

# Zur gesellschaftlichen Bedeutung von Evaluation

Reinhard Stockmann

Stockmann, Reinhard

Zur gesellschaftlichen Bedeutung von Evaluation

Saarbrücken: Centrum für Evaluation, 2008.

(CEval-Arbeitspapiere; 15)

**NICHT IM BUCHHANDEL ERHÄLTlich**

**SCHUTZGEBÜHR:** 5 €

**BEZUG:** Centrum für Evaluation (CEval)  
Universität des Saarlandes  
Postfach 15 11 50  
D-66041 Saarbrücken  
info@ceval.de



oder kostenfrei zum Download:  
<http://www.ceval.de>

## Inhalt

1.	Einleitung .....	1
2.	Zweckbestimmungen .....	2
3.	Institutionalisierung der Evaluation .....	5
4.	Literatur .....	8

## 1. Einleitung

Moderne Gesellschaften sind dadurch gekennzeichnet, dass traditionell und religiös bestimmte Glaubensvorstellungen von der Ordnung der Welt durch das Vertrauen auf Rationalität und Fortschritt ersetzt werden. Schon Ende der 60er Jahre hat Daniel Lerner (1968: 387) die Merkmale der Moderne pointiert herausgearbeitet:

Hierzu zählen

- eine wachsende, sich selbst tragende Wirtschaft,
- demokratische Teilhabe im politischen Bereich,
- eine sich an säkularen rationalen Normen orientierende Kultur und
- eine mit persönlichen Freiheiten ausgestattete räumlich, sozial und geistig mobile, leistungsorientierte Gesellschaft.

Mit dieser Konstruktion ist die Vorstellung verbunden, dass die Moderne ein universelles Phänomen ist, das gestaltet und gesteuert werden kann<sup>1</sup>

*Evaluation ist eine Erfindung der Moderne.*

Sie ist einerseits verknüpft mit der Vision eines wirtschaftlichen und sozialen Fortschritts, eines Aufwärtstrebens und ständig Bessermachens und andererseits mit dem Glauben an die Machbarkeit und steuernde Beeinflussbarkeit der gesellschaftlichen Entwicklung.

Evaluation bietet sich sowohl als *Aufklärungsinstrument* an, das Licht in Entwicklungsprozesse bringt, als auch als *Steuerungsinstrument*, das diese Abläufe gezielt zu beeinflussen sucht. Darüber hinaus eignet sich Evaluation jedoch auch für den reflexiven Einsatz, als ein *Instrument für die Kritik an der Moderne* selbst. Indem mit Hilfe von Evaluation nicht nur die intendierten Wirkungen von Eingriffen erfasst werden, sondern auch ihre nicht-intendierten Folgen, liefert sie die empirische Basis für eine gesellschaftliche Selbstreflexion.

Ulrich Beck (1996: 19ff.) hat den Blick auf die Nebenfolgen der Moderne gelenkt und seinen Ansatz der „reflexiven Modernisierung“ der „einfachen Modernisierung“ gegenüberstellt: Deshalb sieht er nicht länger die Zweckrationalität als Motor des Gesellschaftswandels, sondern die Nebenfolgen in Form von Risiken, Gefahren, Individualisierung und Globalisierung (vgl. Beck 1996: 40 u. 55).

Evaluation unterstützt nicht einfach nur den Glauben an den Fortschritt, dann wäre sie nur ein technokratisches Instrument, sondern stellt den Fortschritt selbst gleichzeitig in Frage, indem sie auch die Nebenfolgen in den Blick nimmt. Dies bedeutet allerdings, dass sich Evaluation nicht auf simple Soll-Ist-Vergleiche reduzieren lassen darf, bei denen die gewünschten Ziele mit den realisierten Zuständen verglichen werden, sondern dass Evaluation die nicht-intendierten Folgen ins Zentrum ihrer Betrachtungen stellt.

---

<sup>1</sup> Vgl. z.B. einführend Degele und Dries 2005, Hill 2001

Evaluation muss eine „*totale*“ *Perspektive* einnehmen, einen umfassenden Wirkungsansatz verfolgen und bei der Auswahl seiner Bewertungskriterien auf die Zukunftsfähigkeit der implementierten Lösungen achten.

Aus diesen Überlegungen kann die Schlussfolgerung gezogen werden, dass Evaluation in unserer Gesellschaft noch nie so notwendig war, wie heute

## 2. Zweckbestimmungen

Aus den bisherigen Ausführungen wird deutlich, dass Evaluation (1.) im Dienste *gesellschaftlicher Aufklärung* betrieben werden kann. In diesem Falle geht es in erster Linie darum, politische Strategien, Programme und Maßnahmen mit dem Instrument der Evaluation dahingehend zu bewerten, ob sie einen Beitrag zur Lösung gesellschaftlicher Probleme leisten.

Indem sie Transparenz über die Ziele und die Wirkungen solcher Strategien und Maßnahmen schaffen, ermöglichen sie Bewertungen auf einer rationalen Grundlage. Indem z.B. offen gelegt wird, welche politischen Ziele erreicht und welche vernachlässigt werden, wer von solchen Maßnahmen profitiert und wer nicht, welche Probleme gelöst werden und welche Gefahren damit verbunden sind etc. können öffentliche Diskussionen ausgelöst werden.

Dabei muss Evaluation transparent machen, welche Beurteilungskriterien sie zugrunde gelegt hat, um sich nicht dem Vorwurf auszusetzen, nur die Perspektive der politischen Eliten und Entscheidungsträger verwendet zu haben.

Evaluationsergebnisse stellen stets bewertende (nicht Wert-) Urteile dar. Sie basieren auf systematisch erhobenen Daten, die anhand verschiedener Kriterien bewertet werden können, z.B. ob Maßnahmen dazu beigetragen haben, die soziale Gerechtigkeit oder die Bildungschancen benachteiligter sozialer und ethnischer Gruppen in einer Gesellschaft zu erhöhen, geschlechtsspezifische Nachteile abzubauen oder ressourcenschonende Lösungen umzusetzen. Durch die getrennte Darstellung von Evaluationsergebnissen, den vorgenommenen Bewertungen und der Offenlegung der dabei verwendeten Kriterien werden nicht nur die im Rahmen einer Evaluation vorgenommenen Bewertungen nachvollziehbar, sondern es bietet sich auch die Möglichkeit, bei der Verwendung anderer Bewertungskriterien zu anderen Beurteilungen zu gelangen. Indem Evaluation, mit Hilfe seiner Konzepte und Verfahren, öffentliches Handeln beobachtet, bewertet und transparent macht, kommt ihr eine ähnliche gesellschaftliche Funktion zu wie dem Journalismus.

Evaluation kann nicht nur Teil einer demokratischen Kontrolle sein, sondern auch (2.) Teil einer *demokratischen Regierungsführung*. Dabei werden Evaluationen sowohl von legislativen als auch exekutiven Einrichtungen genutzt, indem sie in Gesetzen und Verordnungen für bestimmte Zwecke festgeschrieben und von den ausführenden Organen umgesetzt werden müssen.

In welchem Umfang und wie Evaluationen durchgeführt werden, um Programme und Policies zu überprüfen, ist von Land zu Land und von Ministerium zu Ministerium verschieden und unterliegt einem Wandel im Zeitablauf. Wenn öffentliche Einrichtungen Evaluation selbst als Mittel einsetzen um zu belegen, dass sie ihre gesetzten Ziele erreichen (Effektivität), welche

Wirkungen (auch nicht-intendierte) ausgelöst wurden (Impact), wie sich das Verhältnis von Kosten zu Nutzen verhält (Effizienz) etc., kann die *Glaubwürdigkeit und Legitimität von Politik* gesteigert werden. Wenn sich nachvollziehbar begründen lässt, warum bestimmte Programme eingestellt, gekürzt oder ausgeweitet werden, steigt die Akzeptanz oder zumindest das Verständnis für Entscheidungen. Darüber hinaus ermöglicht das Wissen um Zusammenhänge und die durch Politik-Strategien ausgelösten Wirkungen auch eine rationale Politik. Politik kann auf einer breiteren Wissensgrundlage zum Nutzen des Allgemeinwohls betrieben werden.

Donald Campbell (1969) hat diesen Gedanken in seinem Konzept von der „experimentierenden Gesellschaft“ aufgegriffen, in dem eine Art „Arbeitsteilung“ zwischen Evaluation und politischer Entscheidungsfindung propagiert wird. Danach sollte das bei Evaluationen gewonnene rationale Wissen direkt in politische Entscheidungen überführt werden. Diese Form der Verknüpfung zwischen Evaluation und Politik ist als eine Reduktion von politischen auf technische Fragen heftig kritisiert und als „social engineering“ bezeichnet worden. Doch Evaluationsergebnisse lösen natürlich keinen Automatismus aus.)

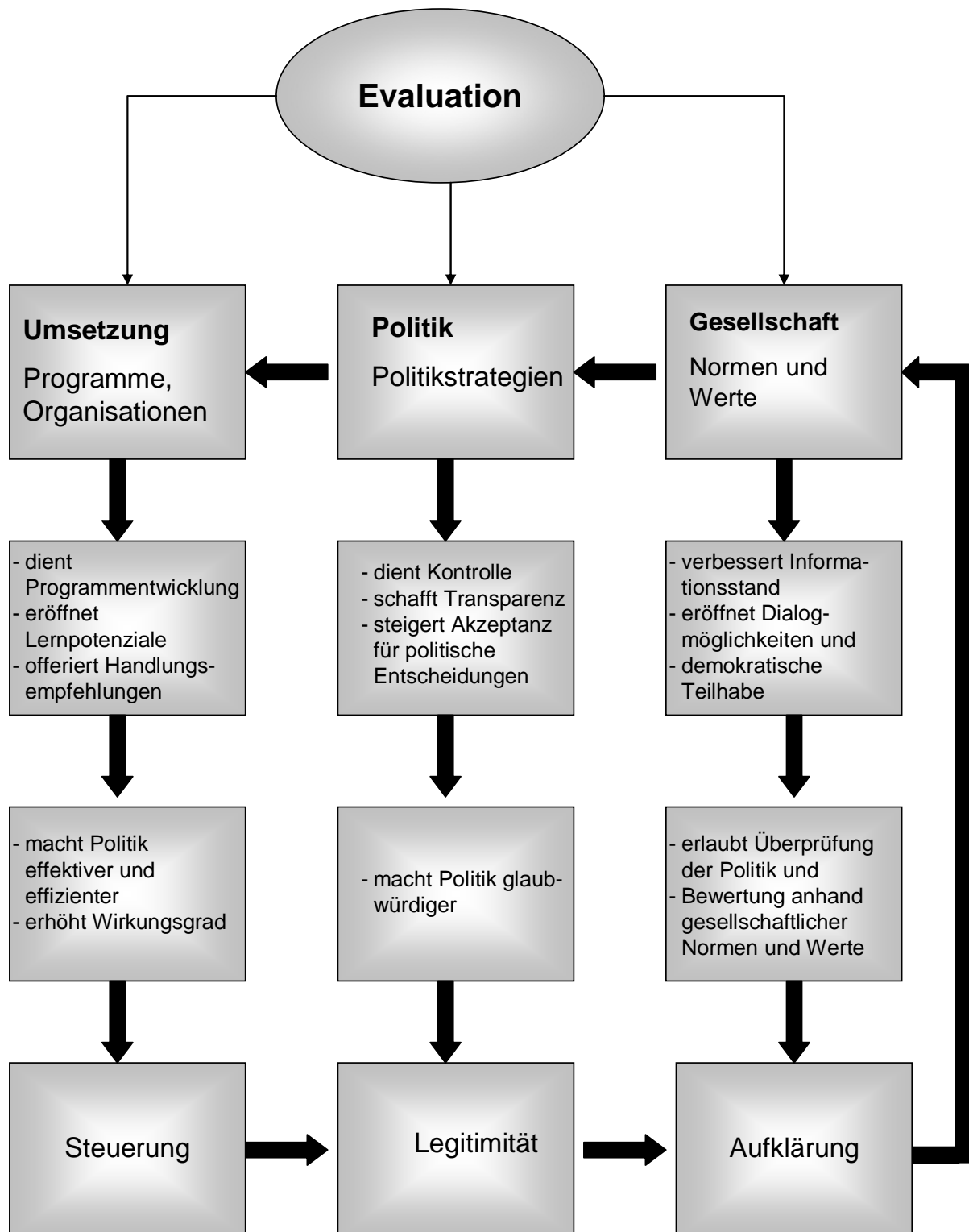
Evaluation kann nicht nur dazu beitragen, gesellschaftliche Aufklärung zu leisten und die demokratische Teilhabe und Regierungsführung zu stärken, sondern auch (3.) dazu, die *Steuerungsfähigkeit* von einzelnen Maßnahmen, Programmen oder gar ganzen Organisationen zu erhöhen. Insbesondere im Zuge der Einführung von Neuen Steuerungsmodellen im Rahmen von New Public Management und der Etablierung weit reichender Qualitätsmanagementmodelle avanciert Evaluation zu einem integralen Bestandteil der organisationalen Struktur und Kultur sowie der Ablaufprozesse in Organisationen. Da sich Evaluationen in organisationale „feed-back“-Schleifen integrieren lassen, fließt erworbenes Wissen, z.B. über den Ablauf und die Wirksamkeit von Programmen immer wieder in deren Steuerung ein.

Da sich Evaluation in allen Phasen des politischen Prozesses einsetzen lässt (in der Planungs-, Durchführungs- und Wirkungsphase), kann sie die Programmsteuerung durchgängig unterstützen. Evaluation dient somit der Programmentwicklung, kann auf Planungs- und Implementationsdefizite aufmerksam machen und dadurch Lernpotenziale erschließen.

Indem mit Hilfe der Evaluation auch Ursache-Wirkungs-Zusammenhänge analysiert werden, wächst ein Verständnis über Programmszusammenhänge und können Handlungsempfehlungen erarbeitet werden. Evaluation lässt sich zudem zur Verbesserung organisationaler Prozesse, zur Qualitätsentwicklung, zum Aufbau eines Wissensmanagements oder zum organisationalen Lernen einsetzen. Dadurch trägt Evaluation dazu bei, Politik effektiver und effizienter zu gestalten und damit den Wirkungsgrad von Politik zu erhöhen. Darüber hinaus ist die Bereitschaft und die Fähigkeit, Evaluation in die Management Strukturen einer Organisation zu integrieren mittlerweile zu einem Kennzeichen moderner Organisationen geworden

In Abbildung 1.1 sind die *drei zentralen Zweckbestimmungen von Evaluation* dargestellt. Dabei wird deutlich, dass die drei Einsatzfelder in einem engen Zusammenhang stehen.

**Abbildung 1.** Dreifache Zweckbestimmung der Evaluation



In der Literatur herrscht keineswegs Einmütigkeit über die *Vielfalt von Evaluationszwecken*. Eleanor Chelimsky (2006: 33) weist darauf hin, dass einige Evaluatoren davon ausgehen, „that evaluation is valuable only when it measures accountability, that is, when it maximizes

tax resources by holding policy makers and program managers accountable for the merit and worth of their policies and programs“. Andere halten Evaluation nur dann für nützlich, wenn sie Wissen (knowledge) generiert: „when it brings new or more profound understandings in some specific area of public endeavor“. Andere wiederum sehen den Wert von Evaluation vor allem in den programmdurchführenden Organisationen selbst: „for them, evaluation is valuable only when it improves institutions, moving them from performance measurement and self-evaluation to the learning organisation“.

Auch wenn diese drei Perspektiven häufig nicht exklusiv verstanden werden, herrscht oft wenig Verständnis für die Tätigkeit des anderen: „There have been veritable battles in the literature between warriors claiming, on the one hand, that ‚the only real evaluation is ultimately a fundamental judgement of merit or worth‘, and on the other, that there is a whole ‚menu of evaluation purposes‘ (...)“ (Chelimsky 2006: 34).

Hier wird die Auffassung vertreten, dass Evaluation keinesfalls nur einem Zweck dient, sondern einer Vielfalt von Zielstellungen, die hier unter den drei Aspekten *demokratische Aufklärung*, *Legitimitätsbeschaffung* für Politik und *Steuerung* von Politik (über Programme, Projekte und Maßnahmen) subsummiert werden. Diese enge Verbindung zwischen Evaluation und Politik ist nicht unproblematisch, wie noch zu zeigen sein wird.

### 3. Institutionalisation der Evaluation

Damit Evaluation als Aufklärungs-, Legitimitäts- und Steuerungsinstrument seine Funktionen entfalten kann, sind Evaluationskapazitäten in verschiedenen Bereichen notwendig:

(1.) Um seiner aufklärerischen Aufgabe gerecht zu werden, müssen Evaluationskapazitäten *in der Gesellschaft* aufgebaut werden, die möglichst unabhängig von Auftraggebern und mittelverwaltenden Stellen agieren können. In einigen Ländern (z.B. Schweden, Niederlande, Großbritannien, USA) haben General Accounting Offices solche Aufgaben übernommen. Der deutsche Bundesrechnungshof prüft nach Artikel 114 des Grundgesetzes „die Rechnung sowie die Wirtschaftlichkeit und Ordnungsmäßigkeit der Haushalts- und Wirtschaftsführung“. Umfassende Evaluationen, die die Wirksamkeit oder gar Nachhaltigkeit politischer Strategien untersuchen, gehören nicht zu seinem originären Aufgabenfeld. Mittlerweile sind in Deutschland eine Reihe unabhängiger Institutionen entstanden, die in bestimmten Politikfeldern tätig sind, z.B. das Institut für Schulqualität der Länder Berlin und Brandenburg<sup>2</sup> und das Institut für Forschungsinformation und Qualitätssicherung<sup>3</sup>. Auch die Idee, einen internationalen „Impact Evaluation Club“ der wichtigsten Geberländer in der Entwicklungszusammenarbeit zu gründen, um ihre Wirksamkeit zu überprüfen, ist hier einzuordnen<sup>4</sup>.

---

<sup>2</sup> <http://www.isq-bb.de>

<sup>3</sup> <http://www.forschungsinform.de>

<sup>4</sup> In Deutschland wurde diese Idee, ein Zentrum für die unabhängige Überprüfung der Wirksamkeit und der Nachhaltigkeit der Entwicklungszusammenarbeit zu gründen, erstmals vom Autor dieses Kapitels während einer Tagung zur „Nachhaltigkeit von Projekten und Programmen der Entwicklungszusammenarbeit“ im November 1992 an der Universität Mannheim vorgeschlagen. Vgl. auch dpa-Interview vom 2. Dezember 1994 mit dem Autor, in dem dieser Vorschlag wiederholt wird. Am 14. März 1996 haben der Bundestagsabgeordnete Dr. R. Werner Schuster und die SPD-Bundestagsfraktion einen Antrag zur „Systematischen Erfolgskontrolle“ gestellt, in dem



Damit Evaluation seine unabhängige Rolle im Rahmen moderner Demokratien erfüllen kann, sind solche Institutionen notwendig, die relativ frei von externen Einflüssen professionell die Wirkungen politischer Programme und Strategien untersuchen können, um die Öffentlichkeit mit fundierten Erkenntnissen zu versorgen. Dass diese Aufgabe nicht leicht ist, wird an der Aussage von Eleanor Chelimsky (2006: 53) deutlich: „Yet evaluation is a fragile reed to send up against all those giant oaks (...) and evaluators need to be ingenious, lucky, and much better protected than they currently are if they are to survive in any government“. Insbesondere wenn Evaluatoren unliebsame Ergebnisse zu verkünden haben, sind sie bei den Programmverantwortlichen oft wenig gut gelitten. Um so wichtiger ist es, dass solche Ergebnisse nicht in tiefen Schreibtischschubladen in der Versenkung verschwinden, sondern die Öffentlichkeit davon erfährt, damit politische Konsequenzen gezogen werden.

Evaluationskapazitäten sind auch auf der *Ebene der politischen Steuerung* notwendig, um die Umsetzung der eigenen Strategien und Politiken zu überprüfen und evaluativ zu begleiten. Dabei können Evaluationen – wie ausgeführt – sowohl (2.) der *Steigerung der politischen Legitimität* dienen, als auch (3.) der *Verbesserung des Steuerungspotenzials*, um die eigene Arbeit effizienter und effektiver zu gestalten. Dies ist nicht nur eine Herausforderung für den staatlichen Sektor sondern auch für den Nonprofit-Bereich insgesamt. Auch und gerade Nonprofit-Organisationen, die mit privaten Spendenmitteln arbeiten, sollten ein besonderes Interesse daran haben, ihre Steuerungspotenziale voll auszuschöpfen und anhand von Evaluationen zu belegen, dass sie eine wirkungsvolle Arbeit leisten und die Spenden bei ihnen gut angelegt sind (vgl. hierzu Stockmann 2006b, Kuhlmann u.a. 2004). Um diese Steuerungsaufgaben lösen zu können, sind sowohl *interne Evaluationskapazitäten*, also in den geldgebenden und durchführenden Organisationen, als auch *externe Evaluationskapazitäten*, in Form wissenschaftlicher Institute, privater Unternehmen und einzelner Gutachter (Experten) notwendig. Diese können dazu beitragen:

„To help agencies develop improved capabilities for policy and program planning, implementation and analysis of results, as well as greater openness and a more learning-oriented direction in their practice“ (Chelimsky 2006: 39).

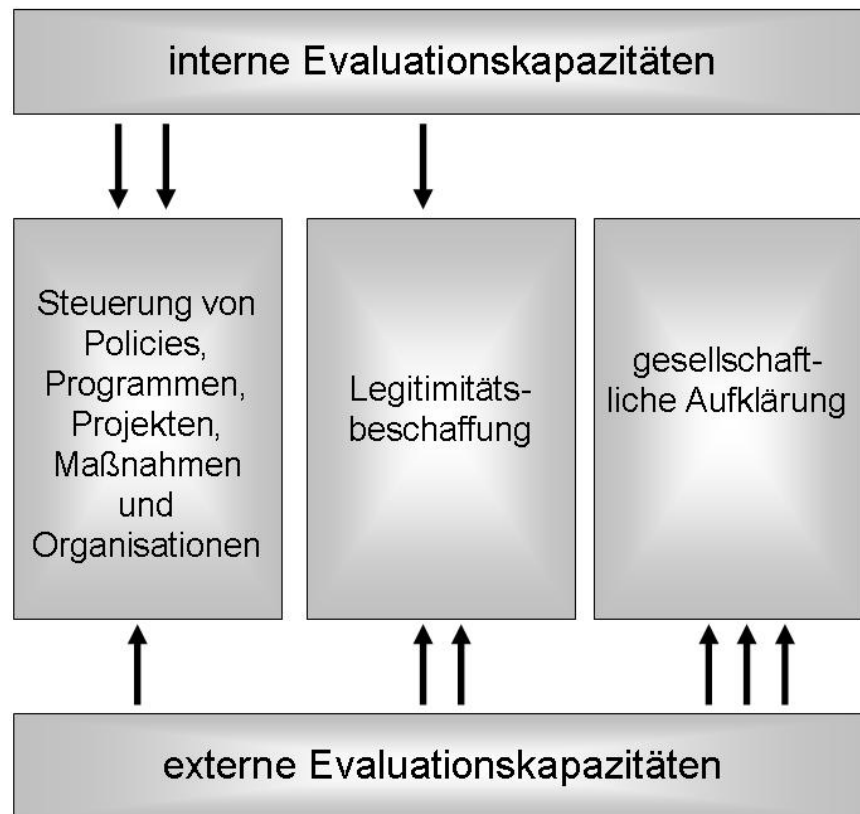
Andererseits sind Evaluationskapazitäten aber auch notwendig, um die geleistete Arbeit gegenüber den Geldgebern und der Öffentlichkeit zu legitimieren. Dabei werden externe Evaluationskapazitäten eher für die unabhängige Analyse im Dienste der gesellschaftlichen Aufklärung und der demokratischen Legitimierung genutzt, während sich für die Programmsteuerung sowohl interne als auch externe Evaluationskapazitäten einsetzen lassen. Es kann festgehalten werden, dass die *multifunktionale Verwendung* von Evaluationskapazitäten deutlich macht, dass es keine passgenaue Zuordnung zu den hier definierten Evaluationszwecken gibt. Externe Evaluationseinrichtungen können für alle drei Zwecke herangezogen werden. Dabei gilt, je unabhängiger sie sind, um so glaubwürdiger ist ihr Beitrag zur gesellschaftlichen Aufklärung, zur demokratischen Legitimitätsbeschaffung und zur Programmsteuerung. Interne Evaluationseinrichtungen tragen selten (Rechnungshöfe können es aber) zur gesellschaftlichen Aufklärung bei und werden wegen ihrer eingeschränkten

---

ebenfalls die Gründung eines Evaluationsinstituts gefordert wird (vgl. BT-Drucksache 13/4120). Allerdings wurden diese Vorschläge schon im Vorfeld der politischen Diskussion zu Fall gebracht, da die staatlichen wie privaten (gemeinnützigen) Nonprofit-Organisationen kein Interesse an einer unabhängigen Institution haben, die ihre Arbeit evaluieren könnte. Statt dessen wird auf eigene Evaluationen verwiesen, die teilweise oder auch ganz von externen Gutachtern durchgeführt werden, wobei die Frage nach der Unabhängigkeit dieser Gutachter immer wieder diskutiert wird.

Glaubwürdigkeit wenig zur Legitimitätsdarstellung programmdurchführender staatlicher oder nicht-staatlicher Organisationen genutzt, stattdessen dienen sie vor allem der internen Steuerung von Projekten und Programmen und manchmal auch von Policies, sowie der Gestaltung des organisationalen Qualitäts- und Wissensmanagements.

**Abbildung 2.** Zusammenhang zwischen Evaluationszwecken und -kapazitäten



## 4. Literatur

Beck, U. (1996): Weltrisikogesellschaft, Weltöffentlichkeit und globale Subpolitik. Ökologische Fragen im Bezugsrahmen fabrizierter Unsicherheiten. In: Diekmann, A.; Jaeger, C.C. (Hg.): Umweltsoziologie. Opladen: Westdeutscher Verlag.

Campbell, D.T. (1969): Reform as Experiments. In: American Psychologist. Bd. 24(4), S. 409-429.

Chelimsky, E. (2006): The Purpose of Evaluation in a Democratic Society. In: Shaw, I.F.; Greene, J.C.; Melvin, M. (Hg.): The Sage Handbook of Evaluation.

Degele, N.; Dries, C. (2005): Modernisierungstheorie. Eine Einführung. München: Fink (UTB).

Hill, H. (2001) (Hg.): Modernisierung - Prozesse oder Entwicklungsstrategie? Frankfurt: Campus.

Kuhlmann, S.; Bogumil, J.; Wollmann, H. (Hg.) (2004): Leistungsmessung und –vergleich in Politik und Verwaltung: Konzepte und Praxis (Stadtforschung aktuell Bd. 96). Wiesbaden: Verlag für Sozialwissenschaften.

Stockmann, Reinhard (2006): Evaluation und Qualitätsentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement. Band 5 der Reihe Sozialwissenschaftliche Evaluationsforschung. Münster: Waxmann.