verfassung sind Organisation, Zusammensetzung und Vollmachten des niederländischen Rechnungshofes durch ein formelles Gesetz geregelt, nämlich die Haushaltsordnung von 2001. Dieses Gesetz enthält auch Regelungen zu den Prüfungsrechten des niederländischen Rechnungshofes . . .

[4] INTERNATIONALE EVALUIERUNGSGEMEINSCHAFT:

Europa ist nunmehr fester Bestandteil der internationalen Evaluierungsgemeinschaft. Einzelne europäische Länder mögen ihre eigenen Wege gehen, sie werden aber gleichzeitig durch ein neues Verwaltungsmodell beeinflusst . . . [D]as Multi-Level-Governance-System der EU (*koordiniertes, auf Partnerschaft beruhendes Vorgehen der Union, der Mitgliedstaaten und der regionalen und lokalen Gebietskörperschaften zur Ausarbeitung und Umsetzung der Politiken der Europäischen Union*) . . . ist eine besondere Mischung aus Traditionen der Evaluierung von Sozialprogrammen und der „neuen öffentlichen Verwaltung“ [*NPM – new public management*].

Nicoletta Stame, Vizepräsidentin der *European Evaluation Society*, „*Evaluation and the Policy Context: The European Experience*,“ (Europäische Erfahrungen mit der Evaluierung im politischen Kontext) Grundsatzrede bei der *International Conference of the Australasian Evaluation Society*, Auckland, Neuseeland, 17. September 2003, S. 10. Canadian Evaluation Society, Grey Literature, Ottawa Ontario. <http://evaluationcanada.ca/site.cgi?s=1> (24. Apr. 2009).

§ 85 der Haushaltsordnung von 2001 verpflichtet den niederländischen Rechnungshof dazu, „die Wirksamkeit und Wirtschaftlichkeit der verfolgten Politik sowie die Mittelbewirtschaftung, die Rechnungen und Belege und die Organisation der Staatsverwaltung zu prüfen.“⁷

Das Handbuch fügt hinzu, dass

vom niederländischen Rechnungshof durchgeführte Wirksamkeits- und Wirtschaftlichkeitsprüfungen auf die Messung von Leistungen (Outputs) und Wirkungen (Outcomes) ausgerichtet sind. Es gibt jedoch verschiedene Möglichkeiten, diese zu prüfen: Man kann die Zielerreichung der geprüften Stelle, die Wirksamkeit der politischen Maßnahmen und die Wirtschaftlichkeit der Outcomes und Outputs analysieren.⁸

[2.2] Einige Institutionen definieren Wirtschaftlichkeitsprüfung allgemein so, dass Programmevaluierung einbezogen ist

Im Juli 2007 gab der GAO eine überarbeitete Fassung seiner staatlichen Prüfungsnormen heraus, in dem die Prüfungen der Rechnungsführung und Wirtschaftlichkeitsprüfungen wie folgt definiert werden:

Wirtschaftlichkeitsprüfungen . . . gewährleisten Sicherheit oder liefern Schlussfolgerungen. Dies geschieht aufgrund der Evaluierung ausreichender und geeigneter Nachweise anhand vorgegebener Kriterien, wie spezieller Anforderungen, Maßstäbe oder festgelegter Verfahrensweisen. Wirtschaftlichkeitsprüfungen liefern eine objektive Analyse, so dass die Leitung und die mit Führung und Aufsicht beauftragten Personen . . . Programmleistungen und Maßnahmen verbessern, Kosten reduzieren, die Entscheidungsfindung der für Aufsicht oder Einleitung von Korrekturmaßnahmen Verantwortlichen erleichtern und die Haushaltskontrolle stärken können. . . . Eine Wirtschaftlichkeitsprüfung ist ein dynamischer Prozess, der im gesamten Prüfungsverlauf die einschlägigen

Normen berücksichtigt. . . . Wirtschaftlichkeitsprüfungen können stark unterschiedliche Ziele haben, u.a. die Beurteilung der Wirksamkeit von Maßnahmen, der Sparsamkeit und Wirtschaftlichkeit, der internen Kontrolle sowie der Recht- und Ordnungsmäßigkeit, und die Vornahme von vorausschauenden Analysen einschließen. Diese allgemeinen Ziele schließen sich nicht gegenseitig aus.⁹

Während die Prüfer für gewöhnlich den allgemein anerkannten staatlichen Prüfungsnormen folgen, können sie bei Wirtschaftlichkeitsprüfungen zusätzlich

andere fachliche Normen anwenden, wie z. B. . . . „Leitlinien für die Evaluierung“ (*Guiding Principles for Evaluators*) der *American Evaluation Association*, „Normen für die Programmevaluierung“ (*Program Evaluation Standards*) des *Joint Committee on Standards for Education Evaluation* sowie die „Standards für edukative oder psychologische Testverfahren“ (*Standards for Educational and Psychological Testing*) der *American Psychological Association*.¹⁰

Beispielsweise hat der Rechnungshof von Mauritius seit Beginn der 60er Jahre eine kontinuierliche Weiterentwicklung erfahren:

Früher haben die staatliche Finanzkontrolle und die Berichterstattung im Wesentlichen zwei Anliegen gehabt:

1. Vollständigkeit und Richtigkeit der veröffentlichten Angaben und
2. Kontrollausübung.

In den 60er Jahren nahmen die Finanzgeschäfte der Regierung an Volumen und Komplexität zu. Die Weiterentwicklung der fachlichen Normen für Rechnungslegung und -prüfung und die geringe Größe des Rechnungshofes erforderten eine umfassende Neubewertung der Rahmenbedingungen für die Prüfung und der Prüfungspraxis. . . .

Die Bestrebungen zur Steigerung der fachlichen Leistungsfähigkeit . . . waren erfolgreich, d. h. der Rechnungshof konnte fachlich qualifiziertes Personal aufstocken. . . .

Gleichermaßen erweiterten sich seine Prüfungsansätze auf Erfolgskontrollen/Evaluierungen von Programmsergebnissen, und seit den 80er Jahren wendet er „Wirtschaftlichkeit“ als Prüfungsmaßstab an. Er führt auch IT-Prüfungen und Organisationsprüfungen durch.

Heute kann er seine Dienste, Erfahrungen und Fortbildungsveranstaltungen anderen Ländern in der Region anbieten . . .¹¹

Obwohl Mauritius den Begriff „Programmevaluierung“ nicht benutzt, so weist ein Prüfungsbericht darauf hin, dass die Programmevaluierung dennoch eine Rolle bei seiner Prüfungstätigkeit spielt. Unter dem Titel „Erosionsbekämpfung in Rodrigues“ wird über ein Projekt berichtet, das auf Minderung und Kontrolle von Bodenerosion abzielt, um „den Lebensstandard der Einwohner von Rodrigues durch Förderung des Gleichgewichts zwischen der Wirtschaftsentwicklung der Insel und der rationalen und nachhaltigen Nutzung ihrer natürlichen Ressourcen zu verbessern.“¹²

Erwähnt wird auch, dass sich der Rodrigues-Bericht „darauf konzentrierte festzustellen, ob die ausführende Dienststelle über die für die Planung, Umsetzung und Überwachung des Projekts ausreichenden Sachkompetenzen verfügt. Die detaillierte Prüfung der Finanztransaktionen trat hinter diesem vorrangigen Prüfungsziel zurück.“¹³ Die Prüfer stellten fest, dass bei dem Projekt zwar alle Betroffenen eingebunden werden sollten, dass eine Strategie zur Umsetzung von Aktivitäten auf Gemeindeebene aber nicht klar definiert, das Projekt- und Fachpersonal in Grundsätzen und Verfahren der Bürgerbeteiligung unzureichend geschult und die Bevölkerung kaum eingebunden war. Überdies war die Überwachung und Evaluierung der Projektmaßnahmen unzureichend.¹⁴ In der Praxis kann es deshalb durchaus vorkommen, dass Wirtschaftlichkeitsprüfungen der ORKB eine gewisse Ähnlichkeit mit den Evaluierungen internationaler Entwicklungsprogramme haben.

[2.3] Evaluierungsdefinitionen anderer internationaler Organisationen

Mehrere internationale Organisationen — ob unabhängige Geberinstitutionen wie die Weltbank oder nationale staatliche Stellen wie die Dänische Internationale Entwicklungshilfe im dänischen Außenministerium — nehmen regelmäßig formale Evaluierungen vor. 2003 hat der Entwicklungshilfesausschuss (*DAC*) der OECD eine Definition der Programmevaluierung empfohlen, die alle wichtigen Geberorganisationen angenommen haben:

Eine Evaluierung ist eine möglichst systematische und objektive Begutachtung von laufenden oder abgeschlossenen Hilfsmaßnahmen, ihrer Planung, Umsetzung und Ergebnisse. Dabei wird die Bedeutung, Zielerreichung, Entwicklungswirtschaftlichkeit, Wirksamkeit, Wirkung sowie Nachhaltigkeit der Maßnahmen beurteilt. Eine Evaluierung sollte glaubwürdige und nützliche Informationen liefern, die die Einbeziehung von Erkenntnissen in den Entscheidungsfindungsprozess sowohl bei den Nehmern als auch den Gebern ermöglichen.¹⁵

Die Definition ist implizit in „Evaluierungsgrundsätze 2006“ enthalten, einem Rahmenwerk, das Danidas Erwartungen zur Qualität der Evaluierungen darlegt. Darin heißt es:

Vor 1982 konzentrierten sich die Evaluierungen vorwiegend auf Einzelprojekte und -maßnahmen. In den meisten Fällen handelte es sich dabei um Zwischen- oder Teilevaluierungen, die durchgeführt wurden, wenn das Projekt sich von einer Phase in die nächste bewegte. Nur wenige Abschlussequalierungen wurden durchgeführt und nur gelegentlich fanden nachträgliche Evaluierungen statt, um die Langzeitwirkung von Projekten zu untersuchen.

Nachdem die Evaluierungsstelle der Danida im Zeitraum 1982–87 aufgebaut worden war, wurde vereinbart, Evaluierungen für systematischere Untersuchungen einzusetzen Zu dieser Zeit fanden zumeist Zwischen- oder Teilprüfungen von Einzelprojekten statt. Diese sollten durch interne

Überprüfungen ersetzt und die Zahl der Abschlussevaluierungen erhöht werden. Der Einsatz von Evaluierungen war systematischer, da er durch einen Evaluierungsjahresplan gesteuert wurde . .

Im Zeitraum 1987–97 verringerte sich die Zahl der Einzelevaluierungen und die Zahl der thematischen und ressortübergreifenden Evaluierungen nahm zu. Grundsätzlich wurden alle Evaluierungsberichte veröffentlicht. Im Jahre 1992 wurde die Information der Öffentlichkeit in Übereinstimmung mit den DAC-Grundsätzen als wesentliches Evaluierungsziel hinzugefügt. In dem genannten Zeitraum wurden die Evaluierungen experimenteller und schlossen sowohl einige Wirkungsevaluierungen als auch den Einsatz von partizipativen Methoden ein . . .

Eine Überprüfung der Evaluierungen bei Danida durch externe internationale Berater ergab im Jahre 2003, dass das Evaluierungssystem der Danida im allgemeinen den internationalen Standards entspricht.¹⁶

[5] ZUNAHME PARTIZIPATIVER EVALUIERUNGEN:

Die Welternährungsorganisation FAO hat im Rahmen ihres *People's Participation Program* in 12 Ländern Afrikas, Asiens und Lateinamerikas Hilfsprojekte unter Beteiligung der Bevölkerung durchgeführt. Etwa 13.200 Personen waren aktive Projektteilnehmer, während die Zahl der Nutznießer insgesamt ungefähr 80.000 beträgt Es ist wichtig, dass möglichst jede Gruppe ihre eigenen Aktivitäten festlegt, plant, durchführt und evaluiert . . .

"People's Participation Program by the FAO [Food and Agriculture Organization of the United Nations]", SD dimensions (Mai 1996).
<http://www.fao.org/sd/spdirect/ppre0006.htm> (24. April 2009)

Die amerikanische Behörde für Internationale Entwicklung nennt auf EvalWeb — ihrer Spezial-Website für Evaluierung — folgende Definition:

Eine Evaluierung ist . . . eine systematische und objektive Bewertung eines laufenden oder abgeschlossenen Projekts, Programms oder einer Strategie. Evaluierungen werden durchgeführt, um:

- a. die Leistung von bereits laufenden Maßnahmen oder bestehenden Strategien zu verbessern,
- b. ihre unmittelbaren und mittelbaren Wirkungen zu bewerten,
- c. als Grundlage für Beschlüsse über die Zukunftsplanung zu dienen.

. . .

Evaluierung ist eine analytische Arbeit zur Beantwortung konkreter Fragen zur Programmsteuerung. Sie bietet einen systematischen Ansatz zur Gewinnung von Erkenntnissen sowie zur Erarbeitung von Schlussfolgerungen über die Wirksamkeit konkreter Maßnahmen, die Gültigkeit von Entwicklungshypothesen, den Nutzen von Leistungsüberwachung und zur Berücksichtigung von übergreifenden Entwicklungsgesichtspunkten, die Auswirkungen auf die Zielerreichung und auf alle Maßnahmen haben können, die zur Verbesserung der Wirtschaftlichkeit und Wirksamkeit eingeleitet werden müssen.¹⁷

Der GAO führt Prüfungen und Evaluierungen gemäß Definition zur Wirtschaftlichkeitsprüfung durch, während seine fachlichen Normen für die Programmevaluierung sowohl die nachgängige Analyse der Wirksamkeit, Wirtschaftlichkeit und Sparsamkeit als auch die vorgängige Analyse einschließt, wie in seinen „staatlichen Prüfungsnormen, Stand: 2007“ festgehalten:

Die Prüfungsziele bei der vorgängigen Analyse umfassen die Auswertung oder Würdigung von Informationen, die auf Annahmen über Ereignisse beruhen, die eventuell in der Zukunft eintreten und die Konzipierung möglicher Maßnahmen, die die geprüfte Stelle als Reaktion auf künftige

Ereignisse ergreifen kann. Beispiele . . . schließen Schlussfolgerungen der nachstehend genannten Arten ein über

- (a) aktuelle und erwartete Tendenzen und deren künftige potenzielle Auswirkung auf staatliche Maßnahmen und Leistungen;
- (b) Programm- oder Politikalternativen einschließlich der Vorhersage von Programmergebnissen unter verschiedenen Annahmen [und]
- (c) Politik- oder Gesetzgebungsvorschläge einschließlich ihrer Vor- und Nachteile sowie die Analyse der Ansichten der Anspruchsberechtigten . . .¹⁸

[2.4] Operative Definition der Programmevaluierung für ORKB

Sowohl Definitions- als auch Praxisfragen ansprechend, veröffentlichte die Kanadische Evaluierungsgesellschaft im Oktober 2002 einen Projektbericht über Grundlagen zur Programmevaluierung. Die CES versuchte, den Nutzen der Programmevaluierung sowie die erforderlichen Kenntnisse und Fähigkeiten zu beschreiben.

Die Gesellschaft kam zu dem Schluss, dass „es keine allgemein anerkannte Definition der Programmevaluierung gibt“, dass aber

einige Evaluierungsbasismerkmale . . . offenbar weithin anerkannt werden:

- Die Evaluierung wendet wissenschaftliche Grundsätze an, um praktische Fragen zu Maßnahmen zu beantworten.
- Daten werden gesammelt und systematisch verarbeitet, um Feststellungen über die Vorgänge innerhalb einer Maßnahme (Prozesse und Ergebnisse) zu treffen sowie die Gründe für die Vorgänge zu erfahren und Verbesserungsmöglichkeiten erarbeiten zu können. Mittels dieser Feststellungen kann der Nutzen bzw. Wert der Maßnahmen beurteilt werden.
- Die Evaluierung ist eine fächerübergreifende Disziplin, die auf die Methoden einzelner wissenschaftlicher Disziplinen zurückgreift.¹⁹

In ihrem Überblick über die Fachliteratur sind mehrere Definitionen des Begriffs „Programmevaluierung“ aufgeführt. Bei allen diesen Definitionen liegt der Schwerpunkt auf der Verbesserung der Programmwirksamkeit zwecks Erzielung von gesamtgesellschaftlichem Nutzen. Nachstehend drei dieser Definitionen.

Zuerst eine Definition von William Shadish, Thomas Cook und Laura Leviton aus ihrem 1995 erschienenen Werk „Grundlagen der Programmevaluierung“:

Charakteristisch für die Evaluierung ist eine idealisierende Problemlösungssequenz für (a) die Identifizierung eines Problems; (b) das Aufzeigen von Alternativen, um seine Symptome zu reduzieren; (c) die Bewertung dieser Alternativen und schließlich (d) die Annahme derer, bei denen die Ergebnisse darauf schließen lassen, dass sie das Problem zufrieden stellend vermindern werden.²⁰

Zweitens eine Definition von Eleanor Chelimsky und William Shadish aus ihrem 1997 erschienenen Werk „Evaluierung für das 21. Jahrhundert“, in welchem drei wesentliche Bereiche des Nutzens genannt werden:

Diese unterschiedlichen Zwecke scheinen zusammen mit Fragen, auf deren Lösung sie abzielen, von selbst unter drei wesentliche Gesichtspunkte zu fallen:

- Evaluierung zur Kontrolle der Wirtschaftlichkeit im weitesten Sinn (z. B. Erfolgs- oder Wirtschaftlichkeitskontrolle)
- Evaluierung zur Förderung der Entwicklung (z. B. als helfende Evaluierung zur Stärkung von Institutionen)
- Evaluierung zur Gewinnung von Erkenntnissen (z. B. Erwerb tieferer Einsicht in spezielle Gebiete oder Felder).²¹

Drittens eine Definition von Peter Rossi, Howard Freeman und Mark Lipsey aus ihrem 1999 erschienenen Werk „Evaluierung: Ein systematischer Ansatz“:

Evaluierungsforscher wenden sozialwissenschaftliche Methoden an, um Sozialprogramme unter deren wichtigen Aspekten zu untersuchen,

einzuschätzen und verbessern zu helfen. Dazu gehören die Diagnose der sozialen Probleme, die sie lösen sollen, die Gestaltung der Maßnahmen, ihrer Umsetzung und Verwaltung, ihre Ergebnisse und Wirtschaftlichkeit.²²

Betrachten wir schließlich in diesem Zusammenhang die Weiterentwicklung der Programmevaluierung im amerikanischen Rechnungshof (GAO), wie sie von einer Definition aus dem Jahre 1992 und einer anderen aus 2005 abgeleitet werden kann. Im Jahr 1992 definierte der GAO Programmevaluierungen in seinem „Überblick über die Evaluierung“ als „Anwendung von wissenschaftlichen Forschungsmethoden, um Programmkonzepte sowie die Umsetzung und Wirksamkeit von Maßnahmen einzuschätzen.“²³ Dreizehn Jahre später definierte der GAO in seiner Veröffentlichung „Leistungsmessung und Evaluierung: Begriffe und Beziehungen“ Programmevaluierungen erneut als

einzelne periodisch oder von Fall zu Fall durchgeführte systematische Untersuchungen mittels derer eingeschätzt werden soll, wie gut eine Maßnahme funktioniert. Sie werden häufig von programmexternen Fachleuten, entweder innerhalb oder außerhalb der Dienststelle sowie von Programmleitern durchgeführt.

Eine Programmevaluierung untersucht typischerweise die Erreichung der Programmziele im Zusammenhang mit anderen Aspekten der Programmleistung oder in dem Kontext, in dem sie erfolgt. Vier Haupttypen können identifiziert werden, die alle Maßstäbe der Programmwirtschaftlichkeit im weitesten Sinne zusammen mit anderen Informationen verwenden, um den Nutzen eines Programms oder Möglichkeiten für seine Verbesserung festzustellen.²⁴

Das Dokument definiert weiterhin die vier Haupttypen als Prozess- oder Umsetzungs-, Ergebnis- oder Wirkungsevaluierung sowie Kosten-Nutzen- bzw. Kosten-Wirksamkeitsanalyse.²⁵

Unter Berücksichtigung des Vorhergesagten könnten ORKB als allgemeine Definition der Programmevaluierung die folgende übernehmen:

Programmevaluierung ist eine systematische Untersuchung einer Organisation oder Institution, eines Programms, Projekts, Prozesses oder einer Politik, die der Gesellschaft Nutzen bringen soll. Programmevaluierungszweck kann sein, bei Verbesserungen zu helfen, die Entscheidungsfindung zu unterstützen oder die Gründe für Erfolge und Misserfolge oder Stärken und Schwächen bei der Zielerreichung festzustellen. Sie kann als Dokument der Rechenschaftslegung dienen oder vorhandenes Wissen erweitern. In der Regel werden durch Programmevaluierungen Kriterien zur Unterscheidung hoher und niedriger Qualität festgelegt, Wirtschaftlichkeit und Wirksamkeit an solchen Kriterien gemessen und stichhaltige sowie glaubwürdige Schlussfolgerungen aus der Evaluierung gezogen.

Programmevaluierungen können unabhängig erfolgen, auch wenn sie unter Beteiligung der örtlichen Bevölkerung sowie von Verbrauchern, Managern und anderen durchgeführt werden, um unterschiedliche Gesichtspunkte zu berücksichtigen. Zum Evaluierungsprozess können die Identifizierung von Zielen und Schlüsselfragen, die Definition der Evaluierungskriterien, die Einholung und Analyse relevanter Daten, die Berichterstattung über Ergebnisse und die Förderung der wirksamen Nutzung der Feststellungen gehören. Evaluierungen werden von Personen mit guter Vorbildung in Forschungsdesign und Methodik, Sachkenntnis allgemein und auf bestimmten Spezialgebieten sowie fachlicher Kompetenz in Evaluierungstheorie, -methodik und -praxis durchgeführt.²⁶

-
- 1 Siehe: Inga-Britt Ahlenius, Auditor General, Swedish National Audit Office, "Auditing Standards and Ethics: An INTOSAI Perspective," address, 17th Annual Conference on New Developments in Government Financial Management for Government Financial Managers, Miami, Florida, April 2003. International Consortium on Governmental Financial Management. www.icgfm.org/downloads/presentations/IngaBritt_Ahlenius.ppt (18. April 2009).
 - 2 Colombia, Contraloría General de la República, Institucional, ¿Qué es la

Contraloría? www.contraloriagen.gov.co/html/home/home.asp (18. April 2009). English translation, University of Richmond, Constitution Finder, Constitution of Colombia (1991), title X, ch. 1, article 267. <http://confinder.richmond.edu/index.php> (18. April 2009).

- 3 People's Republic of China, National Audit Office, Laws and Regulations, Audit Law, Audit Law of the People's Republic of China, S. 2. www.cnao.gov.cn (18. April 2009).
- 4 Siehe: INTOSAI-Komitee für die Richtlinien für die Finanzkontrolle, *Auditing Standards*, 14. INTOSAI-Kongress, Washington, D.C., 1992; überarbeitet auf dem 17. INTOSAI-Kongress, Seoul, Korea, 2001, Norm 1.0.40 Die Wirtschaftlichkeitsprüfung ... , Kap. 1, S. 33. INTOSAI: Internationale Organisation der Obersten Rechnungskontrollbehörden, Arbeitsgruppen Task Forces; Ziel 1 Fachliche Normen; prüfungsbezogene Arbeitspapiere ; Pflichten- und Verhaltenskodex und Richtlinien für die Finanzkontrolle. www.intosai.org (18. April 2009).
- 5 Australian National Audit Office, *Performance Auditing in the Australian National Audit Office* (Canberra: 2008), S. 3. Australia National Audit Office, Publications, Other Publications. www.anao.gov.au/index.cfm (26. April 2009).
- 6 Australia National Audit Office, Publications, Better Practice Guides, *Better Practice Guide: Better Practice in Annual Performance Reporting* (Canberra: April 2004), S. 37. www.anao.gov.au/index.cfm (26. April 2009).
- 7 Netherlands Court of Audit, *Performance Audit* (The Hague: Feb. 2, 2006), Kap. 2, Abschnitt 2.2.1, S. 14. Netherlands Court of Audit, Publications, Manuals. www.rekenkamer.nl/9282400/v (20. April 2009). Fußnote: "The legislation does not actually make clear what exactly is meant by 'the efficiency of financial and material management, of the records kept for this purpose and of the organisation of central government,'" discusses this.
- 8 Netherlands Court of Audit, *Performance Audit*, Kap. 1, Abschnitt 1.1, S. 8. The only definition of performance audit appears in the manual's definition of performance data (Anhang 9, S. 197, Hervorhebung im Original): "There are three types of performance data: policy objectives, which should be operationalised as far as possible in terms of desired and actual outcomes,

linked to certain target groups and to a specified time horizon; information on the cost price and quality of *outputs* delivered or to be delivered by the government (i.e. products and services); [and] programme expenditure, supported as much as possible by *volume and price data*”

- 9 Comptroller General of the United States, *Government Auditing Standards, July 2007 Revision*, GAO-07-731G (Washington, D.C.: U.S. Government Accountability Office, Juli 2007), Abschnitte 1.25–1.28, S. 17–18.
www.gao.gov/govaud/ybko1.htm (20. April 2009).
- 10 Comptroller General of the United States, *Government Auditing Standards, July 2007 Revision*, sec. 1.03, S. 12.
- 11 Mauritius, National Audit Office, The Office, History, The Development of Audit. www.gov.mu/portal/site/auditsite (20. April 2009).
- 12 Mauritius, National Audit Office, Reports and Publications, *Annual Report 1999–2000*, Vol. 2, *Special Reviews: Ministry for Rodrigues*. Rodrigues Administration, Anti-Erosion Maßnahmen in Rodrigues (Port Louis: 2000), Abschnitt 4.1.4.
www.gov.mu/portal/site/auditsite (26. April 2009).
- 13 Mauritius, National Audit Office . . . *Ministry for Rodrigues*. Rodrigues Administration, *Anti-Erosion Maßnahmen in Rodrigues*, Abschnitt 4.1.1.
- 14 Mauritius, National Audit Office . . . *Ministry for Rodrigues*. Rodrigues Administration, *Anti-Erosion Maßnahmen in Rodrigues*, executive summary.
- 15 Hans E. Lundgren, W. Haven North und Ray C. Rist, *A Review of Evaluation in Danida Ministry of Foreign Affairs* (Copenhagen: Januar 2003), S. 8, Fußnote 1. Siehe: Ministry of Foreign Affairs of Denmark, Development Policy, Evaluation of Development, About the Evaluation Department. www.um.dk/en (24. April 2009).
- 16 Ministry of Foreign Affairs of Denmark, Development Policy, Evaluation of Development, Evaluation Policy & Guidelines, *Evaluation Policy 2006* (Copenhagen: Evaluation Dept., 2006), S. 1.
www.um.dk/en (24. April 2009).

-
- 17 USAID EvalWeb, What is an evaluation?
<http://evalweb.usaid.gov/> (24. April 2009).
 - 18 Comptroller General of the United States, *Government Auditing Standards, Abschnitt 1.32*, S. 22. Siehe auch: GAO, *Prospective Evaluation Methods: The Prospective Evaluation Synthesis*, PEMD-10.1.10 (Washington, D.C.: Nov. 1990). www.gao.gov/special.pubs/pe10110.pdf (24. April 2009).
 - 19 Rochelle Zorzi, Martha McGuire und Burt Perrin, *Canadian Evaluation Society Project in Support of Advocacy and Professional Development Evaluation Benefits, Outputs, and Knowledge Elements* (Toronto: Okt. 2002), Executive Summary, S. ii.
<http://consultation.evaluationcanada.ca/english/home.htm> (24. April 2009).
 - 20 William R. Shadish Jr., Thomas D. Cook und Laura C. Leviton, *Foundations of Program Evaluation: Theories of Practice* (Thousand Oaks, Calif.: Sage Publications, 1995), S. 21, in: Zorzi, McGuire und Perrin, *Canadian Evaluation Society Project in Support of Advocacy and Professional Development*, Anh. C, S. 2.
 - 21 Eleanor Chelimsky and William Shadish, *Evaluation for the 21st Century: A Handbook* (Thousand Oaks, Calif.: Sage Publications, 1997), S. 10, in: Zorzi, McGuire und Perrin, *Canadian Evaluation Society Project in Support of Advocacy and Professional Development*, Anh. C, S. 2.
 - 22 Peter H. Rossi, Howard E. Freeman und Mark Lipsey, *Evaluation: A Systematic Approach* (Thousand Oaks: Calif.: Sage Publications, 1999), S. 21, in: Zorzi, McGuire und Perrin, *Canadian Evaluation Society Project in Support of Advocacy and Professional Development*, Anh. C, S. 2.
 - 23 GAO, *The Evaluation Synthesis*, PEMD-10.1.2 (Washington, D.C.: 1. März 1992), S. 126. www.gao.gov (24. April 2009).
 - 24 GAO, *Performance Measurement and Evaluation: Definitions and Relationships*, GAO-05-739SP (Washington, D.C.: Mai 2005), S. 2. www.gao.gov (24. April 2009).
 - 25 GAO, *Performance Measurement and Evaluation: Definitions and*

Relationships, S. 4. Verschiedene Arten der Evaluierung sind weiter unten aufgeführt Kapitel 4.

26 Angelehnt an Western Michigan University, The Evaluation Center, About Us, About the Center, Kalamazoo. www.wmich.edu/evalctr/about.html (24. April 2009).

Kapitel 3: Planung der Programmevaluierung

Bei der Planung von Programmevaluierungen sind Prüfungsanlass, Verwendung der Prüfungsergebnisse und die jeweiligen Berichtsadressaten zu berücksichtigen. Die Planung umfasst auch die Festlegung von Prüfungsziel und -zweck, die Formulierung von Prüfungsfragen sowie die Auswahl geeigneter Prüfer.

Programmevaluierung ist ein interaktiver Prozess, bei dem die zuständige Stelle, das Prüferteam, die Programmleiter, andere am Evaluierungsergebnis interessierte Stellen sowie gelegentlich auch die Medien mitwirken. Durch diese Denk-, Arbeits- und Kommunikationsweise unterscheidet sich die Programmevaluierung von der Prüfung der Rechnungsführung und der Wirtschaftlichkeitsprüfung, obwohl der Kommunikationsprozess zwischen Prüfern und geprüften Stellen auch seinen anerkannten Nutzen hat.¹

Neben dem beabsichtigten erzieherischen und gesellschaftlichen Nutzen einer Programmevaluierung, ist es „das Ziel, die Erfahrungen hieraus für den gesellschaftlichen Lernprozess verfügbar zu machen.“² Zu den Stakeholdern einer Evaluierung zählen Dritte, die ihren Sachverstand einbringen sowie von der Evaluierung Betroffene, darunter Teilnehmer an den evaluierten Programmen.

Prüfungen der Rechnungsführung fragen in erster Linie danach, ob und welche Daten vorhanden sind, während die Planung einer Programmevaluierung die Ausarbeitung eines entsprechenden Konzepts einschließt. Dies umfasst im Allgemeinen die Art der zu beschaffenden Informationen, die Informationsquellen, die bei der Stichprobenprüfung der Daten anzuwendenden Methoden, Vergleichsmaßstäbe und Planung der Datenanalyse. Voraussetzung für die

Ausarbeitung eines geeigneten Konzepts ist natürlich, dass das Evaluierungsteam aus Prüfern mit den erforderlichen Kenntnissen und Fähigkeiten besteht.

[6] ÜBER INTERAKTIVE KOMMUNIKATION:

Die von unabhängigen nationalen Rechnungshöfen durchgeführte Finanzkontrolle ist in vielen Staaten der entwickelten Welt ein fester Bestandteil des Haushaltszyklus und der Abarbeitung der ihm zugrunde liegenden politischen Vorgaben. Sie beinhaltet die Evaluierung der Rechnung, der Haushalts- und Wirtschaftsführung sowie zunehmend der Ergebnisse staatlicher Maßnahmen und Programme. . . . Aus der Unabhängigkeit der meisten nationalen Rechnungshöfe ergibt sich eine deutliche Trennung zwischen der prüfenden und den geprüften Stellen Eine intensive Kommunikation zwischen der Prüfungsinstanz und den geprüften Stellen während der verschiedenen Phasen des Prüfungsverfahrens scheint ein bedeutsamer Erfolgsfaktor zu sein.

Gerard Bukkems and Hans de Groot, Netherlands Court of Audit, "Learning by Auditing: Mission Impossible?" Paper for the 5th biennial conference of the European Evaluation Society, Three Movements in Contemporary Evaluation: Learning, Theory and Evidence, Seville, 10.-12. Oktober 2002, S. 5. www.evaluationcanada.ca/distribution/2002_1010_bukkems_aerard_de_groot_hans.pdf (27. April 2009).

Im vorliegenden Kapitel werden unter dem Planungsgesichtspunkt Unterschiede zwischen der Prüfung der Rechnungsführung und der Wirtschaftlichkeitsprüfung einerseits und der Programmevaluierung andererseits erläutert. Es beschreibt, wie das Prüfersteam seine Kenntnisse und Erfahrungen gemeinsam mit den Auftraggebern der Evaluierung einsetzt, um die Ziele der Evaluierung zu formulieren und schildert, dass das Evaluierungsteam letztendlich ein Konzept auswählt, das zu mehr Informationen über eine Maßnahme, Dienststelle, ein Vorhaben oder die daran Beteiligten bzw. Nutznießer führt.

[3.1] Zusammenarbeit mit den Stakeholdern der Evaluierung

Stellen, die bisher Prüfungen der Rechnungsführung oder Wirtschaftlichkeitsprüfungen durchführten, eröffnet die Programmevaluierung neue Perspektiven und den Aufbau neuer Sachkompetenzen. Anders als die Fachliteratur zur Programmevaluierung erwähnen die Rechtsgrundlagen und Leitlinien für die Planung von Prüfungen der Rechnungsführung oder der Wirtschaftlichkeit selten die Notwendigkeit der Zusammenarbeit mit den geprüften Stellen oder den Berichtsadressaten.

So verlangt beispielsweise der vietnamesische Rechnungshof von seinem Personal bei Prüfungen der Rechnungsführung, der Recht- und Ordnungsmäßigkeit und der Wirtschaftlichkeit, „die richtige Stelle entsprechend den vom Präsidenten des Rechnungshofes gemachten Vorgaben zu Zielen, Inhalt, Umfang, Ort und Dauer zu prüfen“ und

die geeigneten Prüfungsmethoden anzuwenden, um Sachverhalte zu erheben und zu würdigen, Stichproben und Vergleiche durchzuführen und Bestätigungsvermerke zu erteilen sowie Einrichtungen und Einzelpersonen im Rahmen der Prüfungshandlungen zu befragen, um gestützt darauf Bemerkungen, Bestätigungsvermerke, Schlussfolgerungen und Prüfungsempfehlungen formulieren zu können.³

[7] EINE NORM FÜR DIE PRÜFUNGSPLANUNG:

Das Handbuch zur Wirtschaftlichkeitsprüfung des Rechnungshofs von Bangladesch — Kapitel 3 über die “Planung der Wirtschaftlichkeitsprüfung” — schreibt für die Prüfungsplanung die Abarbeitung der nachstehenden Punkte vor.

- Hintergrundinformationen über den zu prüfenden Bereich
- die wichtigsten politischen Ziele und Instrumente für deren Umsetzung (können aus verschiedenen Grundsatzpapieren und strategischen Plänen entnommen werden)
- Parlamentarisches Interesse, ggf. unter Einbeziehung früherer Prüfungserfahrungen
- Einschätzung der Risikofaktoren, die die Wirtschaftlichkeit mindern könnten
- Prüfungsumfang und wichtigste zu prüfende Fragestellungen
- Klärung, ob im selben Bereich andere Prüfungen geplant sind, insbesondere ob es zu Überschneidungen kommen kann
- die ins Auge gefassten Prüfungsmethoden (Prüfungsansatz und -technik) sowie Quellen für Prüfungsnachweise
- erwartete Prüfungswirkungen in Form von Einsparungen bzw. Verbesserungen
- Personalbedarf und Prüfungskosten
- Prüfungszeitplan.

Comptroller and Auditor General of Bangladesh, Performance Audit Manual (Dhaka, Bangladesh: 26. Dez. 2000), S. 25. Office of the Comptroller and Auditor General of Bangladesh, www.cagbd.org/html/perfaudit.html (27. April 2009).

Auf eine Zusammenarbeit mit den geprüften Stellen und den Nutzern der Prüfungsergebnisse wird weder hier noch in dem Kapitel des Gesetzes über den Staatsrechnungshof Bezug genommen, das sich mit abgeschlossenen Prüfungen befasst:

Die geprüften Stellen haben die Prüfungsschlussfolgerungen und -empfehlungen

. . . zu Unregelmäßigkeiten in den Jahresabschlüssen und zu Rechts- und Ordnungsverstößen umzusetzen bzw. die vom Rechnungshof empfohlenen Abhilfemaßnahmen zu ergreifen.

Dienststellen oder zuständige Personen, die die Prüfungsschlussfolgerungen verwenden, entscheiden unter Beachtung der einschlägigen Rechtsgrundlagen selbst über die Annahme der Prüfungsschlussfolgerungen des Rechnungshofes.

Den Prüfungsschlussfolgerungen ist zu folgen, sobald sie von Dienststellen bzw. zuständigen Personen gebilligt worden sind.⁴

Jedoch ist hier eine Reform bezeichnend, die von folgenden traditionellen Aufgaben des Rechnungshofes wegführt:

- 1 — Kontrolle der Abschlüsse, die von staatlichen Stellen vorgelegt werden, um die Recht- und Ordnungsmäßigkeit der Erhebung der Einnahmen sowie der Ausgaben, Rechnungen und Belege nachzuweisen.
- 2 — Kontrolle der öffentlichen Mittelverwalter, um Verstöße gegen das Haushaltsrecht zu ahnden.⁵

und hinführt zu einer Konzentration auf Verbesserungen bei den geprüften Stellen anhand von

vielfältigeren und breiter angelegten [Aktivitäten] einschließlich Wirtschaftlichkeitsprüfungen. Letztere stellen sicher, dass die geprüften Stellen die ihnen zur Verfügung stehenden Mittel wirtschaftlich und gut verwalten. Außerdem kann bei Wirtschaftlichkeitsprüfungen auch eine Zielerreichungskontrolle vorgenommen werden.⁶

- Ein Bezug auf eine begrenzte Abstimmung mit den geprüften Stellen findet sich in mehreren Prüfungshandbüchern. Der vom australischen Rechnungshof veröffentlichte „Allgemeine Leitfaden für Wirtschaftlichkeitsprüfungen“ stellt z. B. fest, dass ein Vorgespräch zu Beginn der Prüfung

-
- dem Prüfungsteam Gelegenheit gibt, die Leitung der geprüften Stelle kennenzulernen; [und]
 - die Prüfungsart, Ziele, Umfang, Zeitplan und Methodik näher zu erläutern;
 - der geprüften Stelle ermöglicht, Stellungnahmen/Vorschläge gegenüber dem Rechnungshof abzugeben;
 - geeignete Vereinbarungen über den Kontakt sowohl auf Leitungs- als auch auf Arbeitsebene einschließlich Absprachen über die laufende Beratung und Besprechung der vorläufigen Ergebnisse zu treffen;
 - sicherzustellen, dass die geprüfte Stelle Klarheit über den Prüfungsablauf hat. . .⁷

Der Leitfaden fügt hinzu, dass „besonders wichtig [ist], dass die für die geprüften Programme verantwortlichen Führungskräfte . . . in allen Phasen der Prüfung in die Diskussionen einbezogen werden, um sicherzustellen, dass das Prüfungsteam mit den Rahmenbedingungen und den Problemen vertraut ist, die sich auf die geprüfte Funktion auswirken.“⁸

In der Fassung von 2002 des Handbuchs für Wirtschaftlichkeitsprüfungen des Rechnungshofes von Neuseeland wurde die Ausrichtung auf breiter angelegte Verfahrensweisen angeregt. Hier heißt es im Abschnitt mit der Überschrift „Da Aufgabenerfüllung und Rechenschaftslegung komplexer werden, muss sich die Berichterstattung durch die externe Finanzkontrolle entsprechend ändern“:

Eine Sichtweise der Rechenschaftslegung ist die, dass der Inhalt eines Berichts auf Themen beschränkt werden sollte, die der Zuständigkeit der Autoren unterliegen und für die diese direkt verantwortlich sind. Jedoch kann diese Auslegung der Rechenschaftslegung dazu führen, dass Stakeholdern wichtige Informationen vorenthalten werden.

...

Die beste Art der Berichterstattung der externen Finanzkontrolle stützt sich selektiv auf eine umfassende Kombination von Leistungsmerkmalen, um diejenigen festzustellen, zu messen und darzustellen, die für die Stakeholder interessant sind. Die Einhaltung der gesetzlichen Anforderungen an die Berichterstattung ist unerlässlich, eine erweiterte Berichterstattung aber gleichwohl erlaubt.

Eine modellhafte Berichterstattung setzt außerdem sowohl die Kenntnis der geprüften Maßnahmen als auch des Informationsbedarfs der Stakeholder voraus, um zu erkennen, welche Elemente der Bericht enthalten sollte und anhand welcher Leistungsindikatoren diese Elemente dargestellt werden. ...⁹

Die angegebene Begründung für eine erweiterte Wirtschaftlichkeitsprüfung war, „zur Verbesserung der Rechenschaftslegung der öffentlichen Stellen beizutragen“ und aufzuzeigen, dass „ein Rahmen für die Berichterstattung durch die externe Finanzkontrolle alle Leistungsmerkmale berücksichtigen und über diese berichten muss, die für die Stakeholder von Interesse sind.“¹⁰

Die Einbeziehung von Beteiligten in die Evaluierung kann z. B. in einem Bericht des französischen Rechnungshofes aus dem Jahr 2002 zum Thema Stadtplanung und -entwicklung nachgelesen werden, der eine Verbindung zwischen Wirtschaftlichkeitsprüfung und Programmevaluierung herstellt. Die Prüfer berieten sich zunächst mit den Programmbeteiligten — hohe Verwaltungsbeamte, gewählte Inhaber kommunaler Ämter und Mitglieder gemeinnütziger Organisationen — um die Evaluierung eines nationalen Stadterneuerungsprogramms zu planen. Sie besuchten dann sieben Standorte in vier Regionen des Landes, um sich mit gewählten Gemeinderäten, Beamten staatlicher Dienststellen und Mitgliedern von Vereinen zu treffen, die ebenfalls Programmbeteiligte waren.¹¹ So gehörten zu den Stakeholdern neben Beamten auch Vertreter von Nichtregierungsorganisationen sowie der Zivilgesellschaft insgesamt.

Eine Veröffentlichung aus dem Jahre 2003 stellt die Bedeutung Dritter bei der Planung von Evaluierungen fest. Unter dem Titel „Evaluierung: Eine Gewerkschaftsperspektive“ erläutert Bandula Kothalawala, dass „bei

gewerkschaftlichen Projekten die Beteiligung der Betroffenen an der Evaluierung bereits in der Konzeptphase . . . berücksichtigt wird und ein wesentliches Element bildet.”¹²

Nach der Feststellung, dass die Ziele von Gewerkschaftsevaluierungen sich vorwiegend auf die Personalentwicklung beziehen, bemerkt er:

Es wird angestrebt, so viele Stakeholder wie möglich in den [Evaluierungs]Prozess einzubeziehen, wobei den primären Nutznießern der Vorrang eingeräumt wird. Bei der Analyse der Daten wird auf eine kooperative Durchführung und auf die Sichtweisen der primären Nutznießer und der übrigen Stakeholder Wert gelegt. Es ist üblich, Workshops zu organisieren, an denen die Stakeholder teilnehmen und die in Zusammenarbeit mit lokalen Partnerorganisationen abgehalten werden. Alle Teilnehmer werden über Programmdetails unterrichtet und entsprechende Absprachen getroffen. Der Zielerreichungsgrad des Projekts wird besprochen und der Gesamt Ablauf so in qualitativer und subjektiver Hinsicht zusätzlich zu den durch andere Methoden erlangten Informationen ergänzt.

Im Zuge einer Evaluierung ist es außerdem üblich, informelle Interviews mit einer repräsentativen Auswahl von Hauptnutznießern durchzuführen.¹³

Diese im Laufe der Datenerhebung durchgeführten Interviews können u. a. den nachfolgend beschriebenen Nutzen haben:

Bei der Evaluierung der Schulung von medizinischem Personal in der Diagnose von Berufskrankheiten boten Interviews mit den betroffenen Mitarbeitern im Zuge örtlicher Erhebungen bei Kliniken und Krankenhäusern im Westteil der Kapprovinz Südafrikas eine gute Gelegenheit für den Prüfer, sich über die Lage vor Ort zu informieren.¹⁴

[8] EVALUIERUNG UND STAKEHOLDER:

Für die Wallonische Evaluierungsgesellschaft (*Société Wallonne de l'Evaluation et de la Prospective*) besteht Evaluierung aus der Analyse der Strategien, Programme, Projekte oder Funktionen in Bezug auf Relevanz, Wirksamkeit, Wirtschaftlichkeit, Stringenz, Nachhaltigkeit und Funktionsfähigkeit, durchgeführt mit der Absicht, die Qualität von Entscheidungen zu verbessern, Ressourcen sinnvoller zu verteilen und gegenüber den Bürgern besser Rechenschaft abzulegen. Evaluierung ist konzipiert als partizipativer und demokratischer Ansatz, der in den Evaluierungsprozess jede Person oder Gruppe einbezieht, die sich von der Evaluierung betroffen fühlt. *Société Wallonne de l'Evaluation et de la Prospective*, Documentation, Evaluation.

www.la-swep.be/index.php (27. April 2009).

Zuletzt zwei Beiträge von Prüfern, die beim Jahrestreffen 2000 der AEA in Honolulu/Hawaii über ihre Arbeit berichteten. In der Sitzung 111 zum Thema „Konsumenten als Forscher: Instrumente für die Beteiligung von Psychiatriepatienten bzw. Patienten, die wegen Drogensucht behandelt werden am Evaluierungsprozess“ wurde bemerkt: „Prüfer und Teilnehmer (Patienten) an einem Bundesforschungsprojekt zu Drogenmissbrauch, psychiatrischen Erkrankungen und Folgen traumatisierender Gewalt [führten] eine Diskussion über Strategien zur wirksamen Einbeziehung von Patienten (und Neulingen in diesem Forschungszweig) in das Forschungsteam.“¹⁵

Die Gründe für diesen Ansatz waren, dass „die Patienten eine einzigartige Perspektive haben, die bei der Steuerung des gesamten Evaluierungsprozesses helfen kann.“ Die Diskussion umfasste eine kurze Beschreibung eines kulturspezifischen Kreativitätsprojekts und der Elemente der Patientenbeteiligung an dem Projekt „einschließlich des Nutzens für die Forschung, Nutzen für die

Patienten, Wege bzw. Hindernisse hinsichtlich einer wirksamen Patientenbeteiligung sowie mögliche Lösungen zur Überwindung dieser Hindernisse.“¹⁶

Im Rahmen der Sitzung 115 und einer Präsentation am runden Tisch mit dem Titel “Klinische und politische Folgen bei der Evaluierung eines bikulturellen Programms“ beschrieb Gerald A. Bunn den Evaluierungszusammenhang:

Bei der Evaluierung von Alkohol- und Drogenprogrammen mit einer kulturellen Komponente sind sowohl die politischen als auch die klinischen Aspekte zu bedenken. Obwohl in Neuseeland jetzt viele verschiedene ethnische Gruppen leben, wird die wichtigste ethnische Trennlinie immer noch zwischen Maori (Neuseelands „tangata whenua“ oder „Leute des Landes“) und Pakeha (Nicht - Maori) gezogen. Der Maori-Status wurde im Vertrag von Waitangi definiert. Dieser wird quasi als Staatsgründungsurkunde von Neuseeland angesehen. Der Vertrag ist ein einzigartiges politisches Instrument, das darauf abzielt, die „rangatiratanga“ (Eigenständigkeit) zu schützen und zu stärken, die das Recht der Maori auf Leben und ihre eigene Lebensweise gewährleistet. Dies wird dahingehend ausgelegt, dass die Maori auch die Kontrolle über ihr eigenes Gesundheitswesen haben sollten. Im Allgemeinen erfolgt in Neuseeland die psychiatrische Versorgung getrennt für die Maori und die übrige Bevölkerung.¹⁷

Dementsprechend sei das zu bewertende stationäre Alkohol- und Drogenmissbrauchsprogramm ein „Parallelprogramm“, „demzufolge Maori in ihrem eigenen kulturellen Umfeld, dem Taha Maori-Programm“, und „Pakeha oder Nicht-Maori im Rahmen des allgemeinen Programms behandelt werden“. Somit würden die beiden Programme getrennt evaluiert. Jedoch werde

anerkannt, dass „auf Grund eurozentrischer Ansätze entwickelte Instrumente bei Farbigen nicht undifferenziert angewendet werden können.“ . . . [und dass] das Fehlen spezifischer Instrumente und Verfahren für Maori die Anpassung westlicher Instrumente und Verfahren erfordere. Um dies zu ermöglichen, seien umfassende und zeitaufwändige Gespräche mit einer Vielzahl von Stakeholdern erforderlich gewesen.¹⁸

[9] LERNEN AUS DER PROGRAMMEVALUIERUNG:

Die fünf Dienststellen, die wir geprüft haben, setzten verschiedene Strategien ein, um größeren Nutzen aus der Evaluierung von Programmen zu ziehen. Genauso wie zwischen den Programmen gab es auch Unterschiede bei Stil und Inhalt der Evaluierungen und der Art der von den einzelnen Dienststellen zu bewältigenden Aufgaben Unabhängig davon, ob Evaluierungstätigkeiten im Werdegang der einzelnen Dienststellen bereits eine große Rolle spielten oder als Reaktion auf neue externe Einflüsse begonnen wurden, ermöglichte das Lernen aus der Evaluierung kontinuierliche Verbesserungen bei der Aufgabenerfüllung und der Umsetzung von Programmen und einen Ausbau des vorhandenen Wissens

GAO, Programmevaluierung: Evaluierungskultur und Partnerschaften als Beitrag zum Aufbau von Sachkompetenzen, GAO-03-454 (Washington, D.C.: 2. Mai 2003), S. 24. www.gao.gov (27. April 2009).

Mit anderen Worten: Bei der Planung von Evaluierungen sind die Prüfer im Allgemeinen bestrebt, mit den Beteiligten und Dritten die Gründe der Evaluierung, ihre Ziele, die Merkmale des Programms, die vorläufigen Evaluierungsfragen, vermutete Probleme des Programms und mögliche Lösungen sowie Evaluierungsalternativen zu besprechen.

Laut GAO z. B.

kann es eine schwierige Aufgabe sein, die Zustimmung der Sponsoren, Nutznießer, Programmverantwortlichen und anderen Betroffenen zu Inhalten und Folgerungen einer Frage zu gewinnen. Einer der Gründe für die Schwierigkeit dieser Aufgabe besteht darin, dass die Formulierung eines Problems sehr große Bedeutung für die folgenden Phasen der Evaluierung hat.¹⁹

Kurz gesagt stehen Prüfungen der Rechnungsführung und der Wirtschaftlichkeit im Allgemeinen in einem Kontext von Rechenschaftspflicht und Ordnungsmäßigkeit, Programmevaluierungen dagegen in einem breiteren Zusammenhang, in dem aus komplexen Grundsatz- und Querschnittsfragen Erkenntnisse gewonnen werden können.

[3.2] Festlegung eines Evaluierungskonzepts

Bei der Festlegung eines Evaluierungskonzepts wissen erfahrene Prüfer, dass nicht unbedingt die von ihnen ausgewählte Methode die Aussagekraft der Evaluierung bestimmt:

Längsschnittevaluierungen sowie experimentelle, quasi-experimentelle, vorgängige und nachgängige sowie auf Fallstudien basierende Evaluierungen können eine unterschiedlich große Aussagekraft haben Ein einfaches vor- und nachgängiges Konzept ohne Kontrollen wird immer Schwächen in seiner Aussagekraft aufweisen. Jedoch ist es möglich, dass Stichprobenerhebungen und Kontrollgruppen aus einer Vielzahl von Gründen nicht in Frage kommen. Das heißt, die Aussagekraft einer Evaluierung muss im Zusammenhang von Fragestellung, gesetzter Fristen und Kostenzwänge, Konzept, fachlicher Angemessenheit der Datenerhebung und -analyse sowie anhand der Berichterstattung der Ergebnisse beurteilt werden.²⁰

Kapitel 3 („Planung einer Evaluierung“) der Veröffentlichung „Evaluierung von EU-Maßnahmen: Praktischer Leitfaden für die Kommissionsdienststellen“ der Europäischen Kommission aus dem Jahre 2004, stellt anschaulich dar, wie ORKB Programmevaluierungen planen können, und umfasst wichtige Punkte wie beispielsweise Festlegung der Ziele, des Zeitrahmens und der Projektsteuerung sowie Formulierung entsprechender Fragen und Vorgaben – d. h. die Erstellung eines Dokuments, das die zu bewertende Maßnahme beschreibt, die für die Prüfer verfügbaren Informationen benennt, eine Reihe von Fragen auflistet, die die

Evaluierung beantworten muss sowie die Phasen des Verfahrens und die anzuwendende Methode festlegt.²¹

In Bezug auf das Evaluierungskonzept erinnert der Leitfaden die ORKB daran, dass

viele unterschiedliche Evaluierungsmethoden möglich [sind] So kann es im Falle einer externen Evaluierung im Interesse des Auftraggebers liegen, nur eine grobe Inhaltsvorgabe zu machen und es dem Prüfer zu überlassen, je nach seiner Erfahrung und Fachkenntnis eine angemessene Lösung vorzuschlagen. . . .

Daneben spielt der Zeit- und Geldfaktor eine wichtige Rolle für das Konzept der Evaluierung. Höhere Budgets und größere Zeitfenster ermöglichen die Sammlung größerer und genauerer Daten sowie die Auftragsvergabe aufwändigerer Analysen. Die Möglichkeiten und Beschränkungen, die sich aus der Budgethöhe ergeben, können wie folgt zusammengefasst werden:

- Kurze Zeitspanne und niedriges Budget: Schreibtischstudie, Interviews, Fokusgruppen.
- Mittlere Zeitspanne und normales Budget: Fallstudien, Erhebungen, Expertengremien.
- Lange Zeitspanne und beträchtliches Budget: ökonomische Modelle, Kosten-Nutzen- und Kostenwirksamkeitsanalysen.²²

Vor der Durchführung einer Programmevaluierung sind Überlegungen über die zu sammelnden Informationen und deren Quellen, die Stichprobenmethoden, die Grundlage für die Anstellung von Vergleichen mit oder ohne Programm sowie den Analyseplan anzustellen. Die Entscheidungen, die bei Erstellung des Evaluierungskonzepts zu treffen sind, bestimmen die Qualität der Informationen sowie der Schlussfolgerungen, Kosten und Nutzen der Evaluierung sowie deren rechtzeitige Beendigung.

Kurz gesagt, wird das Evaluierungskonzept strategisch erstellt. Es wird so gestaltet, dass die Evaluierung Ergebnisse bestmöglicher Qualität hervorbringt und außerdem die Evaluierungskompetenz der Organisation weiter fördert.

[10] EVALUIERUNGSKONZEPT UND WIRKUNGEN DER MASSNAHMEN:

Schließlich haben [neben den zeitlichen und finanziellen Beschränkungen] andere Aspekte einer Evaluierung bedeutsame Auswirkungen auf die Wahl der Evaluierungsmethode, wie z. B.:

- Betrifft die Evaluierung ein Programm oder eine Maßnahme, die sich an klar definierte Nutznießer wendet, oder handelt es sich um eine Politik, die sich mehr oder weniger indirekt an ein breites und möglicherweise schwer zugängliches Publikum richtet?
- Soll sich die Evaluierung nur auf die direkten, kürzerfristigen Auswirkungen auf die Adressaten konzentrieren (z. B. im Falle einer Halbzeitevaluierung) oder sollen eher indirekte, längerfristige sowie unbeabsichtigte Wirkungen untersucht werden?

European Commission, Directorate-General for the Budget, Evaluating EU Activities: A Practical Guide for the Commission Services (Luxembourg: Office for Official Publications of the European Communities, Juli 2004), S. 45. European Commission, Financial Programming and Budgeting, Publications. http://ec.europa.eu/budget/index_en.htm (27. April 2009).

[3.3] Verknüpfung des Konzepts mit Evaluierungsfragen

Gleich ob die Evaluierungsziele auf politische Vorgaben oder Programmziele abgestimmt sind, empfiehlt es sich, sie als Fragen zu formulieren, deren Beantwortung alle Beteiligten im Einvernehmen mit den Prüfern für wünschenswert halten. Diese Fragestellungen sind im Allgemeinen beschreibend, normativ oder kausal und haben daher einen anderen Ansatz als Prüfungsfragen.

Die Fragen werden mit dem am besten geeigneten Konzept verknüpft. Dies geschieht durch immer stärkere Fokussierung der Evaluierungsfrage im Zuge der Auswahl von Konzeptalternativen. Ausgehend von der Evaluierungsfrage erfolgt eine Einteilung, ob diese

-
- beschreibend (z. B. Frage danach, wie das Schulungsprogramm umgesetzt wurde) oder
 - normativ (Frage, ob die Stellenvermittlungsziele des Programms erreicht wurden) oder
 - kausal ist (Frage, wie das Programm die Stellenvermittlungsquoten beeinflusst hat).

Das Konzept wird auch danach ausgerichtet, welche Art von Schlussfolgerungen die Stakeholder erwarten und mit welchen Sachzwängen sie bei dem Projekt rechnen. Möglicherweise klären die Prüfer beispielsweise, ob aufgrund von Feststellungen über einzelne Sachverhalte Aussagen für größere Populationen oder Gesamtheiten getroffen werden können. Dann kann die Evaluierungsfrage mit einer Reihe von Konzeptfragen verknüpft werden, wie z. B.:

- Kann eine Vergleichsgruppe gebildet werden?
- Kann nach dem Zufallsprinzip eine Zuordnung nach Gruppen vorgenommen werden?
- Können Ergebnisse im Zeitablauf gemessen werden?

Und schließlich, wie viel Zeit steht für die Durchführung der Evaluierung zur Verfügung?

Bei dieser Verfahrensweise wird ein Konzept Schritt für Schritt entwickelt. Bei der Erarbeitung des endgültigen Konzepts wägen Prüfer fachliche Anliegen gegen Sachzwänge ab. Gleichzeitig werden die Evaluierungsfragen spezifischer und die Untersuchungsmöglichkeiten enger:

So kann es z. B. notwendig sein, zwischen der Sammlung neuer Daten, die die Evaluierungsfragen umfassend beantworten, und der Nutzung verfügbarer Daten zu wählen, wobei letzteres gewöhnlich billiger und schneller ist, aber bestimmte Untersuchungsmöglichkeiten ausschließt.²³

Die mit Hilfe der Prüfungsfragen angestrebten Ziele können sich auf viele Bereiche beziehen: u. a. wie die Verwaltung die Transparenz verbessern kann, welcher

Bedarf nach einer Neuformulierung oder Änderung der politischen Vorgaben besteht, wie sich ein Programm im politischen Kontext darstellt oder welche Nettowirkung es hat. Einige besondere Fragen, die sich für die Prüfer bei der Erstellung eines Evaluierungskonzepts stellen könnten, sind u. a.

- Sind im Bereich der öffentlichen Hand vorgenommene Evaluierungen fachlich korrekt? Sind die Leistungsindikatoren aussagekräftig? Welche neuen Strategien könnten angewandt werden, um Politik zu bewerten?
- Welche wichtigen Aspekte des aktuellen Zusammenhangs wirken sich auf die Programmstrategie aus? Welche Ansichten der Stakeholder wirken sich auf die Strategien und Programme aus?
- Ist eine Folgenabschätzung erforderlich?
- Welche Erfordernisse bestehen künftig für ein neues Gesetz oder Vorschrift?
- Welche Auswirkungen hat das Programm über das hinaus, was passieren würde, wenn es nicht existierte? Wie ist das Verhältnis zwischen den Programmkosten und dem voraussichtlichen Nutzen?²⁴

Solche Fragen unterscheiden sich von jenen, die üblicherweise bei Prüfungen der Rechnungsführung oder der Wirtschaftlichkeit gestellt werden:

- Sind die Informationssysteme vollständig, genau, schlüssig und betrugsgeschützt?
- Werden bei den Einzelmaßnahmen von Programmen die Mittel wirksam eingesetzt?
- Erreicht das Programm seine festgelegten Ziele?
- Wurden bei der Umsetzung des Programms die gesetzlichen Vorschriften, Grundsätze und sonstigen Regelungen eingehalten?

[3.4] Beurteilung, ob ein Team zur Ausführung von Evaluierungsplänen geeignet ist

Das „Evaluierungshandbuch“ der W. K. Kellogg-Stiftung führt aus, dass

die vorrangigen Ziele einer Evaluierung darin bestehen, dass Stakeholder engagierte und aktive Teilnehmer sind . . . und dass das Evaluierungsverfahren sowie seine Ergebnisse aussagekräftig und nützlich für diejenigen sind, die letztlich für die Verbesserung und Beurteilung des Programms verantwortlich sind. Das bedeutet, dass es nicht nur einen Weg zur Durchführung einer Evaluierung gibt.²⁵

Weiterhin wird vorgeschlagen, dass je nach Beziehung zwischen primären Stakeholdern und dem Evaluierungsteam,

nach einem Prüfer mit guten methodischen Kenntnissen und großer Erfahrung gesucht werden sollte, falls der Evaluierungszweck darin besteht, den Wert und Nutzen eines Programms zu bestimmen. Falls die Evaluierung sich auf Programmverbesserungen konzentriert, wäre nach jemandem zu suchen, der das Programm gut kennt und überlegt vorgeht. Falls das Primärziel die Konzipierung neuer Programme aufgrund von Erfahrungswerten ist, müsste der Prüfer teamfähig sein und gute analytische Fähigkeiten haben.

Die Erfahrung zeigt jedoch, dass „die wichtigsten Eigenschaften eines Prüfers darin bestehen, flexibel zu sein und Problemstellungen lösen zu können“.²⁶

Der Erfolg von Programmevaluierungen kann durch Blick auf den Kontext der zu bewertenden Programme abgeschätzt werden. Durch Auswahl von Evaluierungskonzepten und -methoden wird festgelegt, welche Kenntnisse und Fähigkeiten die jeweiligen Prüfer haben müssen. Vor Beginn einer Evaluierung sind ORKB deshalb gut beraten, mit ihren Evaluierungsteams jedes Konzept und den Gesamtevaluierungsplan genau zu betrachten, um einen Überblick darüber zu erhalten, was die Untersuchung nicht nur für diejenigen mit sich bringt, die von den Ergebnissen profitieren, sondern auch für die evaluierende Institution selbst. Zu diesen Folgen und Auswirkungen gehören Fragestellungen wie:

[11] KENNTHNISSE, FERTIGKEITEN UND FÄHIGKEITEN DES PRÜFERS:

Prüfer sollten für die Evaluierung angemessene Sachkenntnisse besitzen oder sich aneignen sowie kontinuierlich bestrebt sein, ihre methodischen und praktischen Fertigkeiten zu verbessern. Prüfer sollten in folgenden Bereichen Wissen und Können bzw. Erfahrung besitzen:

- Anwendung solider Untersuchungskonzepte, die die ausgewählten Fragen beantworten können;
- Sammlung und Analyse zuverlässiger quantitativer und qualitativer Daten und
- Formulierung stichhaltiger, glaubwürdiger und unparteiischer Schlussfolgerungen und Empfehlungen.

Treasury Board of Canada Secretariat, Centre of Excellence for Evaluation, Evaluation Function in the Government of Canada (Ottawa: 6. Juli 2004), appendix 2, Evaluation Standards in the Government of Canada. Treasury Board of Canada Secretariat, Organization, Evaluation, Tools and Resources, Historical Documentation. www.tbs-sct.gc.ca/tbs-sct/index-eng.asp (27. April 2009).

-
- welche Konzeptbestandteile die Prüfer ausgewählt und auf welche sie verzichtet haben;
 - ob sie ein Konzept gewählt haben, das die Untersuchungsfragen angemessen beantworten wird;
 - ob der Evaluierungsplan Einschränkungen und Sachzwänge berücksichtigt;
 - wie die Prüfer die Stärken und Schwächen der Evaluierung begründen wollen;
 - Erwartungen der Stakeholder und wie der Prüfer ihnen entsprechen will;
 - wie die Erfahrungen aus der Evaluierung den Prüfern dabei helfen werden, neue Kenntnisse und Fähigkeiten für sich selbst, ihre Organisation und die Evaluierungsgemeinschaft zu erwerben.

Ressourcen und Ideen, um solchen Sachverhalten zu entsprechen, werden im Kapitel 5 dargestellt.

-
- 1 Siehe: Gerard Bukkems und Hans de Groot, Netherlands Court of Audit, "Learning by Auditing: Mission Impossible?" Paper for the 5th biennial conference of the European Evaluation Society, Learning, Theory and Evidence, Seville, 10.12. Oktober 2002, S. 5. www.evaluationcanada.ca/distribution/20021010_bukkems_gerard_de_groot_hans.pdf (27. April 2009).
 - 2 Saville Kushner, "The Object of One's Passion: Engagement and Community in Democratic Evaluation," Keynote address, Australasian Evaluation Society International Conference, Wollongong, Australia, Oktober–November 2002, S. 8. Australasian Evaluation Society, Conferences, Previous. www.aes.asn.au (27. April 2009).
 - 3 Vietnam, Office of the State Audit, Socialist Republic of Vietnam, National Assembly of the Socialist Republic of Vietnam, State Audit Law (25. Dez. 2001), Kap. 4, Abschnitt 5, Artikel 52. www.kiemtoannn.gov.vn/ktnnEnglish/thongtinM.htm (27. April 2009).

-
- 4 Vietnam, Office of the State Audit, Socialist Republic of Vietnam, National Assembly of the Socialist Republic of Vietnam, State Audit Law (25. Dez. 2001), Kap. 1, Artikel 9.
 - 5 Lebanese Republic, Court of Audit, About COA, Goals and Mission, Jurisdiction and Procedures of the Audit Court. www.coa.gov.lb (27. April 2009).
 - 6 Saudi Arabia, General Auditing Bureau, The Kingdom of Saudi Arabia's Message to the 7th Global Forum on Building Trust in Government for Further Participation and Transparency, "The Saudi's Message in Brief," S. 12. www.gab.gov.sa/index.php (27. April 2009).
 - 7 Australia National Audit Office, *General Guidance on the Conduct of Performance Audits* (Canberra: Juli 2003), Abschnitt 25, S. 4. Australia National Audit Office, Publications, Other Publications. www.anao.gov.au/ (27. April 2009).
 - 8 Australia National Audit Office, *General Guidance on the Conduct of Performance Audits*, Abschnitt 25, S. 5.
 - 9 New Zealand, Report of the Controller and Auditor- General, *Reporting Public Sector Performance*, 2. Ausg. (Wellington: Jan. 2002), S. 22–23. New Zealand Office of the Controller and Auditor-General, Good Practice Guides. www.oag.govt.nz (27. April 2009).
 - 10 New Zealand, Report of the Controller and Auditor-General, *Reporting Public Sector Performance*, S. 5.
 - 11 France, Cour des Comptes, *La Politique de la Ville* (Paris: Feb. 2002), siehe S. 12–14 und 206–07. France, Cour des Comptes, Les juridictions financiers, Publications, Rapports publics thématiques. www.ccomptes.fr/ (27. April 2009).
 - 12 Bandula Kothalawala, *Evaluation: A Trade Union Perspective* (London, Eng.: Trades Union Congress, European Union and International Relations Department, n.d.), S. 4. www.minefi.gouv.fr/fonds_documentaire/TRESOR/cid/atelier/contrib/27.pdf. (27. April 2009).
 - 13 Kothalawala, *Evaluation*, S. 5.
 - 14 Kothalawala, *Evaluation*, S. 6.
 - 15 Siehe : Sitzung 111 (Mi. 16.30)mers into the Evaluation Process, Chaired by Debi Elliott, Portland State University. American Evaluation Association, Training,

-
- Conference History, Evaluation 2000, 1.-4. November, Honolulu. www.eval.org (27. April 2009).
- 16 Session 111, Consumers as Researchers, Chaired by Debi Elliott. American Evaluation Association, Training, Conference History, Evaluation 2000.
- 17 Siehe: Sitzung 115 (Mi 16.30), Examining Treatment and Prevention Programs, Robert G. Orwin, host, Battelle Centers for Public Health Research Considering Culture in Evaluating Health Programs, Gerald A. Bunn, "Clinical and Political Implications in Evaluating a Bicultural Program." American Evaluation Association, Training, Conference History, Evaluation 2000, 1.-4. November, Honolulu. www.eval.org (27. April 2009).
- 18 Sitzung 115, Examining Treatment and Prevention Programs, Gerald A. Bunn, "Clinical and Political Implications in Evaluating a Bicultural Program." American Evaluation Association, Training, Conference History, Evaluation 2000.
- 19 GAO, *Designing Evaluations*, PEMD-10.1.4 (Washington, D.C.: Mai 1991), S. 11. www.gao.gov (27. April 2009).
- 20 GAO, *Designing Evaluations*, S. 16.
- 21 Europäische Kommission, Directorate-General for the Budget, *Evaluating EU Activities: A Practical Guide for the Commission Services* (Luxembourg: Office for Official Publications of the European Communities, Juli 2004), Kap. 3, S. 37-43. European Commission, Financial Programming and Budgeting, Publications. http://ec.europa.eu/budget/index_en.htm (27. April 2009).
- 22 European Commission, *Evaluating EU Activities*, "Designing an Evaluation," Kap. 3, S. 45.
- 23 GAO, *Designing Evaluations*, S. 66.
- 24 Diese und weitere Fragen in den folgenden Absätzen werden behandelt in Anlehnung an: Valerie J. Caracelli and Judith A. Droitcour, "Evaluation and Audit Cross-discipline Models: A Multiplist Approach to Strengthening Information Quality," Anhang I, Tabelle A1, S. 48, vorgelegt bei der American Evaluation Association, Evaluation 2000, Conference History, Evaluation 2000, Session 277, Honolulu, Donnerstag 2. November 2000. www.eval.org (27. April 2009).
- 25 W. K. Kellogg Foundation, *W. K. Kellogg Foundation Evaluation Handbook* (Battle Creek, Mich.: 1. Jan. 1998),

S. 59. www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf (27. April 2009).

26 W. K. Kellogg Foundation, *W. K. Kellogg Foundation Evaluation Handbook*, S. 59–60.

*K*apitel 4: *D*urchführung *d*er *P*rogrammevaluierung

Die erfolgreiche DURCHFÜHRUNG einer Programmevaluierung durch ein EVALUIERUNGSTEAM bedeutet, dass es die verschiedenen sozialen und kulturellen Perspektiven der Bedürfnisse, Werte und Vorstellungen unterschiedlicher Gruppen in das Konzept einbezogen hat. Das Konzept ist flexibel genug, um eine Neuorientierung zu ermöglichen, falls sich die Projektrahmenbedingungen ändern. Auch wurden Konzept und Evaluierungsfragen verknüpft, so dass die Prüfer eher auf Ziele als auf Methoden konzentriert bleiben. Sorgfältig entworfene Programmevaluierungen sind auch geeignet, Tendenzen der Untersuchung aufzuzeigen, Daten nur zu relevanten Informationen zu erfassen und zu stichhaltigen und zuverlässigen Feststellungen und Würdigungen zu kommen.

In diesem Kapitel werden die Prüfer auf einige der ihnen zur Verfügung stehenden Methoden hingewiesen. Weiter wird ausgesagt, dass die zunehmende Erfahrung der Prüfer mit der Durchführung von Evaluierungen und der Berichterstattung über sie für den Ausbau der Evaluierungskompetenz der Organisation sowie der der größeren Evaluierungsgemeinschaft ausschlaggebend ist.

[4.1] Einsatz der Evaluierungsmethodik

Bei einer Evaluierung kann es sich um eine nachgängige oder eine vorgängige handeln. Die Durchführung von Programmevaluierungen entweder mit einem Blick zurück auf das Vergangene oder mit einem Blick nach vorne auf die künftige

Entwicklung muss nicht, kann aber, komplex sein. Tabelle 1 zeigt einige Aspekte dieser zwei Methoden.

Bei der Wahl der Datenquellen und Analysemethoden der Evaluierung sollte selbstverständlich die Glaubwürdigkeit und Zuverlässigkeit der Ergebnisse angestrebt werden. Dies gibt größere Sicherheit hinsichtlich der Evaluierungsqualität.

Tabelle 1. Programmevaluierungsmethodik: Nachgängig oder Vorgängig

Methoden-merkmal	Nachgängige Evaluierung	Vorgängige Evaluierung
Quellen für Fragen	Für Maßnahmen, Vorschriften, Kriterien, Sachverhalte	Strategien, oder relevante oder Ideen und Annahmen über Probleme, Gründe und mögliche Lösungen
Informationsquellen	Interviews, Beobachtungen, Meinungsumfragen, Dokumente, Verwaltungsdaten	Vorgängige Forschung, Theoriebildung und Evaluierung; Pilotversuche mit dem vorgesehenen Ansatz; Gutachten
Analysearten	Qualitative oder quantitative Ansätze zum Umgang mit empirischen Daten; Informationssynthese gemäß Programmkriterien und -fragen	Simulationen, Modelle und Informationssynthesen zu konzeptuellen und operativen Annahmen der Vorschläge; Delphi-gemäß Analyse Techniken; Voraussichtlicher Wirkungen

Quelle: Aus einer Publikation des GAO entnommen und bearbeitet mit dem Titel, Prospective Evaluation Methods: The Prospective Evaluation Synthesis, [PEMD-10.1.10](#) (Washington, D.C. Nov. 1990), Tabelle 1.2, S. 16. www.gao.gov.

Zu den für die Datensammlung und -analyse im Zuge einer Evaluierung verfügbaren Methoden und Instrumenten zeigt die Tabelle 1 einige nachgängige Ansätze — z. B. Durchführung von Interviews und Erhebungen oder Prüfung von Unterlagen oder Verwaltungsdaten oder Durchführung einer Informationssynthese — und einige vorgängige Ansätze — etwaige Durchführung von Pilotverfahren oder experimentellen Tests zu einer ins Auge gefassten Vorgehensweise oder von Simulationen bzw. die Einrichtung von Modellen zur Überprüfung konzeptioneller oder operativer Annahmen. Viele andere Methoden sind ebenfalls verfügbar.

Während Wirtschaftlichkeitsprüfungen herkömmlich Informationen darüber liefern, wie gut die Dinge in einem Programm laufen und den Blick darauf richten, was sogleich gemessen und berichtet werden kann, untersuchen Programmevaluierungen eher die Programmleistung und warum sie so ist, wie sie sich darstellt und eventuell ihre mögliche künftige Entwicklung. Evaluierungen können Wirtschaftlichkeitsprüfungen durch Untersuchung der Beziehungen zwischen Aktivitäten, Kontext und Ergebnissen eines Programms ergänzen, das Verständnis der Gründe für die Ergebnisse fördern und den Beitrag des Programms zu den Ergebnissen von anderen Faktoren unterscheiden.

Prüfer richten ihre Fragen und die Arten der von ihnen durchgeführten Evaluierungen nach der Programmphase, die sie zu untersuchen beabsichtigen. Tabelle 2 listet die Fragen auf, die Prüfer im Allgemeinen je nach Programmfortschritt stellen und zeigt die Art der Evaluierung, die sie dann wahrscheinlich durchführen.

Tabelle 2. Fragen- und Evaluierungsarten je nach Programmstadium		
Programmstadium	Übliche Evaluierungsfragen	Evaluierungsart
Konzept	Ist das Programmkonzept gut ausgestaltet, machbar und wird es die geplanten Ziele wahrscheinlich erreichen?	Konzeptbewertung
Frühes Stadium oder neue Initiative innerhalb des Programms	<ul style="list-style-type: none"> • Erreicht das Programm die vorgesehene Zielgruppe? • Wird das Programm gut verwaltet? • Welcher Fortschritt wurde bei der Einführung neuer Bestimmungen erzielt? 	Prozess- oder Vollzugs-evaluierung
Reif und stabil mit einem scharf umrissenen Programmmodell	<ul style="list-style-type: none"> • Werden die gewünschten Programmergebnisse erzielt? • Verursachte das Programm unbeabsichtigte Nebenwirkungen? • Unterscheiden sich die Ergebnisse quer durch Programmkomponenten, Lieferanten oder Kundenuntergruppen? 	Ergebnis-evaluierung
	Warum erzielt das Programm nicht mehr die gewünschten Ergebnisse?	Prozess-evaluierung
	Erzielte das Programm die gewünschte Wirkung?	Wirkungs-evaluierung
Quelle: Eric Bernholz und andere, "Evaluation Dialogue between OMB Staff and Federal Evaluators: Digging a Bit Deeper into Evaluation Science," PowerPoint presentation [Washington, D.C.], Juli 2006, Tabelle 1, S. 14, und GAO, <i>Performance Measurement and Evaluation: Definitions and Relationships</i> , GAO-05-739SP (Washington, D.C.: Mai 2005), www.gao.gov .		

Der Zweck dieses Kapitels besteht jedoch nicht darin, alle möglichen Methoden aufzulisten und zu definieren. Es soll ORKB Instrumente für viele unterschiedliche

Situationen anbieten. Die Evaluierungsgemeinschaft stellt eine Vielzahl von unterschiedlich genannten Checklisten, „Kochbüchern“, Matrizes oder ähnliches bereit, und zwar je nach Art der Evaluierungsfrage für die Auswahl geeigneter Datenquellen und Analysemethoden.

In den nachstehenden Abschnitten werden einige Methodenquellen angeführt. Der erste Absatz enthält direkte Quellen aus Leitfäden, herausgegeben von der Evaluierungsgemeinschaft und intern von zu dieser Gemeinschaft gehörenden ORKB. Der zweite Absatz enthält indirekte Quellen, die Methodikdiskussionen in Evaluierungsberichten entstammen. ORKB sollten auf Erfahrung mit einer Vielzahl von Methoden abzielen, da zwei identische Evaluierungen nicht existieren und ein Evaluierungskonzept sich in dem Maße weiterentwickeln kann, wie die Prüfer Informationen über Sachverhalte und Daten einer Evaluierung ansammeln.

[4.1.a] Einige direkt verfügbare Methodenquellen

Bei Erstellung der vorliegenden Abhandlung wurde auf *W. K. Kellogg Evaluation Handbook* zurückgegriffen. Teil 2 des Stiftungshandbuchs „Blaupause für die Durchführung von Evaluierungen auf Projektebene“ (*Blueprint for Conducting Project - Level Evaluations*) führt evaluierende Organisationen durch die Umsetzungsschritte Nr. 5 „Bestimmung von Datenerfassungsmethoden“, Nr. 6 „Datenerfassung“ und Nr. 7 „Datenanalyse und –auslegung“.¹ Das Handbuch vergleicht quantitative und qualitative Analysemethoden, beschreibt Techniken der Beobachtung und der Interviewdurchführung, illustriert die Nützlichkeit von Tests und Begutachtungen und bietet verschiedene Verfahren für die Durchsicht von Unterlagen an. Weiterhin unterweist es Prüfer im Umgang mit diesen Methoden, zitiert Beispiele, wie andere sie angewendet haben und listet Gedächtnishilfen zu ihrer Umsetzung auf. Schritt Nr. 7 beschreibt auch untraditionelle Analyseformen, die das Flexibilitätsziel im Evaluierungskonzept aufzeigen.

Das ins Internet gestellte und regelmäßig aktualisierte „Evaluierungsprüflistenprojekt“ des Evaluierungszentrums der Western Michigan University offeriert laufend rund 32 Dateien, die Prüfern Modelle von

Programmevaluierungen und Metaevaluierungen zur Verfügung stellen, Evaluierungswerte und -kriterien diskutieren, Prüfern bei der Evaluierungssteuerung helfen und sie beim Aufbau von Evaluierungskapazität unterstützen. So helfen z. B. mehrere Prüflisten unter „Metaevaluierung“ bei der Evaluierung der Leistung von Metaevaluierungen durch Scoring der quantitativen und qualitativen Informationsanalyse einer Evaluierung.²

Die USAID, die mit ihrem EvalWeb im Begriff ist, „ihre zentrale Evaluierungsfunktion zu reaktivieren“, entwirft „Funktionelle Angaben und Stellenbeschreibungen für die neue Evaluierungsabteilung“ und plant „einen sichtbaren öffentlichen ‘Start’ der neuen Evaluierungseinheit Anfang 2009“ zusammen mit dem Anlegen „einer Anfangsagenda von Untersuchungen für die zentrale Evaluierungsfunktion.“³ EvalWeb stellt seit langem mehrere Evaluierungsressourcen zur Verfügung, unter denen sich Unterlagen befinden, die verschiedene Evaluierungsmethoden wie partizipative Evaluierungen, Interviews mit Schlüsselinformanten und Fokusgruppen, Direktbeobachtungstechniken, schnelle Auswertungen und Folgeabschätzungen beschreiben.

Sie sind auf den Seiten der Entwicklungshilfe-Datenbank der USAID gesammelt und schließen z. B. die in 2006 erschienene Abhandlung „Übergang von Ausgaben zu Einkommen: Praktische Ratschläge aus Ländern in aller Welt“ (*Moving from Outputs to Incomes: Practical Advice from Countries around the World*) ein. Aufgrund einer zweitägigen Diskussion der Angehörigen von Rechnungshöfen, Finanzministerien, Schatzämtern und anderen nationalen öffentlichen Dienststellen aus 12 Staaten in Afrika, Europa und dem Vereinigten Königreich, Lateinamerika, dem Mittleren Osten und den USA arbeitet dieser Fachtext zur Evaluierung heraus

welche kleinen und größeren, anderswo bereits erfolgreichen Maßnahmen ergriffen werden müssen, um eine Staatsverwaltung auf Ergebnisorientierung neu auszurichten. Dies umfasst bewährte Verfahrensweisen, Schulung und praktische Schritte, die anderen Regierungen bei der Weiterentwicklung eines ergebnisorientierten Ansatzes für die öffentliche Verwaltung nützlich sein können.⁴

So beschreibt z. B. der Beitrag „Übergang von Ausgaben zu Einkommen: Praktische Ratschläge aus Ländern in aller Welt“ unter der Überschrift

„Evaluierung“ im Kapitel mit dem Titel „Die Rolle der Beobachtung und Evaluierung bei Ermöglichung eines ergebnisorientierten Ansatzes“ mehrere Zwecke, die Länder für die Evaluierung fanden. Unter ihnen sind zu nennen:

- Einschätzung der Eignung von Maßnahmen, Vorhaben und Strategien für Gegenwart und Zukunft.
- Feststellung, ob die für die Überwachung verwendeten Ziele, Indikatoren und Vorgaben weiterhin geeignet sind.
- Gewinnung von Informationen zum Warum und Wie, die gebraucht werden, um zu verstehen, unter welchen Umständen ein Programmansatz „funktioniert“ oder nicht.
- Erklärung der Faktoren, die hinter positiven Auswirkungen stehen.
- Identifizierung ungewollter oder unerwarteter Folgen.
- Demonstration von Kausalität oder Zuordnung — Verknüpfung zwischen Inputs, Aktivitäten und Ergebnissen.⁵

Bereits wesentlich früher legte die USAID-Abhandlung aus dem Jahre 1987 mit dem Titel „AID-Evaluierungshandbuch: Bericht Nr. 7 über das AID-Programmkonzept und Evaluierungsmethoden“ (*AID Evaluation Handbook: AID Program Design and Evaluation Methodology Report 7*) die Evaluierungspraxis der Agentur dar. Unter den „Informationsquellen und Daten für AID-Evaluierungen“ in „Evaluierungsplanung: Was ist zu evaluieren“ im Kapitel 3 werden Methoden aufgelistet für das Sammeln „sekundärer und kontextbezogener Daten“ wie

- Entwicklungspläne und –strategien des Gastlandes
- Projektunterlagen und Berichte des Gastlandes
- Berichte von Organisationen des Privatsektors
- Bücher, Periodika und Zeitschriften
- Forschungsstudien
- Sonstige bilaterale/multilaterale Geberprojekt- und Programmdokumentation
- Informantenberichte/Gutachten
- Weltbankberichte/Studien

[12] RISIKOANALYSE—KOSTEN UND NUTZEN DER EVALUIERUNG:

[D]a die Ressourcen für die Evaluierung in Neuseeland sehr begrenzt sind, sind sorgfältige Entscheidungen darüber zu treffen, wann Evaluierungen vorzunehmen sind Deshalb haben wir untersucht, wie gut die Beamten darüber entscheiden, was zu evaluieren ist, wann Evaluierungsaufträge erteilt werden und welche Evaluierungsverfahren im Einzelfall angewandt werden sollen.

Ein zentrales Thema des Berichts mit dem Titel „Review of the Centre“, der sich mit den Mechanismen der Staatsverwaltung insgesamt befasst, war die Notwendigkeit, für die Staatsverwaltung neue Instrumente zu entwickeln, die die Erzielung von Ergebnissen für die Bürger in den Mittelpunkt stellen. Folglich konzentriert sich unser Interesse an Evaluierungen darauf, wie mit ihrer Hilfe die Entscheidungsfindung verbessert werden kann, so dass die staatlichen Interventionen zu besseren Ergebnissen für die Bürger führen. Wir haben deshalb auch geprüft, wie wirksam die Verwaltung die Ergebnisse von Evaluierungen für die Ausrichtung der Politik und die Konzeption sowie Durchführung von Maßnahmen nutzt.

Shenagh Gleisner, Arati Mascarenhas und Sonia Wansbrough, “Improving Evaluative Activity in the New Zealand State Sector,”

Vortrag bei der Australasian Evaluation Society Conference, 16.–18. September 2003, Auckland, S. 3. Australasian Evaluation

Society, 2003 Conference, Paper WE31, “Government Cannot Make This Decision without a Government Study—Yeah, right.” www.aes.asn.au (29. April 2009).

und das Sammeln primärer Daten wie z. B.

Beobachtung: durch Teilnehmer oder Nicht-Teilnehmer . . . entwickelt als Teil regelmäßiger Standortbesuche durch das Projektteam

Erhebung: Durch Interviews oder Einsatz strukturierter Fragebögen

Sonstiges:

- Vor- und nachgängige Fallstudien
- Dokumentation durch das Projektteam in Form von Protokollen usw.
- Gruppensitzungen, um Diskussionen über Projekterfahrungen und -erkenntnisse anzuregen.⁶

Schon erwähnt wurde die Arbeitshilfe „Evaluieren von EU-Aktivitäten: Praktischer Leitfaden für die Kommissionsdienststellen“ der Europäischen Kommission aus dem Jahre 2004, in dem der Anhang E wohlweislich „eine Reihe von häufig in Evaluierungsmethoden vorgefundene Instrumente und Techniken“ beschreibt.⁷ Die Kommission regt z. B. an, dass Evaluierungsgesellschaften bei der Ausarbeitung von Konzepten eine Interventionslogikanalyse oder SWOT-Analyse der Stärken und Schwächen des zu evaluierenden Unternehmens und seine zukünftigen Chancen und Risiken in Betracht ziehen könnten.

Zu den nachgängigen und vorgängigen Methoden, die der Leitfaden aufzählt, gehören Datensammlung, Interviews und Fokusgruppen. Anhang E führt ebenfalls gezielt verschiedene Methoden der qualitativen und quantitativen Datenanalyse an und schließt dann mit einer Diskussion der Kosten-Nutzen- und Multikriterienanalyse ab, die „in der Regel im Rahmen von vorgängigen Analysen verwendet wird und eingesetzt werden kann, um Beurteilungen in komplexen Situationen zu stützen.“⁸ Außerdem differenziert das Kapitel 1 in „Evaluierung in der Kommission, Definition und Umfang“ unter dem Titel „Zeitliche Varianten“⁹ zwischen nachgängigen und vorgängigen Methoden. Schließlich stellt der Leitfadenanhang F kurz einige der methodischen und analytischen Folgerungen der Evaluierungsstrategie im Unterschied zu Evaluierungsagenturen oder -programmen vor.¹⁰

Dies sind nur einige der Leitfäden, Handbücher und Arbeitshilfen, die weltweit elektronisch oder auf Papier zur Verfügung stehen. Weitere Überlegungen zur

Evaluierungsmethodik und ihre Beziehung zur Wirtschaftlichkeitsprüfung können in einer umfangreichen Fachliteratur nachgelesen werden, welche die gesamte ORKB-Prüfungsstruktur und -praxis behandelt und analysiert¹¹

[4.1.b] Einige indirekt verfügbare Methodenquellen

Evaluierungsmethoden lassen sich auch aus vielen Evaluierungsberichten erschließen. Die Beschreibung einer Evaluierungsmethodik, in der Zielsetzungen, Konzept und Umfang diskutiert werden, kann im Hauptbericht zusammen mit den wesentlichen Feststellungen und Schlussfolgerungen oder in einem Anhang zum eigentlichen Bericht untergebracht werden. Hier nun einige Beispiele für die Beschreibung von Methoden in veröffentlichten Evaluierungsberichten oder in von ORKB erstellten Zusammenfassungen ihrer Evaluierung.

Im Oktober 2004 fasste der neuseeländische Rechnungshof nach einer Evaluierung der Forschungs- und Evaluierungsfunktion in vier Strafjustizdienststellen die Hauptfeststellung der Evaluierung in einem kurzen Papier zusammen, betitelt „Erfolgsfaktoren für wirksame Koordinierung und Zusammenarbeit zwischen öffentlichen Dienststellen“.¹² Diese Veröffentlichung erfolgte gesondert von dem Hauptbericht „Abstimmung und Zusammenarbeit im Strafjustizsektor“, der das Grundprinzip der Evaluierung einschließlich ihrer Ziele und Methodik bespricht.¹³ Infolge der von der Regierung erkannten Notwendigkeit der Zusammenarbeit zwischen Behörden mit gemeinsamen Zielen hatte sich der Zweck der Studie erweitert. Der Strafjustizsektor — ein komplexes Netzwerk eigenständiger, aber verfahrensrechtlich verbundener Dienststellen — wurde für Evaluierung ausgesucht, da er „relativ gut definiert und in sich geschlossen ist“, „eine große Menge verfügbarer Informationen besitzt“, „funktionelle Beziehungen zwischen den Sektordienststellen im Allgemeinen gut verstanden [wurden]“ und „die Dienststellen im Strafjustizsektor es häufig mit den gleichen Klienten zu tun hatten“.¹⁴

Im Berichtsabschnitt mit dem Titel „Wie führen wir die Prüfung durch?“ wird die Methodik als eine Reihe von Aktivitäten beschrieben, die durchgeführt bzw. unterlassen wurden. So prüfte das neuseeländische Evaluierungsteam z. B. Strategien und Abläufe für die Abstimmung zwischen den Strafjustizdienststellen auf vier Hauptgebieten, interviewte Beamte an wichtigen Schaltstellen und prüfte

zweckdienliche Unterlagen einschließlich des Werdegangs von Gesetzesvorhaben und „verwaltungsinternen Grundsatzregeln zur strategischen Planung, zur Weiterentwicklung der Grundsätze sowie Informationssysteme und Berücksichtigung der Belange der Maori.“ Das Team prüfte jedoch nicht „wie Strafjustizdienststellen ihre Anstrengungen koordinieren und sich bei der Wahrnehmung ihrer laufenden Aufgaben bei der Betreuung einzelner Klienten und Straftätern miteinander ins Benehmen setzen.“¹⁵

Die in „Koordinierung und Zusammenarbeit in der Strafjustiz“ niedergelegten Ergebnisse enthielten „viele Beispiele für Good Practice im Bereich Strafjustiz und für ein starkes Engagement beim Austausch von Informationen und bei der Zusammenarbeit“, aber gleichzeitig „wurden die Auswirkungen der Pläne oder Aktivitäten einer Dienststelle auf andere . . . nicht immer gut verstanden, was zu Risiken für die Umsetzung politischer Vorhaben führte, die für die Erreichung der strategischen Ziele der Regierung wesentlich waren.“¹⁶ Die Studie ergab u. a., dass „neue Verfahren für Projektaufsicht und Risikomanagement eingeführt“ und eine Empfehlung abgegeben wurde, wonach die Strafjustizbehörden ein Modell übernehmen sollten, das die sozialpolitischen Dienststellen eingeführt hatten, seitdem „der Ausschuss für sozialpolitische Evaluierung und Forschung mit dem Auftrag eingesetzt worden war, die staatlichen Investitionen in sozialpolitische Forschung und Evaluierung zu überwachen.“¹⁷ Man könnte deshalb die in Neuseeland geleistete Arbeit als vorgängige ergebnisorientierte Evaluierung bezeichnen.

[13] Wichtige Bestandteile eines Musterbeispiels für hohe Wirtschaftlichkeit:

Welcher Kontext ist gegeben?

- Gesamtkonfiguration des Programms (Beschreibung, Ziele, Ressourcen)
- Ergebniskette (Programmtheorie)
- bestehende Risiken

Was sollte zu welchen Kosten erreicht werden?

- (klare und konkrete) Beschreibung zu erwarteten Outputs und Outcomes
- geplante Ausgaben

Soll-Ist-Vergleich (inwieweit wurden die Erwartungen erfüllt?)

- relevante Outputs zu welchen Kosten
- realisierte Outcomes im Vergleich zu den Erwartungen
- Diskussion der verfügbaren Nachweise, die den vom Programm erbrachten Beitrag zu besagten Outcomes belegen

Was wurde gelernt und was ist als nächstes zu tun?

- Diskussion darüber, was aufgrund der gemachten Erfahrungen künftig anders gehandhabt wird

Was wurde getan, um die Datenqualität sicherzustellen?

- Beschreibung dessen, was die Organisation tut, um die Qualität der gelieferten Daten und Informationen sicherzustellen

Wesentlicher Bestandteil eines Musterbeispiels für hohe Wirtschaftlichkeit ist der Soll-Ist-Vergleich und die Darstellung der künftigen Verbesserung des Verwaltungshandelns.

John Mayne, Office of the Auditor General of Canada, Reporting on Outcomes: Setting Performance Expectations and Telling Performance Stories (Ottawa: 14. Mai 2003). Office of the Auditor General of Canada, Reports and Publications, Audit Methodology, Discussion Papers. www.oag-bvg.gc.ca (2. Mai 2009)

Mazedonien erstellte im Juli 2008 ein Papier über eine Untersuchung, die man Prozessevaluierung nennen könnte. In einer Präsentation bei einem EUROSAI-Seminar im November 2008 in Prag erklärten Vertreter des mazedonischen Rechnungshofes, dass die Untersuchung mit dem Titel „Bereitstellung angemessener Sozialbetreuung“ auf die Beantwortung folgender Fragen abzielte:

- ob ein angemessener Rechtsrahmen für den Schutz der Familie besteht;
- ob die Sozialzentren ausreichend mit personellen und sonstigen erforderlichen Mitteln ausgestattet sind, um die gesetzlichen Vorschriften umzusetzen;
- ob die verschiedenen Sozialbetreuungsstellen ausreichend miteinander Kontakt halten und zusammenarbeiten, um ihre Aufgaben wirksam erfüllen und negative Vorfälle reduzieren zu können.¹⁸

Die Präsentation umriss die Planung, Datensammlung und Analyse sowie die Berichterstattungsmethoden der Untersuchung zusätzlich zur Beschreibung ihrer Ergebnisse und Empfehlungen. Aufgrund von Besuchen bei acht Sozialbetreuungscentren in Skopje und anderen Städten der Republik Mazedonien und der Analyse anderer Primär- und Sekundärdaten bewertete der Bericht, ob das Familienbetreuungsprogramm des Landes eine solide Rechtsgrundlage besaß, das Personal hinlänglich geschult und ausgestattet war und die zuständigen Schulen, psychiatrischen Einrichtungen, Polizeidienststellen und Justizbehörden ihre Tätigkeit ausreichend miteinander abstimmten, um die Bereitstellung angemessener Betreuung für die vom Programm betroffenen Familien sicherzustellen.¹⁹

Die Evaluierungsergebnisse führten zu mehreren Empfehlungen: Einführung von Normen und operativen Standards für eine angemessene Ausstattung der Institutionen des Sozialbetreuungssystems mit Personal, Räumen und Sachmitteln; Erstellung eines strategischen Budgetplans für den Betrieb der Sozialbetreuungscentren und deren Aufsichtsbehörden; Verpflichtung der Dienststellen zur Erarbeitung und Umsetzung methodischer Richtlinien; Verschärfung der Gesetze und Vorschriften, die Opfer häuslicher Gewalt schützen und unterstützen und Verbesserung „der bestehenden bewährten Formen ambulanter

Betreuung . . . [und] Einführung neuer Formen, mit denen die entwickelten europäischen Länder gute Erfahrungen gemacht haben.“²⁰

Zusätzliche Empfehlungen betreffen die Einführung von zwischenbehördlichen Verfahrensregelungen zur wirksameren Kooperation zwischen Schulen, Sozialbetreuungszentren, Polizei, Gesundheitseinrichtungen und Gerichten — sowie Sozialarbeitern, Psychologen, Lehrern und anderen mit entsprechendem Fachwissen — bei der Aufdeckung und Prävention nachteiliger Sozialphänomene. Besondere Empfehlungen wurden hinsichtlich Schulung sowie Einrichtung von Familien-Notfalldiensten, Beratungsdiensten und anderen Interventions- und Unterstützungsinstrumenten gemacht.²¹ In Anbetracht der Tatsache, dass die Berichtsergebnisse auf größeren Erfolg des Programms ausgerichtet waren, zum Teil durch Übernahme einschlägiger Erfahrung anderer, könnte die Evaluierung als vorgängig angesehen werden.

Im Januar 2003 gab die irische staatliche Stelle für die Modernisierung des öffentlichen Dienstes nach Durchführung einer Evaluierung einen 138 Seiten langen Abschlussbericht unter dem Titel „NGM-Bericht über die Kommunikation von Wandel und Modernisierung des Beamtentums: Eine Befragung irischer Beamter“ heraus. Die Evaluierung kam u. a. zu folgenden Schlüssen: Beamte haben eine positive Haltung gegenüber dem Bedürfnis nach Wandel und Modernisierung; viele hatten den Eindruck, dass das vorliegende Programm nicht angemessen erläutert wurde und schwer verständlich war und alle Beamten hatten einen starken Wunsch nach mehr Information über bestehende Programminitiativen. Das Ergebnis war eine Reihe von Empfehlungen für die Entwicklung „einer formellen Kommunikationsstrategie für den Prozess des Wandels und der Modernisierung“ und die Verbesserung der „Mitteilungsstruktur, des Inhalts und der Zielrichtung des Kommunikationsprozesses.“²²

Ähnlich wie viele Programmevaluierungsberichte besteht der Hauptteil aus Ergebnissen, Schlussfolgerungen und Empfehlungen, wogegen ein Anhang die Evaluierungsmethodik erläutert und den Text mit einigen Datenaufstellungen veranschaulicht. Im Methodikanhang können ORKB und andere Angehörige der Evaluierungsgemeinschaft Einzelheiten über die Zielgesamtheit, die Auswahlgrundlage und das Auswahlverfahren, die Umfragerücklaufquoten und

Analysemethoden des Evaluierungsteams finden.²³ Die Arbeit kann Ergebnissevaluierung unter Anwendung retrospektiver Methoden genannt werden.

Der dänische Rechnungshof, der vor einigen Jahren umorganisiert wurde, um seine Fähigkeit „zur Durchführung von Querschnittsprüfungen zu verbessern und Fachkräfte zur Stärkung der Wirtschaftlichkeitsprüfung und zur Bewältigung der mit der Einführung der elektronischen Datenverarbeitung verbundenen Herausforderungen zu gewinnen,“ aktualisiert laufend seine Arbeitshilfen und macht seine Publikationen auf seiner Website leicht zugänglich.²⁴ ORKB können dort auch ein kleines, aber nützliches Buch mit dem Titel „Wirtschaftlichkeitsprüfung in der Praxis“ (*Performance Audit in Practice*) finden, dessen Zweck darin besteht,

Momentaufnahmen über die Art und Weise der Durchführung von Wirtschaftlichkeitsprüfungen auf kommunaler und regionaler Ebene sowie der anschließenden Berichterstattung zu liefern. Der betreffende Arbeitskreis will damit zur fortgesetzten Weiterentwicklung des Prüfungswesens im öffentlichen Sektor beitragen, um für jede einzelne Gemeinde und Region eine qualitativ hochwertige und nützliche Finanzkontrolle zu gewährleisten, und zwar unabhängig davon, wer der Prüfer ist und wer geprüft wird.²⁵

In dem Buch wird erläutert, dass die in „Wirtschaftlichkeitsprüfung in der Praxis“ zusammengestellten Prüfungen aus zwei Arten bestehen: laufende und erweiterte Wirtschaftlichkeitsprüfungen. Grundlagen einer laufenden Wirtschaftlichkeitsprüfung sind „Analysen und Bewertungen, die zur Unterstützung der Prüfung der Rechnungsführung vorgenommen werden“; in aller Regel „zeigen diese kritische Gebiete auf, in denen Bedarf für eine weitergehende Prüfung besteht.“ Eine erweiterte Wirtschaftlichkeitsprüfung ist dagegen eine Prüfung „tiefergehender bzw. breiter angelegter Natur“, was „in der Regel die Gelegenheit zur Beantwortung kritischer Fragen gibt.“²⁶

Der Abschnitt mit dem Titel „Wie wird eine erweiterte Wirtschaftlichkeitsprüfung durchgeführt?“ merkt im Besonderen an, dass

die erweiterte Wirtschaftlichkeitsprüfung unabhängig von z. B. der laufenden Prüfung, aber häufig als eine Erweiterung derselben durchgeführt wird, wenn die Prüfung Gebiete aufzeigt, die höhere

Aufmerksamkeit erfordern. Im engeren Sinne betreffen solche Prüfungen häufig das Management auf ausgewählten Gebieten einschließlich

- Bewertungen der Mittelbewirtschaftung
- Strukturelle Arbeitsplanung
- Ermittlung des finanziellen Spielraums für Zwecke der politischen Prioritätensetzung
- Verfahrensanalyse
- Bewertungen der Qualität der Sachbearbeitung
- Bewertungen der Zielerreichung und der Angemessenheit der festgelegten Ziele.²⁷

Die 15 Zusammenfassungen, die auf diese Beschreibung folgen, umfassen eine breite Palette von Studien, die von Kinderheimen über Benchmarking zwischen Krankenhausabteilungen, kommunaler Gehaltsabrechnung, Finanzmanagement der örtlichen Feuerwehr bis hin zur IT-Sicherheitspolitik reichen. Mehrere Studien machen den Eindruck, dass vorausschauende Methoden angewandt wurden, obwohl sie als Ergebnis-, Wirkungs- oder Prozessevaluierungen eingestuft werden konnten.

Das Ziel der Studie Nr. 7 z. B. war die Erfassung und Bewertung organisatorischer Bedingungen für die Planung und Umsetzung „eines anstehenden Projekts, das die Verschmelzung der Versorgungsbetriebe [einer] Kommunalverwaltung zu einem übergreifenden Versorgungsunternehmen betraf.“ Zu den Arbeitsmethoden gehörten „Interviews mit Betroffenen sowie [Prüfung] von Finanzdaten, Funktionsbeschreibungen und andere schriftliche Unterlagen.“ Zu den Ergebnissen gehörten Vorschläge zur Einrichtung der gemeinschaftlichen Versorgung, der Entwurf eines Kontenplans für die gemeinschaftliche Versorgung und die Unterstützung der Prüfer bei „der Erstellung eines strategischen Aktionsplans für die gemeinschaftliche Versorgung.“²⁸

In gleicher Weise war das Ziel der Studie Nr. 16,

die Begutachtung des Integrationssektors . . . die Bewertung der Möglichkeiten der Kommunalverwaltung zur Kontrolle der Verknüpfung zwischen Sprachunterricht in Dänisch und Maßnahmen zur Schaffung von

Arbeitsplätzen. So waren z. B. Managementmöglichkeiten sowohl in Bezug auf die aktuelle Struktur als auch auf alternative Organisationsformen zu bewerten.²⁹

Es wurden leitende Mitarbeiter einschließlich des Verwaltungsleiters, der Mitarbeiter der Integrationsstelle und des Verwaltungspersonals der kommunalen Sprachschulen interviewt. Um festzustellen, „ob die aktuelle Struktur des Eingliederungsprozesses ausreichende Gelegenheit bietet, das Ziel der Verknüpfung zu erreichen,“ legten die Prüfer der Kommunalverwaltung am Ende „eine Reihe von alternativen Strukturmodellen sowie die Finanzvor- und -nachteile der verschiedenen Modelle“ vor.“³⁰

Schließlich tauchen Modelle für Evaluierungskonzepte und -methodik häufig in Vorträgen auf, die bei Fachkonferenzen, Seminaren und Arbeitstreffen zum Thema Evaluierung gehalten werden. Um nur ein Beispiel zu zitieren: ein schwedischer Beitrag, der im Jahre 2000 auf einer Sitzung über „Evaluierung aus der Perspektive der Bürger“ bei der Konferenz der Europäischen Evaluierungsgesellschaft in Lausanne, Schweiz, vorgestellt wurde, fasste die Evaluierungsziele wie folgt zusammen: „Gewinnung neuer Erkenntnisse zur möglichen Weiterentwicklung der Strategie und der Geschäftsprozesse der [schwedischen] Steuerverwaltung“, „Verbesserung ihrer analytischen Fertigkeiten und der Steuerzahlererhebungen als Mittel zur Datensammlung“ und „Erprobung der Nützlichkeit der Modellierung struktureller Gleichungen [*SEM - structural equations modeling*] in diesem Zusammenhang.“³¹

Der Methodikabschnitt des Beitrags beschrieb dann Aufbau und Analyse statistischer Modelle, mit denen der Prüfer mehrere Fragen beantworten kann, die dazu gedacht sind, die Evaluierungsziele zu erreichen:

Die Methodik bietet ein wissenschaftliches Arbeitsmittel, um Messfehler zu berücksichtigen, Modelle auf ihre Eignung zu testen, Vertrauensintervalle für die Bedeutung der erklärenden Variablen anzugeben, Gruppenabweichungen zu testen, Feedback aus möglichen Verbesserungen der vorgeschlagenen Modelle zu gewinnen usw. Mit einem eigenen Design und repräsentativen Stichproben ohne übermäßige Antwortausfälle schafft die SEM-Modellbildung sichtbare Beziehungen von echter Bedeutung für die

Verknüpfung von Meinungen und Verhalten der Bürger mit gesellschaftlichen Anstrengungen, um deren Bedarf an Dienstleistungen zu decken und gleichzeitig anderen Aufgaben gerecht zu werden. . . . Kurz gesagt, der statistische Untersuchungsansatz garantiert, dass jeder Bürger eine faire Chance hat, die Modellbildungsergebnisse zu beeinflussen, was unser Weg ist, die Bevölkerung für Evaluierung zu interessieren.³²

Zweifelloos kann über Evaluierungsmethodik und -praxis sowohl aus Anleitungen als auch aus Beispielen viel gelernt werden.

[4.2] Setzung eines Rahmens für die Projektmanagementplanung

Organisationen, die Evaluierungen durchführen, werden einen Rahmen für das Management von mindestens vier Phasen von Programmevaluierungen erarbeiten wollen: Konzepterstellung, Datensammlung, Datenanalyse und Vorlage der Ergebnisse. Jede Phase eines Plans zur Durchführung von Evaluierungen wird einer Reihe von Sachzwängen unterliegen hinsichtlich Kosten, Zeit, Ort, verfügbare personelle und sonstige Mittel. Viele Modelle solcher Pläne sind an vielen Orten verfügbar. So führt z. B. in der Arbeitshilfe „Praktische Evaluierung für Führungskräfte im Öffentlichen Dienst“ das Kapitel 9 unter dem Titel „Was kommt als nächstes?“ den Leser ohne Umstände entlang einer Reihe von vorgeschlagenen Meilensteinen, die auf den Erfahrungen des Büros für Evaluierung und Inspektionen des U.S.-Gesundheitsministeriums beruhen. So wie es dieses Ministerium getan hat, können ORKB ihren eigenen Evaluierungsprozess gestalten, indem sie jene Meilensteine als Fragen formulieren:

- Wie werden die möglichen Fragestellungen identifiziert? Wie werden Beschlüsse über die zu untersuchenden Fragestellungen getroffen? Wer wird einbezogen und wer trifft den endgültigen Beschluss?
- Wie werden die Prioritäten gesetzt?
- Wie werden externe Stakeholder sowohl in die Planung als auch in die Durchführung von Evaluierungen einbezogen?

-
- In welchen Phasen wird das [Evaluierungs]Team ein Produkt erstellen? Wer ist dessen Adressat?
 - Werden in irgendeiner Phase formelle Stellungnahmen eingeholt? Von wem?
 - Welche Art von Produkt entsteht aus der Evaluierung?
 - Wird das [Evaluierungs]Team Empfehlungen abgeben? An wen?
 - Gibt es einen Mechanismus, um die Nutzung der Information sicherzustellen?³³

Ein völlig anderes Modell aus Kanada legt nahe, dass die Steuerung der Berichterstattung über Ergebnisse sich zu einer Berichterstattung über Wirkungen weiterentwickeln kann. Aus der Perspektive von Wirtschaftlichkeitsprüfungen offeriert es unter dem Titel „Darstellung eines Beispiels für hohe Wirtschaftlichkeit“ eine anspruchsvolle Management-Metapher:

Eine Berichterstattung über die Wirtschaftlichkeit im weitesten Sinne erfordert sowohl die Darstellung des Erreichten im Verhältnis zu dem, was erwartet wurde, als auch die Beschreibung der gewonnenen Erkenntnisse. . . . Es ist mehr erforderlich als lediglich die Darstellung des Erreichten im Vergleich mit vorher festgelegten Maßstäben oder Zielen. . . . Notwendig ist die überzeugende Darstellung eines Beispiels für hohe Wirtschaftlichkeit. . . . Eine fachlich solide Arbeitsweise muss sicherstellen, dass die Strategie der Programmergebnismessung Spielraum für die Bewertung unbeabsichtigter, mit dem Programm verbundener Ergebnisse lässt.³⁴

Obwohl die Diskussionsunterlage „Ergebnisberichterstattung: Setzen von Wirtschaftlichkeitszielen und Darstellung von Beispielen für hohe Wirtschaftlichkeit“ hauptsächlich auf die Wirtschaftlichkeitsprüfung abstellt, zeigt sie mögliche alternative Ansätze für Planung und Berichterstattung einer beliebigen Evaluierung auf. Außerdem gibt sie ORKB einige Arbeitshilfen an die Hand, die sie für Zwecke der Steuerung von Programmevaluierungen anpassen können.

Stärker traditionell ausgerichtete Rahmenkonzepte für Evaluierungen von der Planung bis zum veröffentlichten Bericht sind sowohl im Internet als auch gedruckt

in großer Zahl verfügbar. Seitdem die Programmevaluierung in dem Prüfungshandbuch des amerikanischen Rechnungshofes unter Wirtschaftlichkeitsprüfung subsumiert wird, können Einrichtungen, die Evaluierungen durchführen, einschlägige Passagen mit fachlichen Normen für örtliche Erhebungen bzw. für die Berichterstattung in den Kapiteln 7 bzw. 8 als Ausgangspunkt nehmen. Wenn z. B. Abschnitt 7.06 (in Kapitel 7) fordert, dass „Prüfer den Arbeitsablauf planen und dokumentieren müssen“, ist gemeint, dass „Planung ein kontinuierlicher Prozess während des gesamten Verlaufs der Prüfung [ist].“³⁵ Im Einzelnen umfasst das Prüfungsvorhaben die Projektziele sowie Umfang, Methodik und Risikoabschätzung, außerdem Fragestellungen zur Aufsicht, das Nachvollziehen von Vorgängen anhand von Stichproben, die Behandlung der Ergebnisse und die Dokumentation.

Nach Kapitel 8 (Abschnitt 8.04) soll „sich die Prüfungsberichterstattung nach dem beabsichtigten Zweck richten. Berichte sollten in schriftlicher bzw. einer anderen gespeicherten Fassung vorliegen und auf die Bedürfnisse der Berichtsadressaten zugeschnitten sein.“³⁶ Vor allem sollte die Berichterstattung Folgendes umfassen:

Ziele, Umfang und Methodik . . . ; Je nach den Umständen, Prüfungsergebnisse einschließlich Sachverhaltsfeststellungen, Schlussfolgerungen und Empfehlungen, eine Erklärung zur Einhaltung der allgemein anerkannten Normen für die staatliche Finanzkontrolle; Zusammenfassung der Stellungnahmen der zuständigen Führungskräfte; und Angaben zur Art ggf. weggelassener vertraulicher sensibler Information.³⁷

Eine brauchbare konkrete Präsentation zur Programmevaluierung befindet sich in den einschlägigen Abschnitten der EU-Studie „Evaluieren von EG-Aktivitäten: Praktischer Führer für die Kommissionsdienststellen“. Diese Arbeitshilfe behandelt komplett mit Schaubildern und Grafiken die kostenmäßigen und zeitlichen Beschränkungen von Projekten, die Erarbeitung der qualitativen und quantitativen Analysepläne und die Erfordernis, Transparenz und andere Qualitätsfaktoren in der Berichterstattung über Evaluierungsergebnisse zu gewährleisten. So verbindet z. B. der Abschnitt 3.3.6 den Evaluierungsarbeitsplan mit dem jeweiligen Budget:

Der Arbeitsplan ist immer mit dem Entscheidungsbedarf verknüpft (z. B. dem Termin, zu dem die Information aus der Evaluierung nachgefragt wird . . .) und sollte deshalb niemals ohne Beachtung dieses Sachzwangs entworfen werden. Wenn die Verknüpfung zum Entscheidungsprozess hergestellt wurde, ist es wichtig, klare Fristen für die verschiedenen Phasen des Evaluierungsprozesses, die Treffen mit dem Lenkungsausschuss und die Pflichtergebnisse zu setzen, die im Verlauf der Evaluierung zu liefern sind.³⁸

[14] DARSTELLUNG DER ERGEBNISSE:

Evaluierungsberichte werden meist verfasst, um entweder den Fortschritt bei Erreichung der vorgegebenen Zwecke und Ziele darzustellen und Gebiete zu finden, auf denen Verbesserungen möglich sind (formativer Bericht), oder die Gesamtwirkungen Maßnahmen zusammenzufassen (summativer Bericht). Der übergreifende Zweck aller Berichte ist, die Wirkungen des Programms den Programmmitarbeitern, Kunden der . . . Maßnahmen. . . , Mitarbeitern der finanzierenden Stelle sowie der Öffentlichkeit zu vermitteln. Leider werden viele Berichte zum Zwecke der Rechenschaftslegung an die finanzierenden Stellen erstattet und von den möglichen sonstigen Adressaten nicht zur Kenntnis genommen. Dies liegt vielleicht teilweise an den Prüfern, die im Fachjargon schreiben, ohne die Bedürfnisse des Laienpublikums oder der Programmbeteiligten selbst zu berücksichtigen..

Judith Wilde und Suzanne Sockey, Evaluation Handbook (Albuquerque: New Mexico Highlands University, Evaluation Assistance Center-Western Region, Dez. 1995), S. 80-81. ERIC: Education Resources Information Center, Search Results www.eric.ed.gov (2. Mai 2009).

Obwohl die Kommission administrative Vorgaben nennt, die nur für die EU Geltung haben, zeigt der Leitfaden eine Reihe von Schritten auf, die nützlich für jede Programmevaluierung sein können wie z.B.:

Jede Evaluierung ist als Projekt zu planen, das drei getrennte Stadien umfasst: Konzeption, Umsetzung und Nutzung. Der Projektmanager bestimmt, wer an diesen drei Phasen mitwirkt . . .

Bei der Konzeption eines (internen oder externen) Evaluierungsprojekts ist der Zweck der Evaluierung klar und sorgfältig definiert. Das Evaluierungskonzept umfasst folgende Punkte:

- a. Hintergrund, Gründe und Ziele der Evaluierung,
- b. Adressat und Nutzer,
- c. Umfang des evaluierten Bereichs,
- d. Schlüsselfragen,
- e. Einzelheiten zu den verfügbaren Informationen,
- f. Berichte,
- g. Fristen.

Im Falle externer Evaluierung sind auch die vertraglichen, finanziellen und administrativen Klauseln und Qualitätskriterien einzubeziehen. . . ³⁹

Vor allem soll die Organisation, die die Programmevaluierung erfolgreich durchführt, sicherstellen, dass das Evaluierungsteam höchstmögliche Kompetenz sowie Kenntnisse und Fähigkeiten für Konzeption und Durchführung von Evaluierungen und den fachlichen Erfahrungsaustausch besitzt. Ein weiteres Evaluierungshandbuch behandelt diesen Gedanken in vier Kapiteln unter den Titeln „Vorüberlegungen zu Evaluierung“, „Planung“, „Umsetzung“ und „Berichterstattung“. Diese stellen selbst eine Art Checkliste für den Ablauf einer Evaluierung dar.⁴⁰

Am Ende wird auf Entwurf, Konzept und Anfertigung des Evaluierungsberichts eingegangen. Der Aufwand für die Steuerung dieses Projektspekts ist nicht zu unterschätzen. Zu den Erfolgsrezepten gehört es, alle Beteiligten die das Projekt umgesetzt haben wollen und die, die es zuwege bringen werden, an einen Tisch zu

bringen und über das Projekt schon zu Beginn zu diskutieren, wenn die Ideen noch frisch sind. Das Team sollte alle notwendigen Arbeitsschritte in chronologischer Reihenfolge auflisten und daraus einen Gesamtplan mit Fristsetzung für jede Aufgabe zusammenstellen, der an Änderungen im Projektverlauf angepasst werden kann.

[15] STRATEGISCHE PERSPEKTIVE ZUR EVALUIERUNGSKOMPETENZ:

Im Vergleich zu der bestehenden Nachfrage sind die für die Durchführung qualitativ hochwertiger Evaluierungen benötigten Kenntnisse sowohl in der öffentlichen Verwaltung als auch im privaten Sektor kaum ausreichend vorhanden.

. . . Um gute Entscheidungen über die durchzuführenden Evaluierungen zu treffen, müssen Berater und Entscheidungsträger wissen, was die verschiedenen Arten von Evaluierungen zu welchen Kosten und in welchem Zeitrahmen liefern können. Fehlt dieses Wissen jedoch, werden Evaluierungen oft geplant, ohne den Rat von Evaluierungsfachleuten einzuholen. Dies führt zu einer schlechten Auswahl des zu Evaluierenden. . . . Außerdem äußerten viele Befragte ihre Sorge darüber, dass Politiker und Spitzenbeamte im Allgemeinen unzureichende Kenntnisse in Evaluierung und statistischer Analyse haben. . . . Insbesondere hätten diese zu große Erwartungen an das Maß von Gewissheit, welches Evaluierungsergebnisse erbringen können . . .

Shenagh Gleisner, Arati Mascarenhas und Sonia Wansbrough, —Improving Evaluative Activity in the New Zealand State Sector,|| Vortrag bei der Australasian Evaluation Society Conference, 16.–18. September 2003, Auckland, S. 6–7. Australasian Evaluation Society, 2003 Conference, Paper WE31, —Government Cannot Make This Decision without a Government Study—Yeah, right.|| www.aes.asn.au (2. Mai 2009).

Bei der Erstellung des Prüfungskonzepts sind die Akten bzw. Dateien sorgfältig zu gliedern und in handlicher Größe zu halten. Quellenmaterial kann nach Abschnitt (Titelseiten, Kapitel, Anhänge) getrennt und nach Elementen (Text, Tabellen, Anschauungsmaterial) gruppiert werden, wobei für jede Datei eine Sicherheitskopie zu erstellen ist. Im Bericht sollte eher Aktiv als Passiv bzw. eine für Laien verständliche Sprache verwendet werden und möglichst kein Fachjargon. Das Prüfungsteam sollte sich mit Graphikern und Druckern des Endprodukts wegen einer ansprechenden Gestaltung des Berichts sowie auch der Einhaltung der vorgegebenen Fristen abstimmen. Zur Qualitätskontrolle könnte schließlich jeder, der dazu in der Lage ist, alles Korrektur lesen, was den späteren Lesern vor die Augen kommen wird. Die Transparenz sollte sich nicht nur auf das Evaluierungskonzept, die Nachweise, Feststellungen und Schlussfolgerungen sondern auch auf Ausdruck und Darstellung erstrecken.

- 1 W. K. Kellogg Foundation, *W. K. Kellogg Foundation Evaluation Handbook* (Battle Creek, Mich.: 1. Jan 1998), S. 69–95. www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf (29. April 2009).
- 2 Evaluation Checklists, The Evaluation Center, Western Michigan University, Kalamazoo, Michigan. www.wmich.edu/evalctr/checklists (29. April 2009).
- 3 USAID, EvalWeb, <http://evalweb.usaid.gov> (29. April 2009).
- 4 Burt Perrin, *Moving from Outputs to Outcomes: Practical Advice from Countries around the World* (Washington, D.C.: IBM Center for The Business of Government, Januar 2006), S. 5. USAID, EvalWeb, Evaluation Resources, <http://evalweb.usaid.gov> (29. April 2009).
- 5 Burt Perrin, *Moving from Outputs to Outcomes*, S. 43. (29. April 2009).
- 6 Siehe USAID's EvalWeb *AID Evaluation Handbook: AID Program Design and Evaluation Methodology Report 7* (Washington, D.C.: April 1987), S. 25. Siehe auch S.6, auf der "Evaluierung als Tätigkeit definiert ist, die bei Bedarf im Auftrag der Leitung erfolgt, um vor bedeutenden

Grundsatzentscheidungen bestimmte Angelegenheiten auf deren Relevanz, Wirksamkeit, Wirtschaftlichkeit, Folgen und Nachhaltigkeit zu prüfen.“ USAID Development Partners Resources, USAID Evaluation Publications, USAID Development Experience Clearinghouse, Advanced Search. <http://dec.usaid.gov/partners/evalweb/index.cfm#> (29. April 2009).

- 7 Europäische Kommission, Directorate-General for the Budget. *Evaluating EU Activities: A Practical Guide for the Commission Services* (Luxembourg: Office for Official Publications of the European Communities, Juli 2004), Anhang E, S. 85. European Commission, Financial Programming and Budgeting, Publications. http://ec.europa.eu/budget/index_en.htm (29. April 2009).
- 8 European Commission, *Evaluating EU Activities*, S. 85–90.
- 9 European Commission, *Evaluating EU Activities*, S. 12–13.
- 10 European Commission, *Evaluating EU Activities*, Anhang F, S. 91–96.
- 11 Siehe z. B., Yutaka Suzuki, “Basic Structure of Government Auditing by a Supreme Audit Institution,” *Government Auditing Review* 11 (März 2004): 39–53, besondere Abbildungen 1 und 2. www.jbaudit.go.jp/pr/pdf/e11d03.pdf, as well as the extensive 2007 bibliography at www.iiasiisa.be/egpa/aggroup/Bibliography%20on%20performance%20auditing-july2007.pdf (29. April 2009).
- 12 K. B. Brady, Controller and Auditor-General, *Key Success Factors for Effective Co-ordination and Collaboration between Public Sector Agencies*, Report of the Controller and Auditor-General, Tumuaki o te Mana Arotake (Wellington: New Zealand Audit Office, Okt. 2004). New Zealand, Controller and Auditor-General, Good Practice Guides. www.oag.govt.nz (30. April 2009).
- 13 K. B. Brady, Controller and Auditor-General, *Co-ordination and Collaboration in the Criminal Justice Sector*, Report of the Controller and Auditor-General, Tumuaki o te Mana Arotake (Wellington: 1. Okt. 2003). New Zealand, Controller and Auditor-General, Reports by Type, Performance Audit Reports. www.oag.govt.nz (30. April 2009).

-
- 14 K. B. Brady, Controller and Auditor-General, *Co- ordination and Collaboration in the Criminal Justice Sector*, S. 10–11.
 - 15 K. B. Brady, Controller and Auditor-General, *Co- ordination and Collaboration in the Criminal Justice Sector*, S. 14.
 - 16 K. B. Brady, Controller and Auditor-General, *Co- ordination and Collaboration in the Criminal Justice Sector*, S. 19.
 - 17 K. B. Brady, Controller and Auditor-General, *Co- ordination and Collaboration in the Criminal Justice Sector*, S. 19 und 51.
 - 18 Biljana Ivanovska, State Audit Office, Republic of Macedonia, "Providing Proper Social Care," presentation at EUROSAI Training Event, Seminar on Auditing of Social Security Systems, Prague, 10. November 2008, S. 3. Siehe: Czech Republic, Supreme Audit Office, International Cooperation, Conferences [and] Seminars, 2008, EUROSAI Training Event, Seminar on Auditing of Social Security Systems, Prague, 10.-12. November 2008. www.nku.cz/default-en.htm (30. April 2009).
 - 19 Biljana Ivanovska, State Audit Office, Republic of Macedonia, "Providing Proper Social Care," S. 5.
 - 20 Biljana Ivanovska, State Audit Office, Republic of Macedonia, "Providing Proper Social Care," S. 14–15.
 - 21 Biljana Ivanovska, State Audit Office, Republic of Macedonia, "Providing Proper Social Care," S. 16–17.
 - 22 NGM Market Research and Strategic Planning, *NGM Report on Communicating Change and Modernisation in the Civil Service: A Survey of Irish Civil Servants*, report prepared for Civil Service Communications Strategy Group (Waterford, Ireland: Jan. 22, 2003), S. 5–6. Publications, View Complete List of Publications Organized by Thematic Areas, Quality Customer Service. www.bettergov.ie/eng/index.asp (2. Mai 2009).
 - 23 NGM Market Research and Strategic Planning, *NGM Report on Communicating Change and Modernisation in the Civil Service*, sec. 3.0, Anh.

-
- 1, Methodology, S. 70–75.
- 24 Denmark, Audit of the State Accounts, About Us, History and Background. www.rigsrevisionen.dk (2. Mai 2009).
- 25 Denmark, Audit of the State Accounts, Committee for Public Audit, *Performance Audit in Practice: Examples from County and Municipality Audits* (Copenhagen: Jan. 2003), S. 5. Denmark, Rigsrevisionen, Performance Audit, Methodology and Planning. www.rigsrevisionen.dk (2. Mai 2009).
- 26 Denmark, Audit of the State Accounts, Committee for Public Audit, *Performance Audit in Practice, Tabelle 1*, S. 6.
- 27 Denmark, Audit of the State Accounts, Committee for Public Audit, *Performance Audit in Practice*, S. 11.
- 28 Denmark, Audit of the State Accounts, Committee for Public Audit, *Performance Audit in Practice*, S. 11–12.
- 29 Denmark, Audit of the State Accounts, Committee for Public Audit, *Performance Audit in Practice*, S. 18.
- 30 Denmark, Audit of the State Accounts, Committee for Public Audit, *Performance Audit in Practice*, S. 18.
- 31 Maria Karanta und andere, "A Citizen's Perspective on Public Sector Performance and Service Delivery: Progress in Measurement and Modelling of Data from a Swedish Taxpayer Survey," S. 3, Vortrag bei der European Evaluation Society conference, Lausanne, Switzerland, 12.-14. Oktober 2000. Canadian Evaluation Society, Grey literature. <http://evaluationcanada.ca/site.cgi?s=1> (2. Mai 2009).
- 32 Maria Karanta und andere, "A Citizen's Perspective on Public Sector Performance and Service Delivery," S. 6.
- 33 June Gibbs Brown, *Practical Evaluation for Public Managers: Getting the Information You Need* (Washington, D.C.: U.S. Department of Health and Human Services, Office of Inspector General, Nov. 1994), S. 76–77. www.epa.gov/evaluate/DHHSPracticalEvaluation.pdf (2. Mai 2009).

-
- 34 John Mayne, Office of the Auditor General of Canada, *Reporting on Outcomes: Setting Performance Expectations and Telling Performance Stories* (Ottawa: 14. Mai 2003). Office of the Auditor General of Canada, Reports and Publications, Audit Methodology, Discussion Papers. www.oag-bvg.gc.ca (2. Mai 2009).
 - 35 Comptroller General of the United States, *Government Auditing Standards, July 2007 Revision*, GAO-07-731G (Washington, D.C.: U.S. Government Accountability Office, Juli 2007), S. 124. www.gao.gov (2. Mai 2009).
 - 36 Comptroller General of the United States, *Government Auditing Standards, 2007 Revision*, S. 160.
 - 37 Comptroller General of the United States, *Government Auditing Standards, 2007 Revision*, S. 161.
 - 38 European Commission, *Evaluating EU Activities*, S. 46.
 - 39 European Commission, *Evaluating EU Activities*, S. 69.
 - 40 Judith Wilde und Suzanne Sockey, *Evaluation Handbook* (Albuquerque: New Mexico Highlands University, Evaluation Assistance Center–Western Region, Dez. 1995). ERIC: Education Resources Information Center, Search Results. www.eric.ed.gov(2. Mai 2009).

Kapitel 5: Aufbau von Organisationskapazität für Programmevaluierung

ALLGEMEIN BETRACHTET besitzt KEIN Prüfer im notwendigen Umfang umfassende Evaluierungsfertigkeiten und häufig haben Organisationen, die Evaluierungen durchführen, nicht das gesamte notwendige Fachwissen, um das offenbar breite Spektrum der möglichen Evaluierungskonzepte abzudecken. Evaluierungsfachkräfte können in Teams arbeiten, um Synergieeffekte zu erzielen bzw. Evaluierungsfachkräfte werden für neue Posten eingestellt oder externe Sachverständige werden fallweise eingeschaltet. Teamarbeit oder die Beratung durch externe Experten bei der Durchführung von Evaluierungen kann hilfreich sein und dem Einzelnen und der Organisation neue Kenntnisse und Fertigkeiten als Vorbereitung für die weitere Evaluierungstätigkeit vermitteln.

[5.1] Evaluierungsteam

Inwieweit die praktische Evaluierungstätigkeit dazu dient einschlägige Erfahrungen zu sammeln, wird ausführlich diskutiert. So wurde z. B. auf der Konferenz der Europäischen Evaluierungsgesellschaft im Jahre 2002 in einer Sitzung über die Förderung von Evaluierungen und deren Lerneffekt die Meinung geäußert, dass „es sich meist um Organisationen und nicht um Einzelpersonen handelt, die aus

Evaluierungen lernen sollen, was Fragen darüber aufwirft, wie Organisationen Informationen nutzen und lernen.”¹

Auf der gleichen Konferenz in einer Sitzung über Evaluierung und Restrukturierung wurde auch folgender theoretischer Ansatz vertreten:

Evaluierung hat sich in einer für die Transaktion förderliche und für das Lernen hilfreiche Weise entwickelt. Die Rolle der Evaluierung ist die Vermittlung von Informationen und Kenntnissen an die Organisation, um dieses Fachwissen in der Praxis anwenden zu können. So spielt die Evaluierung eine wichtige Rolle beim Aufbau von lernenden Organisationen, indem sie diese in die Lage versetzt, sich des bei ihnen vorhandenen implizierten Wissens bewusst zu werden bzw. neues Wissen zu schaffen und zu bewahren. . . . Die Prüfer . . . stützen sich auf ihre Kenntnisse über partizipative Evaluierungsformen, Wissen- und Lernarten sowie das Lernen von Organisationen, um diese zu befähigen, . . . selbständig lernende Organisationen zu werden.²

Außerdem wurde erörtert, was genau ein Evaluierungsteam ausmacht. Bei der Konzeption von Programmevaluierungen arbeiten die Prüfer mit Auftraggebern und anderen Stakeholdern zusammen, um Ziele in Fragen umzuwandeln, die die Evaluierung in den durch die verfügbare Zeit und die verfügbaren Mittel gesetzten Grenzen beantworten kann. Insoweit hilft die Mitwirkung von Dritten dem Evaluierungsteam bei der Erstellung eines Konzepts.

Ein anderes Dokument einer früheren Europäischen Evaluierungskonferenz im Jahre 2000 aus einer Sitzung über Evaluierung als Instrument für die Zivilgesellschaft behandelt dieses Thema und stellt fest, dass „die ständige Herausforderung für die Evaluierung darin besteht, genügend Kapazität aufzubauen, um ein breites Spektrum unterschiedlicher Stakeholder in die Lage zu versetzen, Evaluierungen durchzuführen und zu nutzen“ und dass „Befassung mit Evaluierungen nicht nur die Fähigkeit einer Organisation zur Weiterentwicklung und zur Rechenschaftslegung ausbaut, sondern auch gesellschaftliches Kapital generiert.“³

[16] AUFBAU EINER EVALUIERUNGSKULTUR:

1998 gab es neun nationale und regionale Evaluierungsgesellschaften. Diese Zahl stieg bis Ende 1999 auf 26 an. . . . In den letzten Monaten wurden in Westeuropa mehrere neue Evaluierungsgesellschaften gegründet; in Ostmitteleuropa geht die Entwicklung schnell voran; in Russland wurde ein Evaluierungsnetzwerk gebildet; die sehr erfolgreiche Afrikanische Evaluierungskonferenz im Jahre 1999 führte zur Gründung nicht nur der Afrikanischen Evaluierungsgesellschaft, sondern auch zahlreicher nationaler Verbände wie auch in Süd- und Südwestasien, u. a. Japan. In der Karibik und Lateinamerika sind sie in der Planung.

. . . . der Aufbau nationaler und regionaler Evaluierungsgesellschaften ist für das Fachgebiet der Evaluierung von entscheidender Bedeutung, da diese Gesellschaften Fachleute aus mehreren Disziplinen zusammenbringen, um Erfahrungen auszutauschen und fachliche Grenzen zu überwinden, Themen von grundlegender Bedeutung zu diskutieren, fachliche und berufsethische Normen zu setzen, Kenntnisse auf- und auszubauen und als Gruppe mit starker gemeinsamer Identität den künftigen Kurs zu bestimmen. . . . Wenn die Institutionalisierung der Evaluierung als wünschenswert angesehen wird, ist die Häufigkeit der Besetzung von Stabsstellen mit Evaluierungsfachkräften auch ein Indikator für das Angebot an solchen Fachkräften und das Bestehen einer starken Evaluierungskultur..

A. J. Love und C. Russon, —Building a Worldwide Evaluation Community: Past, Present, and Future,|| Evaluation and Program Planning 23 (2000): 449-59. (2. Mai 2009).

Den „Leitlinien für Evaluierungsfachkräfte“ der Amerikanischen Evaluierungsgesellschaft (AEA) zufolge ist sicherzustellen, dass ein Evaluierungsteam über die notwendige Vorbildung bzw. Fachkenntnisse und Erfahrungen verfügt, um zu bestimmen, was für eine konkrete Evaluierung erforderlich ist; außerdem die Fähigkeit, unter fachlichen Gesichtspunkten zu beurteilen, ob das Team die Evaluierung durchführen kann, und die Fähigkeit zur Zusammenarbeit mit anderen Evaluierungsfachleuten, um von diesen zu lernen. Die Leitlinien weisen darauf hin, dass

Evaluierung eine Tätigkeit ist, die von Personen mit unterschiedlichen Interessen ausgeübt wird, die sich zwar auch, aber nicht ausschließlich auf die Evaluierung von Maßnahmen, Produkten, Personal, Politik, Leistung, Vorschlägen, Technologie, Forschung, Theorie und sogar auf die Evaluierung der Evaluierung erstreckt

und dass angesichts der

Unterschiede zwischen den Prüfern im Hinblick auf Vorbildung, Erfahrung und Arbeitsumfeld unterschiedliche Auffassungen über den primären Evaluierungszweck bestehen. Zu den möglichen Zielen gehören u.a. . . . die Verbesserung von Produkten, Maßnahmen, Personalwirtschaft, Organisationen, der öffentlichen Verwaltung, eine bessere Wahrnehmung von Verbraucherinteressen und gesamtgesellschaftlichen Belangen; die Förderung sachlich fundierter Entscheidungen und ein umsichtigeres Vorgehen bei Veränderungsprozessen, das Anstoßen notwendiger Veränderungen; allen Stakeholdern durch Sammlung ihrer Daten mehr Einfluss zu geben und sie am Evaluierungsprozess zu beteiligen sowie innovative Einsichten zu gewinnen.⁴

Die „Nationale Wissenschaftsstiftung“ (NSF) in den USA gab einen Hinweis, wie man solche Fachleute finden kann, wenn sie nicht schon Mitglieder eines Evaluierungsteams einer Organisation sind. Es sollten Fachbereichsleitungen von Colleges und Universitäten kontaktiert bzw. selbständige Unternehmer und Berater ermittelt werden sowie bei Forschern oder Prüfern sowie auch den Autoren, die in veröffentlichten Evaluierungsberichten genannt werden, Empfehlungen eingeholt werden.⁵

[17] EVALUIERUNG: Eine französische Evaluierungsgesellschaft

Warum? „Um zur Weiterentwicklung der Evaluierung beizutragen und ihre Nutzung innerhalb öffentlicher und privater Organisationen zu fördern“. Für wen? „Die SFE (Société Française d’Evaluation) möchte ein Versammlungs-, Diskussions- und Schulungsort für alle mit Evaluierung in Kontakt stehende Interessengruppen und Nutzer sein unabhängig davon, ob sie gewählte Amtsträger, öffentliche Bedienstete, Beamte, Forscher, Akademiker, Berater, Innenrevisoren oder andere Evaluierungsfachleute aus dem öffentlichen oder privaten Sektor sind“. Wie? „Sie setzt sich selbst das Ziel, die fachliche Weiterentwicklung sowie die Einhaltung ethischer Normen und ordnungsgemäßer Verfahren zu fördern, um die Qualität ihrer Evaluierungen sowie die angemessene Nutzung ihrer Ergebnisse zu gewährleisten“.

French Evaluation Society (SFE), Evaluation.
www.sfe.asso.fr (3. Mai 2009).

Die Zusammenstellung eines Evaluierungsteams, so die NSF,

ist eine entscheidende Phase im Evaluierungsprozess. Von vornherein muss der soziokulturelle Zusammenhang, auf dem die Maßnahmen oder Vorhaben basieren, berücksichtigt werden. . . . Zumindest sollte der Prüfer bzw. das Team sich der Kultur der Beteiligten und Stakeholder bewusst sein, vor allem in Bezug auf die jeweilige geprüfte Maßnahme.⁶

Manche Einrichtungen nennen die Kenntnisse und Fähigkeiten, die Prüfer besitzen sollten. So entwickelte z. B. die Schweizer Evaluierungsgesellschaft (SEVAL) Normen für die Förderung der Evaluierungsqualität:

„Wer Evaluationen durchführt, ist sowohl vertrauenswürdig als auch kompetent, damit bei den Evaluationsergebnissen ein Höchstmass an Glaubwürdigkeit und Akzeptanz erreicht wird.“ „Planung, Durchführung und Darstellung einer Evaluation ermuntern die Beteiligten und Betroffenen (Stakeholder) dazu, dem Evaluationsprozess zu folgen und die Evaluation zu nutzen.“ Die Evaluierungsfachkräfte sind in der Lage, „[herauszuarbeiten] welche Informationen zur Beantwortung der Evaluationsfragen unabdingbar sind und welche Informationen zwar interessant und wünschenswert, aber für die Beantwortung der Kernfragen unerheblich sind.“⁷

Andere fachliche Normen, die Prüfer einhalten, bestehen laut SEVAL darin, dass sie in der Lage sind, „theoretische Modelle und Wertvorstellungen einzusetzen“, um „die gesammelten Informationen sowie die Ergebnisse zu interpretieren.“ Sie können statistische Methoden anwenden und die kontextbezogenen Faktoren bestimmen, die „genau untersucht werden müssen, um die Planung, Durchführung und Botschaft der Evaluierung passend zu positionieren.“ Sie wissen, wie man Evaluierungen „. . . kostenbewusst“ durchführt und haben die Fachkenntnisse, um „die Methoden und Verfahren [auszuwählen, die] so zweckmäßig wie möglich sind.“ Sie können „die Stärken und Schwächen des Evaluationsgegenstandes beurteilen und darstellen, und zwar in einer Weise, die die Stärken betont und die Problembereiche aufzeigt.“⁸

[5.2] Entwicklung organisatorischer Kompetenz

Es wurde erklärt, dass „jeder Evaluierungsüberlegungen anstellen sollte“ und eigentlich

ein bedeutender Teil der Arbeit eines internen oder externen Prüfers darin besteht, Kollegen und andere Interessengruppen beim Erwerb von Fertigkeiten, Fachwissen und Fähigkeiten zu unterstützen. Ein Prüfer, der der geprüften Stelle die Evaluierung so vermittelt, dass ihre Ergebnisse sich im laufenden Projektmanagement und -vollzug niederschlagen, ist besser als ein Prüfer, der eine perfekt konzipierte Evaluierung mit fundierten Empfehlungen durchgeführt hat, die aber niemand nutzt bzw. weiter umsetzen kann.⁹

Fallstudien von Bundesbehörden in den USA stellten ähnliche Sachverhalte dar. So kam z. B. der GAO-Bericht „Programmevaluierung: Evaluierungskultur und Partnerschaften als Beitrag zum Aufbau von Sachkompetenzen“ zum Schluss, dass

die langfristige Kooperation zwischen der „Behörde für Kinder und Familien“ (*ACF*) und der „Abteilung für Planung und Evaluierung“ (*ASPE*) des US-Ministeriums für Gesundheit und Soziales (*HHS*) dazu beigetragen hat, das Fachwissen der Behörde unmittelbar durch Beratung zu konkreten Evaluierungen und mittelbar durch Aufbau des Fachwissens des Evaluierungsteams zu bereichern. *ASPE* ist zuständig für die Koordination und Beratung bei Evaluierungen. Die *ACF* schilderte die fachliche Unterstützung durch *ASPE* sowie die Beteiligung an gemeinsamen Beschlüssen und die Bündelung von Finanzmitteln, was die Glaubwürdigkeit ihrer Arbeit bei der *ACF* gefördert habe. Bei der *ACF* sind nur unzureichende Fachkenntnisse in Statistik oder Forschung vorhanden. Sie braucht Mitarbeiter, die den Bundesstaaten und Kommunalverwaltungen den Nutzen der Mitwirkung an Evaluierungen erläutern können. Jahrzehntlang finanzierte der *ASPE* Evaluierungen sowie Armutsforschung mit Hilfe von Hochschulen, beauftragten Firmen und staatlichen Stellen. Die *ASPE* erklärte, ihr Engagement in der Armutsforschung habe ihre Sachkompetenz im Bereich Evaluierung gestärkt, weil auf dem Gebiet der Armutsforschung Schnittmengen zwischen den Hochschulen und den mit der Forschung beauftragten Unternehmen bestünden. Sie ist der Auffassung, dass (1) bessere Forschung geleistet wird, weil prominente Wirtschafts- und Sozialwissenschaftler mitwirken und (2) Armutsforschung stärker mit der politischen Analyse als mit anderen Bereichen verknüpft ist.¹⁰

Ein Evaluierungsteam sollte in der Lage sein, einschlägiges Wissen und berufliches Können zu demonstrieren. Eine Organisation, deren Programmevaluierungen erfolgreich geplant, konzipiert, gesteuert und mit fachlichem Austausch der Beteiligten über Evaluierungstechniken einhergeht, baut damit automatisch entsprechende Fachkenntnisse auf und aus.

[18] PRÜFUNG UND EVALUIERUNG ERGÄNZEN SICH GEGENSEITIG:

Im der vorliegenden Studie . . . wird untersucht, wie Prüfungs- und Evaluierungsansätze als fachübergreifende Modelle Informationen für staatliche Maßnahmen und Vorhaben liefern. Der GAO (US-Rechnungshof) und andere Fachorganisationen mit dem Aufgabenschwerpunkt staatliche Finanzkontrolle führen eine Vielzahl von Untersuchungen durch. Typischerweise stellen diese (1) prüfungsbezogene Fragen, bei denen es um die Prüfung der Rechnungsführung, der Recht- und Ordnungsmäßigkeit, Controlling und Zielerreichung bei der Aufgabenerfüllung geht sowie (2) evaluierungsbezogene Fragen, welche die Entwicklung neuer Kriterien, Bewertung der Wirkung oder des Gesamt- bzw. Einzelergebnisses (einschließlich unbeabsichtigter Ergebnisse), Prüfung der fachlichen Stichhaltigkeit von Verwaltungsinformationen (über Vollständigkeit, Genauigkeit und Folgerichtigkeit hinaus) und vorausschauende Analysen umfassen. . . . Prüfungs- und Evaluierungsstudien können verschiedene Informationen zusammenführen und so die Wirtschaftlichkeit umfassender darstellen.

Valerie J. Caracelli und Judith Droitcour, U.S. General Accounting Office, —Evaluation and Audit Cross-discipline Models: A Multiplist Approach to Strengthening Information Quality,|| Vortrag bei der Sitzung 277, Evaluation 2000, American Evaluation Association Conference, Honolulu, 2. November 2000. American Evaluation Association, Training, Conference History, www.eval.org (5. Mai 2009).

Wer Evaluierungen leitet, entwickelt durch die praktische Zusammenarbeit ein besseres Verständnis für deren wesentliche Begriffe, deren Konzeption, Methodik, Berichterstattung und die Bewertung der Evaluierung selbst. Unabhängig davon, ob die Evaluierungen innerhalb der Organisation, die Evaluierungen durchführt, entwickelt und umgesetzt oder mit Hilfe von externen Auftragnehmern durchgeführt werden, bringt der Evaluierungsprozess einen Feedback-Kreislauf mit sich: Die Ergebnisse werden in praktische Informationen für die Berichtsempfänger umgewandelt und der Evaluierungsprozess stärkt die Kompetenzen der Evaluierungsfachkräfte bzw. der Einrichtung, für die sie tätig sind.¹¹

[5.3] Evaluierungsgemeinschaft

Das Internetportal *WWW Virtual Library: World Wide Evaluation Information Gateway* bietet Links zu Evaluierungsgemeinschaften, Bildungs- und Schulungs-Websites, öffentlichen und privaten Evaluierungsorganisationen sowie verschiedenen Methoden- und Analysequellen.¹² Unter *Popular Sites* besteht eine Verknüpfung zu freien Ressourcen für Programmevaluierung und Sozialforschungsmethoden, die führende Website in dieser Kategorie und das beste Beispiel dafür, was heute in den verschiedenen Evaluierungsgemeinschaften verfügbar ist. Von Evaluierungsexperten eingerichtet und gepflegt, konzentriert sie sich auf Methodik und Theorie der Evaluierung und Sozialforschung, liefert aber auch eine Fülle von Informationen über „den Schutz der von Forschung Betroffenen, Regelwerke und Leitlinien für die Forschung, Rechte der Öffentlichkeit [und] berufsständische Pflichten- und Verhaltenskodizes . . .“¹³

Das Portal bietet nicht nur Links zu anspruchsvollen, wissenschaftlichen Quellen, sondern auch zu Leitfäden für Laien, die nur geringe oder keine fachspezifischen Vorkenntnisse haben. Weiterhin hat es Links zu Rechnungshöfen auf der ganzen Welt sowie Hunderten von Evaluierungsgesellschaften und Forschungszentren in der ganzen Welt. Neben der Verlinkung mit Fachzeitschriften von A - Z weist es hin auf Online-Diskussionsgruppen und Adresslisten, bei denen man u. a. Informationen zur Klärung methodischer Probleme erhalten kann.

Neben der *WWW Virtual Library* öffnet die „Internationale Organisation für Evaluierungskooperation“ (*IOCE*) Prüfern den Zugang zur wachsenden globalen Evaluierungsgemeinschaft. Rund 50 nationale und regionale Evaluierungsnetzwerke und -gesellschaften können mittels *IOCE* direkt über Links auf ihren Websites erreicht werden. Neuigkeiten aus weltweiten Evaluierungskonferenzen und -tagungen werden bei der *IOCE* ins Internet gestellt, ebenso Bekanntmachungen über die Gründung neuer Evaluierungsgesellschaften. Fallstudien darüber, wie 14 verschiedene Evaluierungsgesellschaften ihre Gesellschaften in Afrika, Nord- und Südamerika, Asien, Australien, Neuseeland, Ozeanien und Europa aufbauten und betreiben, sind dort ebenfalls verfügbar.¹⁴

Hilfreich kann auch eine Evaluierung sein, die der GAO im Jahr 2003 unter dem Titel „Programmevaluierung: Evaluierungskultur und Partnerschaften als Beitrag

zum Aufbau von Sachkompetenzen“ durchgeführt hat und zu folgender Erkenntnis kam:

Vier Hauptelemente der Evaluierungskompetenz waren über das breite Spektrum der von uns geprüften Stellen hinweg offensichtlich, obwohl sie verschiedene Formen annahmen. Diese Elemente umfassen eine Evaluierungskultur, hohe Datenqualität, analytische Fähigkeiten und auf Kooperation beruhende Partnerschaften. Die Evaluierungskultur der geprüften Stellen zeigte sich durch Bereitschaft zur Selbstevaluierung und Lernen durch Experimentieren. Hohe Datenqualität und analytischer Sachverstand waren der Schlüssel, um die Glaubwürdigkeit von Evaluierungsergebnissen und -schlussfolgerungen sicherzustellen. Die Zusammenarbeit der geprüften Stellen mit anderen Bundesstellen sowie Programmpartnern außerhalb der Bundesverwaltung half dabei, den Nutzen der eingesetzten Ressourcen und des verfügbaren Fachwissens für die Evaluierung zu optimieren.¹⁵

Der GAO-Bericht beschreibt die Evaluierungskultur, die sein Team bei den Evaluierungsfachkräften der geprüften Stellen vorfand:

In drei Fällen – ACF, U.S.-Bundesbehörde für Straßen- und Fahrzeugsicherheit (NHTSA) und NSF – konnte das Vorhandensein einer Evaluierungskultur bestätigt werden: die betreffenden Stellen verfügen über einen formalen, geregelten Prozess für die Planung und Durchführung von Evaluierungen sowie die Nutzung ihrer Ergebnisse. Sie zeigen eine Bereitschaft, aus Analyse und Experimentieren zu lernen. Beim U.S.-Wohnungsbau- und Stadtentwicklungsministerium (HUD) und der Küstenwache finden nach Bedarf Evaluierungen statt. Beamte des HUD beschreiben einen jährlichen konsultativen Prozess, in dem darüber entschieden wird, welche Studien im Rahmen der vorhandenen Haushaltsmittel unternommen werden können.¹⁶

Jenseits dieser institutionellen Grenzen besteht die Gemeinschaft aktiver Prüfer, wie z. B. die mehr als 300 Prüfer aus 35 Ländern — einige von ihnen Mitarbeiter in Finanzministerien und ähnlichen Dienststellen — die im September 1999 an der Konferenz der Afrikanischen Evaluierungsvereinigung AfrEA in Nairobi teilnahmen. Das Thema „Zunehmende Evaluierungskompetenz in Afrika“

generierte rund 80 Abhandlungen aus allen wichtigen Gebieten der Evaluierungsforschung.

[19] ZUSAMMENARBEIT IN DER EVALUIERUNGSGEMEINSCHAFT:

Programmevaluierung begann in den USA mit dem Sozialreformprogramm der US-Regierung unter Präsident Lyndon B. Johnson (*Great Society*). Was damals geschaffen wurde, [waren] neue Vorgehensweisen, um soziale Probleme anzugehen, und eine neue Praxis, die der Evaluierung . . . aber es bestand ein Gegensatz zwischen . . . positivistische und konstruktivistische Paradigmen [W]ir führen noch ein Rückzugsgefecht zwischen qualitativen und quantitativen Methoden.

. . .

Jedoch sollten Auftraggeber und Auftragnehmer . . . gemeinsame Ziele haben Beide . . . sollten zur Ausarbeitung der Maßnahmen, Definition der Ziele usw. beitragen. Sie sollten begreifen, dass sie beide von Evaluierung profitieren können. . . . falls eine kontinuierliche Interaktion zwischen unteren und höheren Hierarchieebenen zustande kommt. . . .

. . . Wüssten die Menschen auf der unteren Ebene, was sie tun, was sie zu erwarten haben und wie sie zum Ergebnis beitragen können, stünden sie der Evaluierung wohlwollender gegenüber und würden begreifen, dass es zu ihrem Besten ist. Auf höherer Ebene wären von unten erhaltene Theorien besser fundiert und besser geeignet, eine komplexe und vielgestaltige Situation zu verstehen.

Nicoletta Stame, —Evaluation and the Policy Context: The European Experience,|| S. 37-39 und 42, keynote address, Australasian Evaluation Society 2003 conference, Auckland, 14.-18. November 2003; siehe Evaluation Journal of Australasia, New Series, 3:2 (Dezember 2003): 36-43.

Australasian Evaluation Society, Publications

www.aes.asn.au/publications (4. Mai 2009)

Eine andere Arbeitsgruppe ist die Europäische Evaluierungsgesellschaft (EEA). Ähnlich wie bei einer Reihe anderer Evaluierungsfachgesellschaften besteht ihr erklärter Zweck darin, die Theorie, Praxis und Nutzung der Programmevaluierung dadurch zu fördern, dass sie Seminare, Konferenzen im zweijährigen Turnus, Schulungen, Arbeitsmöglichkeiten und andere Formen der Vernetzung organisiert und dadurch den Fachpraktikern Gelegenheit zum gegenseitigen Austausch gibt. Auf der Website der EES z. B. wird für das Jahr 2009 nicht nur mitgeteilt, dass ein neues Buch zu Überwachungs- und Evaluierungssystemen einzelner Länder, das gemeinsam von der UNICEF, der Weltbank, der UN-Wirtschaftskommission für Europa, IDEAS (*International Development Evaluation Association*), IOCE und anderen Institutionen herausgegeben wurde, online kostenfrei verfügbar ist, sondern auch, dass schon in den ersten sechs Monaten des Jahres folgende Fachkonferenzen zur Evaluierung abgehalten worden sind oder werden

- in Berlin zum globalen Gesundheitswesen;
- in Brüssel zur Evaluierung humanitärer Aktionen;
- in Wukowar, Kroatien, über die Stärkung der Leistungsfähigkeit örtlicher Gemeinschaften;
- in Potsdam über Politikevaluierung in der Europäischen Union und
- in Südkalifornien an der Claremont Graduate University über Evaluierung und angewandte Forschungsmethoden.¹⁷

[20] BEWERTUNG DER EVALUIERUNG — KONTROLLE ODER KOOPERATION

Im Kontext von Verwaltungsreformen ist es den Obersten Rechnungskontrollbehörden zunehmend wichtig, ihre eigene Leistung zu bewerten und darüber zu berichten Jedoch . . . gibt es verschiedene Schwierigkeiten . . . bei der Bewertung der Wirksamkeit einer Wirtschaftlichkeitsprüfung. So ist es z. B. nicht gewährleistet, dass die Umsetzung von Prüfungsempfehlungen wirklich zu einer besseren Wirtschaftlichkeit führt. Um die Akzeptanz der Prüfungsempfehlungen durch die geprüften Stellen zu fördern, müssten die Prüfer dem Verhältnis zu diesen mehr Gewicht beimessen. Dadurch würde eine größere Nähe zum Tagesgeschäft der geprüften Stelle entstehen, was jedoch wiederum die Unabhängigkeit und Objektivität der Prüfer und der ORKB gefährden könnte.

. . .

Der Zusammenhang zwischen der Prüferrolle und ihrer Wahrnehmung durch die geprüfte Stelle konnte [in der vorliegenden Studie] nicht belegt werden. Die vorliegende Studie zeigt jedoch auf, dass die geprüfte Stelle die beratende Rolle des Prüfers im Rahmen einer konstruktiven Zusammenarbeit stärker anerkannte, als es bei einer kontrollfixierten Aufpasserrolle der Fall gewesen wäre.

Jane Etverk, —Wirksamkeit von Wirtschaftlichkeitsprüfungen: Der Fall Estland,|| M.A. thesis, Department of Public Administration, University of Tartu, Tartu, Estonia, 27. Juni 2002, S. 46–48. National Audit Office of Estonia, About NAO, Documents, Other Documents, Tallinn, Estonia. www.riigikontroll.ee/?lang=en (17. April 2009).

ORKB, welche die Programmevaluierung in ihre Organisationskultur einbeziehen wollen, werden feststellen, dass sie damit Neuland betreten und – wie in Kapitel 1 erwähnt – sich neuen Überzeugungen und Werten öffnen müssen. Der marokkanische Rechnungshof hat dies erkannt und auch festgestellt, dass die Aufgabe von Prüfungsinstitutionen darin besteht, ihre Rolle und Verantwortung als „Wächterinnen der Integrität, Rechtmäßigkeit und Wirtschaftlichkeit von Staatsausgaben“ zu verdeutlichen, wohingegen „wenn man eine Prüfung durch eine Evaluierung ergänzen will, Grenzen beachtet werden müssen“, da „Evaluierung häufig dazu dient, vorhandene Maßnahmen zu verbessern“ anstatt zu grundlegenden Änderungen in der Ressourcenverteilung zu führen. Ferner „[kann] eine spezifische Evaluierung sogar zu dem Vorschlag führen, dass ein Programm unverändert bleibt.“¹⁸

Bei der abschließenden Analyse ist die Frage: Wie bauen evaluierende Organisationen Evaluierungskompetenz auf und wie treten sie in die Evaluierungsgemeinschaft ein? Einen Grundsatzkommentar lieferte die Generalinspektorin des Büros für Evaluierung und Inspektionen im U.S. Gesundheitsministerium. Sie sagte,

[W]ir müssen oft Maßnahmen beurteilen und evaluieren. . . [mit] Methoden, die im Einklang mit Evaluierungsprinzipien stehen, sich aber von der peinlich genauen Arbeit der Wissenschaftler ziemlich unterscheiden Wir betrachten unseren Ansatz als weit realistischer und unsere Ziele als denen der meisten Programmverantwortlichen sehr ähnlich. Wir prüfen und berichten über Themen wie Wirtschaftlichkeit, Betrug und Missbrauch sowie Programmwirksamkeit. . . . Ähnlich wie viele Programmverantwortliche gehen wir schnell von einem Thema zum anderen über, da wir die Interessen und Bedürfnisse der Entscheidungsträger in der Exekutive und der Legislative antizipieren und auf sie eingehen.

Wir sind daran interessiert, unseren praxisbezogenen Evaluierungsansatz für Programmverantwortliche und andere, die sich nicht selbst als Prüfer ansehen verfügbar zu machen, *indem wir ergebnisorientierte, praktische Information zum sofortigen Gebrauch liefern*. Es ist uns klar, dass die Einführung einer besonderen Art von Evaluierung für viele Maßnahmen wie die unsere ein unerreichbares oder sogar ungeeignetes Ziel sein

kann. Sie verlangt die Unterstützung erfahrener Prüfer, langfristiges Engagement und erhebliche Risikobereitschaft. Unser Ansatz verlangt harte Arbeit. Stellen, die selbst keine Evaluierungen durchführen, können solche jedoch zur Informationsgewinnung wirksam nutzen, wenn sie lernen, wie Evaluierungstechniken auf das Management von Maßnahmen angewendet werden können. Ähnlich wie bei Werkverträgen gilt, dass Sie am Ende umso bessere Ergebnisse in Form von für Sie brauchbaren Informationen erhalten, je enger die Zusammenarbeit mit dem Auftragnehmer ist und je mehr Sie über die Begriffe und Anwendungen einer Evaluierung wissen.¹⁹

Kurz gesagt, muss ein Netzwerk von Fachkollegen nicht örtlich beschränkt sein, sondern kann Experten aus der ganzen Welt umfassen.

- 1 Barbrol Anell und Marja E. Lemne, "Learning from Evaluations: Who Learns What, When, How and Why?" paper presented at the European Evaluation Society 5th biennial conference, Seville, Oktober 2002, S. 7. Canadian Evaluation Society, Grey literature. <http://evaluationcanada.ca/site.cgi>. (3. Mai 2009).
- 2 Barbara Rosenstein und Miri Levin-Rozalis, "Program Evaluation and Organizational Learning: A Theoretical Perspective," Vortrag bei der European Evaluation Society 5th biennial conference, Seville, Oktober 2002, S. 17. Canadian Evaluation Society, Grey literature. <http://evaluationcanada.ca/site.cgi> (3. Mai 2009).
- 3 Mark Bitel, "Taking Evaluation to the People? Who Wants It?!!" S. 1, paper presented at the European Evaluation Society 4th biennial conference, Lausanne, Oktober 2000. Canadian Evaluation Society, Grey literature. <http://evaluationcanada.ca/site.cgi> (3. Mai 2009).
- 4 American Evaluation Association, *Guiding Principles for Evaluators* (Fairhaven, Mass.: Juli 2004), S. 1. American Evaluation Association, Publications. www.eval.org (3. Mai 2009).

-
- 5 John Frechtling, *The 2002 User Friendly Handbook for Project Evaluation*, NSF 02-057 (Arlington, Va.: National Science Foundation, Jan. 2002), Anh. A, S. 84. National Science Foundation, Publications, Search/Browse Publications. www.nsf.gov/index.jsp (3. Mai 2009).
 - 6 John Frechtling, *The 2002 User-Friendly Handbook for Project Evaluation*, Kap. 7, S. 65.
 - 7 Thomas Widmer, Charles Landert, and Nicole Bachmann, *Evaluation Standards of Seval, The Swiss Evaluation Society (SEVAL Standards)* (Friborg: Swiss Evaluation Society, 5. Dez. 2000), S. 4–6. Swiss Evaluation Society, Seval Standards. www.seval.ch/en (3. Mai 2009).
 - 8 SEVAL Standards, S. 6–10.
 - 9 W. K. Kellogg Foundation, *W. K. Kellogg Foundation Evaluation Handbook* (Battle Creek, Mich.: 1. Jan. 1998), S. 68. www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf (3. Mai 2009).
 - 10 GAO, *Program Evaluation: An Evaluation Culture and Collaborative Partnerships Help Build Agency Capacity*, GAO-03-454 (Washington, D.C.: 2. Mai 2003), S. 20. www.gao.gov (3. Mai 2009).
 - 11 Siehe z.B. die diesbezügliche Diskussion der Europäischen Kommission, *Evaluating EU Activities*, European Commission, Directorate-General for the Budget, *Evaluating EU Activities: A Practical Guide for the Commission Services* (Luxembourg: Office for Official Publications of the European Communities, Juli 2004), S. 25–27. European Commission, Financial Programming and Budgeting, Publications. http://ec.europa.eu/budget/index_en.htm (3. Mai 2009).
 - 12 WWW Virtual Library: Evaluation, The World Wide Evaluation Information Gateway, www.policy-evaluation.org (3. Mai 2009).
 - 13 Free Resources for Program Evaluation and Social Research Methods, <http://gsociology.icaap.org/methods> (4. Mai 2009).
 - 14 IOCE, International Organisation for Cooperation in Evaluation.

<http://ioce.net/index.shtml> (4. Mai 2009).

- 15 GAO, *Program Evaluation: An Evaluation Culture and Collaborative Partnerships Help Build Agency Capacity*, GAO-03-454 (Washington, D.C.: 2. Mai 2003), S. 9. www.gao.gov (4. Mai 2009).
- 16 GAO, *Program Evaluation: An Evaluation Culture and Collaborative Partnerships*, S. 9–10.
- 17 European Evaluation Society, News. www.europeanevaluation.org (5. Mai 2009).
- 18 Kingdom of Morocco, Cour des Comptes, "Suggestions and Comments on the Draft Report of the INTOSAI Working Group on Program Evaluation," September 2005, S. 3.
- 19 June Gibbs Brown, *Practical Evaluation for Public Managers: Getting the Information You Need* (Washington, D.C.: U.S. Department of Health and Human Services, Office of Inspector General, Nov. 1994), S. iii–iv. [www.epa.gov/evaluate/DHHSPacticalEvaluation.p d f](http://www.epa.gov/evaluate/DHHSPacticalEvaluation.pdf) (3. Mai 2009).

Kapitel 6:

Zukunftsperspektiven

WIE WIRKEN SICH PROGRAMMEVALUIERUNGEN auf die Beziehungen AUS, die ORKB zu den nationalen Prüfungs- und Evaluierungsgesellschaften, INTOSAI-Arbeitsgruppen sowie öffentlichen und privaten Evaluierungsinstitutionen haben? Welche Rolle haben sie beim Auf- und Ausbau der Evaluierungskompetenzen von Einzelpersonen und der Evaluierungsgemeinschaft in ihrer Gesamtheit? Welche künftige Entwicklung ist zu erwarten?

[6.1] ORKB als Vermittlerinnen des Wandels

Forschungsergebnisse belegen, dass mit der Zunahme von Wirtschaftlichkeitsprüfungen staatliche Prüfer tendenziell, im Rahmen der Kontrolle des Verwaltungshandelns stärker auf Evaluierungen zurückgreifen. Ein solcher Wandel wurde manchmal von der Erwartung begleitet, dass ORKB durch ihre Tätigkeit die Umsetzung von Reformen beeinflussen können, oder dass sie Änderungen im Verwaltungshandeln bewirken, selbst wenn dies im Gegensatz zu ihrer traditionellen Rolle steht. Häufig wurde festgestellt, dass

Rechnungshöfe als Kontrollinstanz im Namen der Öffentlichkeit . . . fungieren sollen. In dieser Rolle jedoch, in der sie ggf. auf der Einhaltung förmlicher Verfahren bestehen muss, können sie als Gegner des Wandels gesehen werden. Es kann der Eindruck entstehen, dass sie sich mehr als geboten einmisch. Dies kann zu Misstrauen bei den geprüften Stellen führen.¹

Gleichzeitig erlebt die Programmevaluierung weltweit ein enormes Wachstum.

Gerade jetzt arbeiten Organisationen wie die Weltbank und die OECD mit den Staaten auf hoher Ebene zusammen, um Anerkennung für die Evaluierung und Nachfrage nach ihr zu schaffen. Diese Anstrengungen sind wesentlich. . . . Wenn Tagungen stattfinden, legen Dutzende, wenn nicht Hunderte von Personen große Entfernungen zurück, um daran teilzunehmen. Nicht bloß ein paar Leute, eine Elite oder Politiker und Beamte. Bei den Prüfern und bei den Nutzern von Evaluierungen besteht ein enormes Interesse an solchen Treffen. In diesen Basisnetzwerken ist enorme Kraft vorhanden, um den gesamten Prozess voranzutreiben. Die Vergangenheit zeigt, dass die Interessenten kommen, wenn die Aufbauarbeit geleistet wird.²

Wie im vorliegenden Papier dargelegt, sind die Teilnehmer an und Redner bei Konferenzen der Evaluierungsgesellschaften und anderen Fachveranstaltungen, häufig Beamte aus staatlichen Finanz-, Prüfungs- und Revisionsdienststellen, die sich mit anderen Fachleuten im Evaluierungsnetzwerk austauschen und von ihnen lernen.

Auf der Suche nach dem, was beim Aufbau der Programmevaluierungskompetenz einer Organisation hilft, hat der GAO festgestellt, dass

der Bedarf an Erkenntnissen über erfolgreiche Vorgehensweisen einige Behörden dazu anspornte, sich als Institution für die Evaluierung zu engagieren. Die von uns geprüften Stellen machten nicht den Eindruck, dass sie den Aufbau einer Evaluierungskultur „nach Plan“ betrieben. Es schien sich eher um einen systematischen, sich selbst verstärkenden Prozess der Selbstprüfung und Verbesserung mit Unterstützung und Engagement der jeweiligen Leitung und der Aufsichtsorgane zu handeln.³

Der GAO zog daraus den Schluss, dass

gleich ob die Evaluierung im Werdegang der jeweiligen Stelle eine wesentliche Rolle spielte oder ob sie erst als Antwort auf neue externe Einflüsse aufgegriffen wurde, ermöglichte das Lernen aus Evaluierungserfahrungen kontinuierliche Verbesserungen bei der Aufgabenerfüllung und der Umsetzung von Maßnahmen sowie den Ausbau eines Wissensfundus.⁴

Für die staatliche Finanzkontrolle scheint der Weg klar zu sein: Eine Umkehr ist unmöglich.

[6.2] Wege zu neuem Wissen

Einer der Aufträge der INTOSAI-Arbeitsgruppe für Programmevaluierung ist, allen ORKB Informationen über Evaluierungen zur Verfügung zu stellen. Der vorliegende Leitfaden mit Quellenangaben wird diesem Auftrag teilweise gerecht. Was bedeutet dies allerdings für Rechnungshöfe ohne bedarfsgerechten Internet-Zugang oder für ORKB, denen für Reisen zu Konferenzen oder Seminaren nur ein schmales Budget zur Verfügung steht? Um dieses Problem zu lösen, müsste die globale gegenseitige Unterstützung zwischen den vielen Evaluierungsfachleuten, die sich individuell oder in Vereinigungen vernetzen können, wohl verdoppelt werden.

Jedes Jahr finden weltweit viele Konferenzen, Seminare oder Arbeitstreffen über Evaluierungstheorie oder Evaluierungsmethoden statt. Neuigkeiten aus diesen Veranstaltungen werden täglich mittels Internet Fachgesellschaften, einer wachsenden Zahl von Prüfer-LISTSERVS und mündlich verbreitet, und dies sogar in den Diensträumen der Rechnungshöfe auf allen Kontinenten. Dabei muss nicht nur das Wissen erworben werden, das in den Fachveranstaltungen verfügbar gemacht wird; es geht auch darum zu lernen, wie man Wissen breiter zugänglich machen kann.

Für den Anfang stehen viele Wege offen. Neue Gelegenheiten können sich z. B. mittels IDI oder der mit ihr verbundenen Institutionen bieten. AEA hält Konferenzen über Evaluierungstheorie, -methodik und -praxis ab und arbeitet auch mit thematischen Interessengruppen von internationaler Bedeutung zusammen — u. a. der Gruppe „Internationale und kulturübergreifende Evaluierung“, häufig unter Vorsitz von Mitgliedern der Weltbank, der lateinamerikanischen PREVAL oder ähnlichen Organisationen. Die Interessengruppe „Internationale und kulturübergreifende Evaluierung“ der AEA gewährt Reisekostenzuschüsse an Evaluierungsfachleute, die in Entwicklungsländern leben und praktizieren, damit sie an deren Veranstaltungen teilnehmen können.

Es wurde schon erwähnt, dass die Zahl der Gesellschaften und demzufolge die Zahl der Evaluierungsfachleute laufend zunimmt, man also die Gelegenheit nur nutzen muss.

[6.3] Unterstützung einer Evaluierungskultur

Die Prüfung der von der öffentlichen Verwaltung durchgeführten Evaluierungen durch die Rechnungshöfe kann einen wirksamen Beitrag zur Verbesserung der nationalen Evaluierungskultur leisten. Die Finanzkontrolle kann die Bedeutung von Evaluierungen für die Sicherstellung eines verantwortlichen Umgangs mit öffentlichen Mitteln aufzeigen, auf Mängel beim Einsatz dieser Mittel hinweisen und Empfehlungen zur Verbesserung geben. In Deutschland, wo die Verwaltung nach dem Haushaltsgesetz verpflichtet ist, Programmergebnisevaluierungen (sogenannte Erfolgskontrollen) durchzuführen, veröffentlichte der Bundesrechnungshof schon 1998 einen Bericht, der viel Unterstützung im Sinne der Fortsetzung und Verstärkung der Evaluierungstätigkeit fand. Im Zuge einer Querschnittsprüfung stellte der Bundesrechnungshof fest, dass die Verwaltung ihre Verpflichtung zur Durchführung systematischer und objektiver Erfolgs- und Wirtschaftlichkeitskontrollen nicht angemessen erfüllt. Der Bundesrechnungshof betrachtet diese Unzulänglichkeiten als Ergebnis ungeeigneter und unklarer Definitionen der Politikprogrammziele und einer fehlenden Planung der anschließenden Evaluierungstätigkeit. Wo die Verwaltung Evaluierungen durchführte, wurden die Ergebnisse bei der Entscheidung über vergleichbare oder Folgeprogramme häufig nicht angemessen genutzt. Im Einklang mit dem im vorerwähnten Bericht gegebenen Rat verbesserte und konkretisierte die Bundesregierung das Rahmenwerk für die Evaluierung staatlicher Ausgabenprogramme. Zurzeit führt der Bundesrechnungshof eine Querschnittsprüfung durch, mit der der gemachte Fortschritt gemessen und Probleme erkannt werden sollen, die in der öffentlichen Verwaltung noch bestehen.

Worin besteht letztendlich das Ziel der Evaluierung? Einerseits, wie der Niederländische Rechnungshof auf der Konferenz der Europäischen Evaluierungsgesellschaft im Jahre 2002 hervorhob, kann die staatliche Finanzkontrolle einer komplexen Art von Isolierung unterliegen:

Ein Rechnungshof kann als Teil der institutionalisierten Lernfähigkeiten der Staatsverwaltung angesehen werden. Im traditionellen politischen Kreislauf der Vorbereitung, Umsetzung und Evaluierung von politischen Vorgaben sowie des Feedbacks der Ergebnisse zwecks Feinjustierung der politischen Vorgaben fällt die Finanzkontrolle eindeutig in die Phase der Evaluierung. Überdies erhöht der relativ hohe Status und der Bekanntheitsgrad einer ORKB die Wirkung ihrer Arbeit. . . .

Die staatliche Finanzkontrolle kann eine Rolle im Lernverhalten der geprüften Stellen, d. h. der staatlichen Behörden und Betriebe, spielen. Von unabhängigen nationalen Rechnungshöfen durchgeführte Prüfungen sind ein institutionalisierter Teil des politischen Kreislaufs in den meisten Staaten der entwickelten Welt. Sie beinhalten die Evaluierung der Rechnungslegung, der Haushalts- und Wirtschaftsführung und — zunehmend — der Ergebnisse öffentlicher Strategien und Maßnahmen. Diese privilegierte Position — eine kontinuierliche Beziehung zu den geprüften öffentlichen Stellen — fördert die Entwicklung des Lernverhaltens und kann letztlich zur Verbesserung von öffentlichen Institutionen und Maßnahmen führen. Einige Warnungen sind jedoch angebracht. Die unabhängige Position der meisten nationalen Rechnungshöfe bringt auch eine relativ starke Trennung zwischen der Welt der Prüfer auf der einen und derjenigen der geprüften Stellen auf der anderen Seite mit sich. Zudem sind die Rechnungshöfe in den meisten Fällen Monopolisten kraft Gesetzes In der Praxis kann sich dies in der Auswahl von Prüfungsthemen mit geringerer Relevanz und einem fehlenden Blick für die Komplexität der Politik in der realen Welt niederschlagen.⁵

Andererseits muss die Evaluierung, um tatsächlich wirksam zu sein, wie alle Lernprozesse Raum bieten für neue Ideen, Entwicklungen und Menschen, die sich gegenseitig unterstützen. Der Leitende Forscher im Zentrum für Ergebnisanalyse und Evaluierung im Sozial- und Gesundheitswesen an der neuseeländischen Massey Universität, Paul Daignan, stellte auf dem Treffen der AEA 2001 Folgendes fest:

Notwendig ist, dass alle Mitarbeiter auf allen Ebenen unserer Organisationen und Strategieplanungen bei allem, was sie tun, stärker evaluierungsorientiert sind. Dies bedeutet jedoch nicht, dass sie alles *Evaluierung* nennen müssen, was sie tun. Sie müssen lediglich dafür sorgen, dass die *Evaluierungsorientierung* der Mitarbeiter gestärkt wird. Sie müssen über geeignete Evaluierungsfertigkeiten, -systeme, -strukturen und -ressourcen verfügen, um sie als Instrumente einzusetzen, wenn sie einen stärker evaluierungsorientierten Arbeitsansatz wählen.

So gesehen, kann die Aufgabe, die Evaluierung in der Tätigkeit einer Organisation zu verankern, besser als eine des Aufbaus von *Evaluierungsfähigkeit*, *Evaluierungskapazität* und sogar *Evaluierungskultur* in allen unseren Maßnahmen, Organisationen und Strategieentwicklungen beschrieben werden. Eine solche Evaluierungstätigkeit muss nicht zwangsläufig Evaluierung heißen, sollte aber zur Erreichung des Ziels beitragen, die Evaluierungsorientierung zu stärken und die Wirtschaftlichkeit und Wirksamkeit von Organisationen, Strategien und Maßnahmen zu verbessern. ... Wenn wir [die Verankerung der Evaluierung] als den Versuch ansehen, die Evaluierungsorientierung von Mitarbeitern zu stärken, müssen wir bereit sein, Evaluierungswissen weiterzugeben, um tatsächliche Evaluierungsfähigkeit aufzubauen.⁶

Auf welcher Grundlage kann altes oder neues Fachwissen weitergegeben werden? Der Präsident des mexikanischen Rechnungshofes, Arturo Gonzáles de Aragón, der 2003 bei der Eröffnungssitzung zur Feier des 50. Gründungstags der INTOSAI sprach, machte dabei folgende Ausführungen:

INTOSAI ist eine strategisch ausgerichtete Basisorganisation für die Bewertung, Prüfung und Weiterentwicklung politischer Vorgaben geworden. Ausgehend von den 34 Ländern, die sich im Jahre 1953 auf diesen Weg machten und zu einer Organisation anwuchsen, die heute 185 Mitglieder umfasst, gab es eine große Zahl von Erfolgen und neuen Zielsetzungen, an deren Erreichung wir alle mitwirkten. ...

Die Komitees und Arbeitsgruppen, die sich um die Entwicklung von

Analysen, Forschungsansätzen und Richtlinien bemühen, sind für INTOSAI von besonderer Bedeutung. Unter ihnen können wir die . . . Arbeitsgruppen für Privatisierung, Umweltprüfung [sowie] Programmevaluierung nennen . . .

Die Rechnungshöfe ziehen aus diesen Bemühungen große Vorteile . . .

...

. . . Wenn die pluralistische Welt, in der wir heute leben, besser als die Welt von 1953 ist, als INTOSAI gegründet wurde, so können wir zutreffend behaupten, dass diese Organisation dazu erheblich beigetragen hat. . .

Dank INTOSAI haben wir gelernt, dass unsere Mission darin besteht, hauptsächlich das Wohlergehen der Gesellschaften anzustreben, die uns mit der großen Verantwortung betrauten, anderen zu dienen.⁷

Indem er die Ausführungen der Gründungstagsfeier zusammenfasste, fügte der INTOSAI-Generalsekretär, Franz Fiedler, hinzu, dass nach 1953 INTOSAI

dramatischer Mitgliederzuwachs Auswirkungen auf die Strukturen der INTOSAI hatte und nach einer breiteren organisatorischen Grundlage verlangte, woraus sich . . . die Einrichtung von Komitees und Arbeitsgruppen ergab, [die, nachdem ihnen] spezifische Studienthemen zugeteilt wurden, sich zu „Think Tanks“ innerhalb von INTOSAI entwickelten. Die grundlegenden Dokumente, die sie im Laufe der Zeit zu einem breiten Spektrum von Themen ausarbeiteten, sind allgemein anerkannt als praxisorientierte Leitlinien für die tägliche Prüfungsarbeit der INTOSAI-Mitglieder (und haben sich als solche bewährt).

...

Die INTOSAI begrüßt eine neue, moderne Strategie, da sie niemals eine in überholten Auffassungen verwurzelte Organisation war, aber es muss eine sich ständig wandelnde Strategie sein. Dieser Wandel war jedoch

nie revolutionär und auf abrupte Änderungen gerichtet, sondern evolutionär in seiner Art unter Berücksichtigung der Interessen der Mitglieder . . .⁸

Erst kürzlich äußerte sich die „Internationale Organisation für Kooperation in Evaluierung“ in ganz ähnlichem Sinne, in dem sie feststellte, dass

wir in einer Welt leben, die kontinuierliche Verbesserungen bei der Wirtschaftlichkeit des Verwaltungshandelns, der Haushaltskontrolle und Transparenz; . . . und wirksame Erzielung von Ergebnissen im öffentlichen und privaten Sektor sowie in der Zivilgesellschaft verlangt. Die Evaluierung erfüllt diese Bedürfnisse durch Lieferung von Feedback über das, was sich bewährt hat, vertieft unser Verständnis der Prozesse zur Umsetzung politischer Vorgaben, konzipiert Wissensmanagementsysteme, mit denen Institutionen und Systeme lernen können und entwickelt Kapazitäten, um wirksam zu verwalten und Innovationen durchzuführen. In vielen Ländern und auf allen Kontinenten schließen sich Prüfer zu Fachgesellschaften und –vereinigungen zusammen. . . . In den 80er Jahren gab es nur drei nationale und regionale Evaluierungsgesellschaften; Ende der 90er Jahre waren es neun und zu Anfang des 21. Jahrhunderts ist die Zahl auf fünfzig gestiegen.⁹

Um es mit den Worten eines alten chinesischen Sprichwortes auszudrücken: „Aus vielen Schaufeln Erde wird ein Berg. Aus vielen Eimern Wasser ein Fluss.“

-
- 1 Jane Etverk, "Measuring Performance Audit Effectiveness: The Case of Estonia," M.A. thesis, Department of Public Administration, University of Tartu, Tartu, Estonia, Juni 27, 2002, S. 16–17. National Audit Office of Estonia, About NAO, Documents, Other Documents, Tallinn, Estonia. www.riigikontroll.ee/?lang=en (30. April 2009).
 - 2 A. J. Love und C. Russon, "Building a Worldwide Evaluation Community: Past, Present, and Future," *Evaluation and Program Planning* 23:4 (2000):

450. (30. April 2009).

- 3 GAO, *Program Evaluation: An Evaluation Culture and Collaborative Partnerships Help Build Agency Capacity*, GAO-03-454 (Washington, D.C.: 2. Mai 2003), S. 16. www.gao.gov (30. April 2009).
- 4 GAO, *Program Evaluation: An Evaluation Culture and Collaborative Partnerships Help Build Agency Capacity*, S. 24.
- 5 Gerard Bukkems und Hans de Groot, Netherlands Court of Audit, "Learning by Auditing: Mission Impossible?" paper for the 5th biennial conference of the European Evaluation Society, Three Movements in Contemporary Evaluation: Learning, Theory and Evidence, Seville, Oktober 10–12, 2002, S. 2 und 5. www.evaluationcanada.ca/distribution/20021010_bukkems_gerard_de_groot_hans.pdf. (30. April 2009).
- 6 Paul Dignan, "Mainstreaming Evaluation or Building Evaluation Capability? Three Key Elements," in: J. Jackson Barnette und James R. Sanders, Hg., *The Mainstreaming of Evaluation*, New Directions for Evaluation Nr. 99 (San Francisco: Jossey-Bass, 2003), S. 12.
- 7 Arturo Gonzáles de Aragón, Auditor General of Mexico, "The Benefits of INTOSAI to Supreme Audit Institutions," S. 122–23 und 127, in: Internationale Organisation der Obersten Rechnungskontrollbehörden, INTOSAI: 50 Jahre (1953–2003), Sonderausgabe (Wien: 2004). INTOSAI: Internationale Organisation der Obersten Rechnungskontrollbehörden, Dokumente, INTOSAI, Allgemeine www.intosai.org (30. April 2009).
- 8 Franz Fiedler, "50 Jahre INTOSAI," S. 142 und 147, in: Internationale Organisation der Obersten Rechnungskontrollbehörden, INTOSAI: 50 Jahre (1953–2003).
- 9 International Organisation for Cooperation in Evaluation, "IOCE as Part of a Broader Movement." <http://ioce.net/overview/general.shtml> (30. April 2009).

Anhang A: Nützliche Links

Die nachstehenden Links dienen zur Orientierung für Prüfer/innen, die ihre Fertigkeiten vertiefen und auf Teamarbeit mit der Evaluierungsgemeinschaft ausdehnen möchten. Sie zeigen einschlägige Foren zur Vernetzung mit anderen Prüfern auf sowie Stipendien und Fördermöglichkeiten, Diskussionsgruppen, Seminare, Fernkurse, Tagungen, Stellenangebote in aller Welt und umfangreiche Fachinformationen.

Die Liste stellt eine Auswahl an geeigneten Websites dar, die mit vielen weiteren interessanten Fachwebsites verlinkt sind.

1. Auszeichnungen und Preise	Kontakt
<p>Canadian Evaluation Society www.evaluationcanada.ca/site.cgi?s=1&ss=0&_lang=en</p>	
<p>Die kanadische Evaluierungsgesellschaft „CES“ verleiht jährlich eine Auszeichnung für Evaluierungsbeiträge in Kanada, auch für Nichtmitglieder. Eine weitere Auszeichnung für besondere Verdienste für die CES ist nur für Mitglieder vorgesehen (unter „Awards“).</p> <p>Fachwettbewerbe im Bereich in Programmevaluierung sind auf Englisch und Französisch für Studenten kanadischer Hochschulen verfügbar. Preise umfassen Geldprämien, sowie Erstattung von Reise- und Teilnahmekosten an der Jahreskonferenz der CES („Student Competitions“).</p>	<p>Canadian Evaluation Society Secretariat 1485 Laperriere Avenue Ottawa, Ontario Canada K1Z 7S8 Tel.: (613) 725.2526 Fax: (613) 729.6206 secretariat@evaluationcanada.ca</p>
<p>Italian Evaluation Association (Associazione Italiana di Valutazione) www.valutazioneitaliana.it</p>	
<p>Die italienische Evaluierungsgesellschaft „AIV“ verleiht jährlich Auszeichnungen für die beste Abschlussarbeit über einschlägige Themen aus dem Bereich der Programmevaluierung. Weitere Informationen finden sich unter „Convegni e scuole, Iniziative extra AIV, Premi di Valutazione“.</p>	<p>Associazione Italiana di Valutazione Segretario nazionale Piazza Vittorio Emanuele 25 00185 ROMA c/o Parsec Tel.: 06.4465409 Fax: 06.45438049 segreteria@valutazioneitaliana.it</p>

2. Stipendien und Kursarbeiten	Kontakt
<p>CCAF-FCVI www.ccaf-fcvi.com/english/index.html</p>	
<p>Das internationale Assistenzprogramm für verbesserte Governance und Rechnungslegung dieser nationalen Stiftung dient der Stärkung der nationalen Finanzkontrolle durch praktische Mitarbeit beim Rechnungshof. Schulung und praktische Prüfungserfahrung werden vom kanadischen Rechnungshof und vom Rechnungshof in Quebec vermittelt. CCAF-FCVI fördert auch Fachleute aus ausgewählten nationalen Rechnungshöfen, die in Kanada eine 9-monatige theoretische und praktische Schulung durchlaufen.</p> <p>CCAF - FCVI veranstaltet internationale Symposien, um Gelegenheit zum Gedanken und Erfahrungsaustausch zu bieten und Strategien zur Verbesserung der Wirksamkeit im Bereich Rechnungslegung und Leistungsberichterstattung zu vergleichen. Kurse und Maßnahmen finden in Europa, USA, Australien, Südafrika und Lateinamerika statt. Weitere Informationen finden sich unter: „International Business“.</p>	<p>CCAF - FCVI Inc. 291 Olmstead Street Ottawa (Vanier), Ontario Canada K1L 7J9 Tel.: (613) 241 - 6713 Fax: (613) 241 - 6900 info@ccaf-fcvi.com</p>

3. Fernstudien	Kontakt
<p>Centre for Program Evaluation www.edfac.unimelb.edu.au/cpe</p>	
<p>Das Zentrum für Programmevaluierungs der Universität Melbourne „CPE“ hat Förderangebote wie z.B. Spezialkurse für Nichtstudenten. Informationen über ein Aufbaustudium- und Masterabschlusses in Evaluierung finden sich unter „Course Information“. Einige Lehrveranstaltungen sowie ein Forum sind auch über das Internet zugänglich.</p>	<p>Centre for Program Evaluation Room 422, Alice Hoy Bldg. University of Melbourne Melbourne, Australia Tel.: +61 3 8344 - 8394 Fax: +61 3 8344 - 8490 cpe - enquiries@unimelb.edu.au u</p>
4. Vernetzung und Stellenangebote	Kontakt
<p>UK Evaluation Society www.evaluation.org.uk</p>	
<p>Die Evaluierungsgesellschaft des Vereinigten Königreichs „UKES“ stellt Informationen über ihre Angebote und Informationen über andere Evaluierungsaktivitäten, Stellenangebote, Verträge, Veranstaltungen und weiterführende Literatur in GB und international“ bereit. Auf der Webseite sind Stellenangebote, Foren, Konferenzen, Seminare und Arbeitstagungen (viele kostenlos) in GB und Europa aufgelistet.</p> <p>Interessierte Laien finden hier unter „About us“, dass UKES ein „Mitglied der Akademie der Sozialwissenschaften ist und eng mit der internationalen Gemeinschaft zusammenarbeitet — insbesondere mit der Europäischen Evaluierungsgesellschaft und der Internationalen Organisation für Kooperation bei Evaluierung.“</p>	<p>UK Evaluation Society 37 Star Street Ware, Hertfordshire England SG127AA Tel.: +44 (0) 1920 462411 Fax: +44 (0) 1920 462730 ukes@profbriefings.co.uk</p>

5. Vernetzung und Fernkurse	Kontakt
<p>Nigerien Network of Monitoring and Evaluation www.pnud.ne/rense/HOMEUK.HTML</p>	
<p>ReNSE ist eine informelle Expertengruppe von Prüfungs- und Evaluierungs-fachleuten, deren Ziel es ist, Kapazitätsaufbau und die Einrichtung eines Forums zu fördern, das bei der Definition von Normen, Methodologien für die Berufspraxis helfen soll. Unter „Self - Training in M&E“ finden sich mehrere Fachübungen auf Französisch und Englisch. „Strategic Alliances“- ist mit Informationen über unterschiedliche Evaluierungsaktivitäten einschließlich Finanzierungsmöglichkeiten hinterlegt. Unter „Calls for Evaluators“-findet sich ein Forum zur Erleichterung des Kontakts zwischen afrikanischen Prüfern und anderen internationalen Stellen.</p>	<p>Mr. Robert Ndamobissi M&E Specialist UNICEF, BP 12481, Niamey (Niger) Tel.: (+227) 72 30 08 Fax: (+227) 73 34 68 rndamobissi@unicef.org</p>
6. Vernetzung und Standards	Kontakt
<p>The Swiss Evaluation Society (SEVAL) www.seval.ch/en/index.cfm</p>	
<p>Die Schweizer Evaluierungsgesellschaft „SEVAL“ fördert den Austausch von Informationen zwischen Prüfern in Politik, Verwaltung, Hochschulen, Nichtregierungsorganisationen und dem privaten Sektor. Sie engagiert sich aktiv bei der Verbesserung von Erfahrung in und der Qualität der Evaluierung und ihrer Verbreitung. SEVALs Evaluierungsstandards definieren die Anforderungen an Evaluierungen aller Art und richten sich an alle, die an Evaluierungen mitwirken oder diese nutzen.</p>	<p>SEVAL University of Fribourg Ressort Evaluation, Rectorat Av. de Beauregard 11 (1.217) CH-1700 Fribourg Tel.: +41 (0) 26 300 82 82 Fax: +41 (0) 26 300 96 55 secretariat@seval.ch</p>

7. Vernetzung		Kontakt
<p>GOVTEVAL www.angelfire.com/id/govteval</p>		
<p>GOVTEVAL in Malaysia ist eine selbstständige, weltweite,elektronische Diskussionsgruppe, die jedem offensteht, der an Programmevaluierung im öffentlichen Sektor interessiert ist. Sie wird gemeinsam mit dem Programm für Evaluierung des öffentlichen Sektors am Royal Melbourne Institute of Technology in Australien betrieben. Eines ihrer Ziele ist, Sachverständige und Nutzer der Evaluierung, Akademiker und Forscher in Programmevaluierung im öffentlichen Sektor aus aller Welt zusammenzubringen, um Erfahrungen und Meinungen auszutauschen.</p>	<p>majordomo@nasionet.net</p>	
8. Reisekostenzuschüsse		Kontakt
<p>International & Cross Cultural Evaluation Topical Interest Group www.encompassworld.com/index.html</p>		
<p>I&CCE ist der American Evaluation Association angegliedert. Die Gruppe soll an kulturübergreifenden Themen Interessierten, Gelegenheiten für persönliche Weiterentwicklung geben. So wandte I&CCE in 2005 z. B. 7.500 USD an Reisekostenzuschüssen für internationale AEA-Zuwendungen auf.</p>	<p>Tessie Tzavaras Catsambas EnCompass LLC 10500 Alloway Drive Potomac, MD 20854 USA Tel.: (301) 299 3266 Email: tcatsambas@encompassworld.com</p>	

9. Evaluierungsbezogene Datenbank	Kontakt
<p>The Foundation Directory Online http://fconline.fdncenter.org</p>	
<p>Dies ist eine Datenbank mit über 96.000 Stiftungen, Gebern und anderen gemeinnützigen Einrichtungen mit Angaben über die jüngsten ca. 900.000 Stipendien. Die Einschreibung beim Foundation Center öffnet Zugang zu Stipendien, Schulungen, Büchern und Artikeln, Stellenangeboten sowie Antragsformularen.</p>	<p>The Foundation Center 79 Fifth Avenue/16th Street New York, NY 10003-3076 USA Tel.: (212) 620-4230 Tel.: (800) 424-9836 Fax: (212) 807-3677 fdonline@foundationcenter.org</p>
10. Verschiedene Kontakte	
<p>Viele Evaluierungsgesellschaften weltweit bieten Mitgliedern Vergünstigungen bei Reisen zu und der Teilnahme an Konferenzen, dem Erwerb von Fachliteratur u.ä. an. Einige veröffentlichen auch Stellenangebote.</p>	

Anhang B: Literaturhinweise

In dieser Liste ist eine kleine Auswahl von einschlägiger Fachliteratur aus dem Bereich der Evaluierung zusammengestellt. Es handelt sich um dabei um Standardwerke bzw. um häufig zitierte Fachbücher.

American Evaluation Association. *Guiding Principles for Evaluators*. Fairhaven, Mass.: Juli 2003. www.eval.org.

Bamberger, Michael J., Jim Rugh, und Linda Mabry. *RealWorld Evaluation: Working under Budget, Time, Data, and Political Constraints*. Thousand Oaks, Calif.: Sage Publications, 2006.

Chelimsky, Eleanor, und William Shadish. *Evaluation for the 21st Century: A Handbook*. Thousand Oaks, Calif.: Sage Publications, 1997.

Comptroller General of the United States. *Government Auditing Standards, Juli 2007 Revision*, GAO-07-731G. Washington, D.C.: U.S. Government Accountability Office, Juli 2007. www.gao.gov.

Europäische Kommission, Directorate-General for the Budget. *Evaluating EU Activities: A Practical Guide for the Commission Services*. Luxemburg: Office for Official Publications of the Europäische Kommission, Juli 2004. Europäische Kommission, Financial Programming and Budget Publications,

http://ec.europa.eu/budget/library/publications/financial_pub/eval_activities_en.pdf.

Evaluation: The International Journal of Theory, Research, and Practice. Hg. Sage Publications, in association with Tavistock Institute, London, England. www.uk.sagepub.com/home.nav. *Evaluation* can also be accessed through Sage Journals Online at <http://online.sagepub.com>.

Evaluation and Program Planning: An International Journal. Hg. Elsevier Ltd., Oxford, England. www.elsevier.com/wps/find/journaldescription.cws_home/593/description#description.

Evaluation Checklists, The Evaluation Center, Western Michigan University, Kalamazoo, Michigan. www.wmich.edu/evalctr/checklists.

GAO. *Designing Evaluations*, GAO/PEMD-10.1.4. Washington, D.C.: Mai 1991. www.gao.gov.

INTOSAI-Komitee für die Richtlinien für die Finanzkontrolle. Grundsätze für die Durchführung von Wirtschaftlichkeitsprüfungen Richtlinien und Grundsätze für die Wirtschaftlichkeitsprüfung, beruhend auf den INTOSAI-Richtlinien für die Finanzkontrolle sowie den Erfahrungen der Praxis Stockholm, Schweden: Juli 2004. INTOSAI: Internationale Organisation der Obersten Rechnungskontrollbehörden, Dokumente, INTOSAI, Prüfungsbezogene, INCOSAI Ziel 1. www.intosai.org.

Rossi, Peter H., Mark W. Lipsey und Howard E. Freeman.

Evaluation: A Systematic Approach, 7th ed. Thousand Oaks, Calif.: Sage Publications, 2003.

Sanders, James R. und The Joint Committee on Standards for Educational Evaluation. *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs*. Thousand Oaks, Calif.: Sage Publications, 1994.

Shadish Jr., William R., Thomas D. Cook und Laura C. Leviton. *Foundations of Program Evaluation: Theories of Practice*. Thousand Oaks, Calif.: Sage Publications, 1991.

Shaw, Ian F., Jennifer C. Greene und Melvin M. Mark. *The Sage Handbook of Evaluation*. Thousand Oaks, Calif.: Sage Publications, 2006.

Stufflebeam, Daniel L. und Anthony J. Shinkfield. *Evaluation Theory, Models, and Applications*. Hoboken, N.J.: Jossey-Bass, 2007.

Anhang C: Glossar

Im Folgenden sind Definitionen für die wesentlichen Fachtermini in diesem Leitfaden aufgeführt. Die Definitionen stammen aus einem oder mehreren der folgenden Glossare, von denen vier im Internet unter dem angegebenen Link abrufbar sind.

1. Europäische Kommission, *Evaluating EU Activities*, Europäische Kommission, Directorate- General for the Budget, *Evaluating EU Activities: A Practical Guide for the Commission Services* (Luxemburg: Office for Official Publications of the European Communities, July 2004), Annex I, Glossary, pp. 101–10. Europäische Kommission, Financial Programming and Budgeting, Publications. http://ec.europa.eu/budget/index_en.htm (6. Mai 2009). [EC]
2. The Evaluation Center, Western Michigan University, Kalamazoo, Michigan. Tools & Resources, Glossaries, Glossary of Program Evaluation Terminology. <http://ec.wmich.edu/glossary/prog-glossary.htf> (6. Mai 2009). [WMich]
3. GAO, *Performance Measurement and Evaluation: Definitions and Relationships*, GAO-05-739
SP. Washington, D.C.: Mai 2005. www.gao.gov (6. Mai 2009).
4. Sandra Mathison, ed., *Encyclopedia of Evaluation* (Thousand Oaks, Calif.: Sage Publications, 2005). [EE]

-
5. U.S. Department of Justice, Bureau of Justice Assistance, Center for Program Evaluation and Performance Measurement, Glossary. <http://www.ojp.usdoj.gov/BJA/evaluation/glossary/index.htm> (6. Mai 2009). [BJA]

Die Definitionen wurden in gekürzter Form aus obigen Quellen zusammengetragen. Begriffsdefinitionen, die nur in einem der Glossare enthalten sind, wurden in der Regel weggelassen. Ebenso wurde auf Einträge verzichtet, die sich auf die Prüfungstätigkeit im Allgemeinen beziehen und deren Kenntnis bei den Rechnungshöfen vorausgesetzt werden kann.

Daneben gibt es eine Reihe weiterer Glossare zu diesem Themenbereich in Papierform und im Internet. Die fünf obenstehenden Glossare wurden wegen ihrer Vollständigkeit und Genauigkeit sowie ihres internationalen Bezugs ausgewählt.

Das auf jede Definition folgende Kürzel gibt an, aus welcher der fünf Quellen sie ursprünglich stammt: (1) EC, (2) WMich, (3) GAO, (4) EE, (5) BJA.

Begriff	Definition
Auswahl, Auswahldesign	Eine Reihe von Einzelpersonen, ausgewählt als Teil oder Untermenge einer Versuchsgruppe, um für diese Gruppe gewollt repräsentativ zu sein; benutztes Verfahren, um die Auswahl zusammenzustellen. [EC, WMich, EE, BJA]
Auswahlgrundlage	Komplette Liste der Versuchsgruppen, die in einer Evaluierung untersucht werden sollen, wie z. B. alle Nutznießer des Programms. [BJA]
Beteiligter. <i>Siehe auch</i> Interessierte Person	Einzelperson, Familie, Nachbarschaft oder Rechtssubjekt, welche Nutzen, Aktivitäten oder Dienstleistungen eines Programms erhalten. [EE, BJA]
Datenanalyse	Systematische Anwendung von statistischen und logischen Techniken einschließlich Modelle, nicht statistischer Analysen und wertender Techniken, Kosten-Nutzen-, Kostenwirksamkeits- und Multikriterienanalysen, um Informationen über ein Programm für Nutzung in einer Evaluierung zu beschreiben, zusammenzufassen und zu vergleichen. [EC, BJA]

Begriff	Definition
Delphi-Technik	Systematische Methode, um Gruppenkonsens zu erreichen; in jeder von einigen aufeinander folgenden „Runden“ werden den Experten die Ansichten der Experten in der vorherigen Runde übermittelt. Die Technik kann abgesandte Fragebögen oder andere Methoden des kontrollierten Feedback an die Befragten verwenden, bis Konsens erreicht wird. [EC, WMich, EE]
Design. Siehe auch Recherche-Design	Gesamtplan für die Durchführung einer Programmevaluierung, der Elemente beschreibt wie z. B. anzusprechende Fragen, Analyseeinheiten, anzustellende Vergleiche, Datensammlung und Analyseplan, zu bemessende Variablen und Konzepte, Untersuchungsgrenzen in Bezug auf Versuchsgruppen, Geografie und Zeit, benötigter Genauigkeitsgrad für Ermittlung nützlicher und zuverlässiger Ergebnisse, Handhabungs- sowie Berichterstattungsplan. [WMich, EE, BJA]
Ergebnisbezogenes Management	Management-Strategie, die sich auf Erbringung und Erreichung von Outputs, Outcomes und Wirkungen konzentriert. [EE]
Ergebnisse	Effekte, Output, Outcome oder Wirkung, egal ob beabsichtigt oder unbeabsichtigt bzw. positiv oder negativ. [EC]

Begriff	Definition
Evaluierung, Evaluierungsprojekt	Systematische Bewertung, bestehend aus einem evaluativen Arbeitsplan innerhalb eines logischen Rahmens für Design, Umsetzung und Ergebnisse eines Programms oder einer Politik, um seinen Wert, Relevanz, Wirtschaftlichkeit, Wirksamkeit oder Nachhaltigkeit zu bestimmen. [EC, WMich, EE, BJA]
Evaluierungsteam	Personal, Berater und andere, die zusammen eine Programmevaluierung planen, konzipieren und durchführen und dazu Datensammlungsinstrumente entwickeln, Daten sammeln und analysieren und den Bericht schreiben. [BJA]
Ex-Ante-und Ex-Post-Design	Ein Design, das Ergebnisindikatoren sowohl vor als auch nach Verfahrenseinführung misst; jede Änderung kann dem Verfahren zugerechnet werden, nachdem alternative Annahmen verworfen wurden. [BJA]
Experimental-Design	Design zwecks Bewertung der Wirkungen eines Verfahrens oder einer Intervention. Der Prüfer ordnet wahllos Beteiligte dem Untersuchungsverfahren und Kontrollgruppen zu und handhabt Auswahl der Beteiligten, Verfahrensart, Reihenfolge und Verwaltung, wie die Daten aufgezeichnet werden sowie Durchführung der statistischen Analysen. [WMich, EE, BJA]

Begriff	Definition
Expertenmeinung	Methode, die berücksichtigten Expertenurteile auf einem bestimmten Gebiet zu sammeln. [EC]
Externe Evaluierung	Eine von einem Prüfer objektiv durchgeführte Evaluierung, welcher der Organisation, die für das evaluierte Programm verantwortlich ist, nicht angehört. [EC, WMich, EE]
Fallstudie, Fallstudiendesign	Eine umfassende Beschreibung und Analyse eines einzelnen Projekts oder Programms, als Ganzes und im Zusammenhang genommen; sie kann auf einer ausführlichen Untersuchung eines oder mehrerer Fälle oder Situationen beruhen. [EC, WMich, BJA]
Feldaufzeichnungen	Schriftlicher Bericht des Evaluators über Beobachtungen, Interaktionen, Gespräche, situationsbezogene Einzelheiten und Gedanken während einer Studienperiode. [WMich, BJA]
Feldexperiment, Feldtest	Eine randomisierte kontrollierte Studie eines Programm in einem Umfeld wie das, in dem es angewendet werden soll, reichend von einer Vorerkundung bis zu einer summativen Evaluierung in Originalgröße. [WMich, EE, BJA]

Begriff	Definition
Implementations- evaluierung	Bewertung, inwieweit ein Programm wie angestrebt arbeitet. Sie bewertet typischerweise die Übereinstimmung der Programmaktivitäten mit gesetzlichen und verordnungsrechtlichen Bestimmungen, Programmdesign und berufsständigen Standards oder Kundenerwartungen. [GAO]
Interessierte Person. <i>Siehe auch Beteiligter</i>	Einzelpersonen oder Gruppen, die direktes oder indirektes Interesse am Programm und seiner Evaluierung haben. [EC, WMich, EE]
Interne Validität. <i>Siehe auch Validität</i>	Feste Überzeugung, dass die Schlussfolgerungen einer Evaluierung eine kausale Verknüpfung zwischen dem Programm und seiner Wirkung herstellen; ein Evaluierungsdesign, in dem die kausale Verknüpfung zwischen der Programmintervention und den beobachteten Wirkungen unsicher bleibt, ist eine Bedrohung der internen Validität. Andere Bedrohungen umfassen Unzuverlässigkeit von Maßnahmen, Verschleiß der Mitwirkenden und statistische Regression. [EC, EE, BJA]

Begriff	Definition
Kontrollgruppe	<p>Eine wahllos zugeordnete Gruppe, um eine Aussage über die Wirkung der Dienstleistungen, Arbeiten oder Produkte des Programms auf die Programmbeteiligten zu machen. Die gleiche Information wird für Personen im Programm (Experimentalgruppe) und Personen in der Kontrollgruppe eingeholt. Die Einzelpersonen in der Kontrollgruppe besitzen Merkmale, die denen der Programmbeteiligten ähneln, beziehen jedoch keine Dienstleistungen, Arbeiten oder Produkte aus dem evaluierten Programm. In einer Evaluierung werden die beiden Gruppen verglichen, um zu bestimmen, welche der Dienstleistungen, Arbeiten oder Produkte des Programms die erwartete Änderung hervorriefen. [EC, BJA]</p>
Kosten-Nutzen-Analyse	<p>Ein gesamtheitlicher Vergleich des sozialen und privaten Nutzens eines Programms mit den anfallenden Kosten, um zu bestimmen, ob der Nutzen die Kosten übersteigt und um wie viel. Ziel der Kosten-Nutzen-Analyse ist die Wahl der Alternative, die den Nettonutzen des Programms maximiert. [EC, EE, BJA]</p>

Begriff	Definition
Kostenwirksamkeitsanalyse	Ein wertender Vergleich von Kosten und Nutzen eines Programms in Bezug auf Zeit, Aufwand, Ressourcen oder Leistung, um Vergleiche mit anderen Wegen zur Erreichung der gleichen Ziele anzustellen. Die Kostenwirksamkeitsanalyse konzentriert sich auf die Kosten der Ergebnisse, begrenzt auf die messbaren Programmziele. [EC, WMich, EE, BJA]
Langzeitstudie	Studie über die gleiche Gruppe über einen längeren Zeitraum. [EE, BJA]
Leistungsbeschreibung	Überblick über die vom Prüfer zu leistende Arbeit einschließlich Zweck, Umfang und Methode der Evaluierung: zu beantwortende Fragen; zugeordnete Ressourcen und Fristen sowie die Berichterstattungsanforderungen. [EC]
Leistungsevaluierung	Vergleich der aktuellen Leistung eines Programms mit dem, was bezüglich Ressourcen und Outputs geplant war, und dies mit dem Ziel der Verbesserung des Programms durch Neuausrichtung der Ressourcen und Neukonzeptionierung seiner Struktur. [BJA]
Leistungsindikator. Siehe Leistungsmaß	

Begriff	Definition
Leistungsmaß, Leistungsmessung	Beurteilung des Fortschritts eines Programms in Richtung auf seine Zwecke oder Ziele oder deren Erreichung, beobachtet anhand von messbaren Indikatoren. Leistung kann in Bezug auf Qualität, Wirtschaftlichkeit oder Wirksamkeit (sowohl Outputs als auch Outcomes) beurteilt werden. [EE, BJA]
Meta-Analysis. <i>Siehe auch</i> Meta-Evaluierung	In den Sozialwissenschaften eine Synthese mehrerer Forschungsstudien über das gleiche Phänomen, um zu einem Gesamtfazit zu kommen; die Studien können evaluativ sein oder auch nicht. [EE]
Meta-Evaluierung	Systematische Analyse einer Reihe von vorhandenen Evaluierungen ähnlicher Maßnahmen, um nicht bloß zusammenzufassen, sondern deren Qualität zu evaluieren, Gesamtprogrammwirkungen abzuschätzen, Unterstützung für Annahmen zu entwickeln oder die Arbeit der Prüfer zu beurteilen. [WMich, EE, BJA]
Methodologie	Verfahren für Sammlung und Analyse von Informationen, wie z. B. Fallstudien und Delphi-Technik. [BJA]

Begriff	Definition
<p>Monitoring. <i>Siehe auch</i> Partizipative Evaluierung</p>	<p>Laufende Durchsicht eines Programm, mit der systematisch Daten zu festgelegten Indikatoren gesammelt werden, um die Unternehmensleitung und Interessengruppen zu informieren, ob die Programmaktivitäten spezifischen Standards oder Anforderungen entsprechen. Bei Monitoring und Evaluierung in partizipativer Weise entscheiden die Interessengruppen zusammen, wie man Fortschritt einschätzt, Daten sammelt und analysiert und auf Feststellungen hin agiert; Monitoring und Evaluierung auf Leistungsbasis kombinieren traditionelle Monitoring-Umsetzung mit Bewertung von Ergebnissen. Programmevaluierung kann die Ergebnisse der Monitoring-Informationen nutzen. [EC, EE, BJA]</p>
<p>Multivariate Analyse</p>	<p>Analyse der Beziehungen zwischen mehreren Variablen. [BJA]</p>
<p>Norm</p>	<p>Einzelner Wert oder Verbreitung von Werten, welche die typische Leistung einer Gruppe bilden. [WMich]</p>
<p>Normative Frage</p>	<p>Frage, die eines Evaluators bedarf, um zu vergleichen, was in einem Programm gegenüber der Norm passiert, nämlich so wie es passieren sollte. [BJA]</p>

Begriff	Definition
Outcome	Dauerhafte Änderungen, die aus Dienstleistungen oder Outputs des Programms resultieren, wie z. B. Änderungen in den Kenntnissen oder Fertigkeiten einer Person. [EC, EE, BJA]
Outcome-Evaluierung	Beurteilung, inwieweit ein Programm seine ergebnisorientierten Ziele erreicht. Sie konzentriert sich auf Outputs und Outcomes (einschließlich unbeabsichtigter Wirkungen), um Programmwirksamkeit zu beurteilen, kann aber auch den Programmprozess abschätzen, um zu begreifen, wie Outcomes erzielt werden. [GAO]
Output	Konkrete Maßnahmen, die durch ein Programm direkt produziert werden. So ist z. B. die Zahl der Personen, die ein Programm ausbildet, Output. [EC, EE, BJA]
Paneldaten, Panel-interviews	Sammlung von Beobachtungen durch wiederholte Interviews mit der gleichen Befragtenauswahl in einem bestimmten Zeitraum. [BJA]
Partizipative Evaluierung	Vom Programmteam oder Aufsichtsbehörde sowie Beteiligten und anderen Interessengruppen konzipierte, durchgeführte, ausgelegte und mit einem Bericht versehene Evaluierung, häufig mit dem Ziel des Evaluierungskapazitätsaufbaus. [EE]

Begriff	Definition
Politik	Ein Leitprinzip, das eine Reihe von Tätigkeiten betrifft, die möglicherweise in Bezug auf Zeit oder Geld grenzenlos sein und unterschiedlichen Menschen zugutekommen können, jedoch auf gemeinsame Ziele gerichtet sind, deren Belange sich nicht ausschließlich auf Grundlage der Realität und Logik lösen. [EC, BJA]
Prozessevaluierung. <i>Siehe Implementations- evaluierung</i>	
Prüfer	Person, die glaubwürdige und nachprüfbare Unterlagen heranzieht, um zu bewerten und Schlussfolgerungen über Wirtschaftlichkeit, Planung, Design und Durchführung eines Programms sowie die Berichterstattung über eine Programmevaluierung zu ziehen. [WMich, EE]

Begriff	Definition
Prüfung, Prüfung einer Evaluierung	<p>Eine unabhängige, objektive und systematische Beurteilung von Kontrollwesen, Risikomanagement und Führungswesen einer Organisation oder eine unabhängige, objektive und systematische Beurteilung der Qualität eines Evaluierungsplans und seiner Umsetzung sowie der Exaktheit und Validität der Evaluierungsergebnisse und -schlussfolgerungen. Prüfungsstandards können auf die Zielerreichung einer Regierungsorganisation angewendet werden, nämlich dass Maßnahmen und Geldmittel sowohl für sparsamen als auch wirtschaftlichen Einsatz freigegeben werden. In den letzten Jahren wurden Prüfungs- und Evaluierungspraktiken als Ressourcen für Entscheidungsträger, verantwortlich für öffentliche Maßnahmen, zusammengelegt. [EC, WMich, EE, BJA]</p>
Qualitative Daten, Qualitative Information	<p>Fakten und andere Informationen, die Kategorisierung oder Auslegung darlegen, können als Beschreibung oder Schilderung präsentiert werden. [WMich, BJA]</p>

Begriff	Definition
Quasi-Experimental-Design	Design, das einige Merkmale eines Experimental-Designs hat; Prüfer können Vergleichsgruppen und maximale Kontrollen heranziehen, um Bedrohungen der Validität zu minimieren, Zufallsauswahl ist jedoch typischerweise unmöglich oder untauglich. Alternativen sind Ex-Ante- und Ex-Post-Vergleiche, unterbrochene Zeitreihen, ungleichwertige Gruppen und Regression-Diskontinuität-Designs. [EC, EE, BJA]
Recherche-Design	Plan der Datensammlung für eine Evaluierung, der das Wo und Wann der Datensammlung und das Wie ihrer Analyse umfasst. [EE, BJA]
Replizieren, Replizierung	Wiederholung oder Duplizierung eines Programm oder einer Evaluierung, in der alles Wesentliche unverändert bleibt. [WMich, BJA]
Restriktion	Jede Einschränkung, wie z. B. verfügbare Geldmittel, in der Planung oder Umsetzung einer Programmevaluierung. [BJA]
Sektorübergreifende Daten	Mit der Zeit zu einem einzigen Punkt angestellte Beobachtungen über Programmbeteiligte. [BJA]

Begriff	Definition
Sekundäre Daten, Sekundärdatenanalyse	In einer Evaluierung benutzte Daten, häufig in großen Datensätzen, die zu Anfang für andere Zwecke gesammelt wurden; erneute Analyse solcher Daten unter Anwendung gleicher oder anderer geeigneter Verfahren, um die Genauigkeit der Ergebnisse der Anfangsanalyse zu überprüfen oder auf verschiedene Fragen zu antworten. [WMich, EE, BJA]
Umfang, Anwendungsbereich	Untersuchungsfeld der Evaluierung einschließlich u. a. der Hauptziele oder von der Evaluierung geprüften Belange, Diskussion über Validität und ihre Belange, Zeitraum und abgedeckte geografische Zone, Komplexität der Aufgabe, Auswahl des Evaluierungsteams und eine Analyse alternativer Wege bezüglich Evaluierungsdurchführung. [EC, BJA]
Validität	Inwieweit ein Instrument genau misst, was es messen soll. Interne Validität bezieht sich meist auf einen spezifischen Fall; externe Validität untersucht, ob es möglich ist, von einem spezifischen Fall auf andere Versuchsgruppen, Einstellungen oder Zeiten zu verallgemeinern. [WMich, EE, BJA]

Begriff	Definition
Vergleichsgruppe	Eine Gruppe von Einzelpersonen, deren Merkmale denen der Beteiligten am zu evaluierenden Programm (Experimentalgruppe) ähnlich sind. Die Einzelpersonen in der Vergleichsgruppe können verschiedene Dienstleistungen, Arbeiten oder Produkte von denen in der zu evaluierenden Gruppe beziehen oder auch nicht; auf jeden Fall erfahren sie nicht die gleiche Behandlung wie die Experimentalgruppe. In einer Evaluierung werden die beiden Gruppen verglichen, um zu bestimmen, welche der Dienstleistungen, Arbeiten oder Produkte des Programms die erwartete Änderung hervorriefen. [WMich, BJA]
Wirksamkeit	Inwieweit festgelegte Zwecke oder Ziele eines Programms sowohl in Bezug auf Ergebnis als auch Wirkung erreicht werden. Wirksamkeit verlangt eine kausale Verknüpfung zwischen dem, was evaluiert wird, und der beabsichtigten Änderung. [EC, EE, BJA]
Wirkung, Wirkungsbeurteilung, Wirkungsevaluierung	Programmeffekt, positiv oder negativ, vorausgesehen oder nicht vorausgesehen, direkt oder indirekt, beabsichtigt oder unbeabsichtigt; eine Wirkungsevaluierung kann sich auf eine Programmbreite, Langzeitergebnisse, einen spezifischen Einfluss des Programms auf ein gewünschtes Ergebnis oder die Netto-programmwirkung konzentrieren. [EC, EE, BJA]

Begriff	Definition
Wirtschaftlichkeit	Maßstab, wie gut ein Programm seine Inputs oder Ressourcen, wie z. B. Geldmittel, Sachverstand oder Zeit, in Outputs, Ergebnisse oder Wirkungen angesichts festgelegter Zwecke und Ziele umwandelte. [EC, EE, BJA]
Wirtschaftlichkeitsprüfung	Systematische Bewertung der Wirtschaftlichkeit eines Programms oder der direkten Outputs einer Intervention und der Handhabung; während sowohl Wirtschaftlichkeitsprüfung als auch Programmevaluierung darauf abzielen, Qualität zu verbessern, untersucht die Evaluierung Relevanz und Langzeitfolgen des Programms. [EC]
Zuverlässigkeit, zuverlässiges Maß	Maß, das konsistente, stabile und vorhersehbare Ergebnisse mit wiederholten Beobachtungen unter ähnlichen Bedingungen im Zeitablauf erbringt; Beschaffenheit oder Verlässlichkeit der Daten im Hinblick auf Qualität der Verfahren und Analysen, die Prüfer nutzen, um sie zu sammeln und zu interpretieren. [WMich, EE, BJA]

Verzeichnis der Veröffentlichungen der ORKB

Im untenstehenden Verzeichnis sind Fallstudien zu Programmevaluierungen und Wirtschaftlichkeitsprüfungen aufgeführt, auf die der vorliegende Leitfaden Bezug nimmt mit Angaben zu Seitenangaben, Prüfungsart, Titel und Verfasser. Die Übersicht beschränkt sich auf Studien, die von Organen der staatlichen Finanzkontrolle erstellt wurden.

Evaluierung	Prüfung	Leitfaden	Titel	ORKB
		✓	Der wirksame Einsatz öffentlicher Mittel: Wales, Strategie 2006–2009	Wales
		✓	Ergebnisberichte: Leistungserwartungen und -resultate	Kanada
		✓	Vergleichende Studie über die Berichte von Aufsichtsorganen: Gute Beispiele und Ratschläge für bessere Berichte	Niederlande

Evaluierung	Prüfung	Leitfaden	Titel	ORKB
	✓		Reform der öffentlichen Verwaltung: Fortschritte, Rückschläge und Herausforderungen	Kanada
	✓		Überwachung der Produktsicherheit	Finnland
✓			Evaluierung von Maßnahmen zur Erkennung und Behandlung von Fehlsichtigkeit	Brasilien
		✓	Querschnittsprüfungen zur Wirtschaftlichkeit	Philippinen
		✓	Organisationskultur: Vermittlung und Wandel von Überzeugungen und Wertvorstellungen	USA
		✓	Strategie zur Verbesserung der Planung von Wirtschaftlichkeitsprüfungen	Thailand
		✓	Wirksamkeit von Wirtschaftlichkeitsprüfungen: Der Fall Estland	Estland
		✓	Jahresbericht, Haushaltsjahr 2002	Brasilien

Evaluierung	Prüfung	Leitfaden	Titel	ORKB
✓			Programmevaluierung: Evaluierungskultur und Partnerschaften als Beitrag zum Aufbau von Sachkompetenzen	USA
		✓	Die Wirtschaftlichkeitsprüfung beim Australischen Rechnungshof	Australien
		✓	Verbesserungsmöglichkeiten bei der Prüfung der jährlichen Leistungsberichterstattung	Australien
		✓	Handbuch für Wirtschaftlichkeitsprüfungen	Niederlande
		✓	Staatliche Prüfungsnormen, Stand: 2007	USA
	✓		Erosionsbekämpfung in Rodrigues	Mauritius
		✓	Evaluierungsrichtlinien	Dänemark,
		✓	Überblick über die Evaluierung	USA
		✓	Leistungsmessung und Evaluierung: Begriffe und Beziehungen	USA
		✓	Das Rechnungshofgesetz	Vietnam
		✓	Handbuch zur Wirtschaftlichkeitsprüfung	Bangladesch

Evaluierung	Prüfung	Leitfaden	Titel	ORKB
		✓	Allgemeiner Leitfaden für Wirtschaftlichkeitsprüfungen	Australien
		✓	Bericht über Wirtschaftlichkeit des öffentlichen Sektors, 2. Ausgabe	Neuseeland
✓			Stadtplanung und -entwicklung	Frankreich
		✓	Planung einer Evaluierung	USA
		✓	7. Global Forum über Vertrauensbildung und Staatsführung, Bürgerbeteiligung und Transparenz	Saudi Arabien
		✓	Erfolgsfaktoren für wirksame Koordinierung und Zusammenarbeit zwischen öffentlichen Dienststellen	Neuseeland
	✓		Koordinierung und Zusammenarbeit in der Strafjustiz	Neuseeland
	✓		Wirtschaftlichkeitsprüfung in der Praxis: Beispiele aus Regional- und Kommunalprüfungen	Dänemark