

Evaluation in Wissenschaft und Gesellschaft¹

Helmut Kromrey

In der Beschäftigung mit dem Thema, auf das ich mich ohne langes Zögern gern eingelassen habe, ist es immer mehr angewachsen und immer komplexer geworden. Um dennoch den vorgegebenen Rahmen nicht zu sprengen, werde ich mich beschränken und vor allem von der steten Notwendigkeit des Bemühens der Evaluator(inn)en sprechen, den Standort „zwischen allen Stühlen“ nicht zum ständigen Wohnsitz zu nehmen. Dabei wende ich mich zunächst dem schwierigen Verhältnis der Evaluation zu ihrer „sozialen Herkunft“ - der Wissenschaft - zu, die sich durchaus nicht in allen ihren Teilen darüber einig ist, ob sie die Evaluation - oder allgemeiner: eine direkt anwendungsbezogene Forschung - überhaupt mögen soll. Danach kommt das zweite Stichwort in meiner Themenstellung, die Gesellschaft - speziell Politik und Planung - als Adressat von Anwendungsforschung zur Sprache. Und schließlich soll am Ende auch die Evaluation selbst nicht vergessen werden.

Doch zunächst ein paar *Vorbemerkungen* zum schon angedeuteten schwierigen Verhältnis der Evaluation sowohl zur Wissenschaft wie zur Gesellschaft.

In welchen größeren Rahmen gehört Evaluation?

Sie ist - natürlich - *bezogen* auf außerwissenschaftliche, auf "gesellschaftliche" Sachverhalte, wird aber *durchgeführt* mit dem Anspruch auf Wissenschaftlichkeit: sie ist also *praxis- und anwendungsbezogene Wissenschaft*. Und damit sind Probleme mit "der Gesellschaft" ebenso vorprogrammiert wie Probleme mit "der Wissenschaft".

Allerdings ist Evaluation nur ein *Spezialfall* anwendungsbezogener wissenschaftlicher Betätigung, und mit ihrer *direkten* Verwertungsorientierung steht sie in Konkurrenz zur *nur indirekt* anwendungsbezogenen problemorientierten, problemdiagnostizierenden Wissenschaft, die sich - ebenso wie "reine" Grundlagenforschung - auf so hochrangige Werte wie Objektivität und Wertneutralität verpflichtet fühlt, wogegen die direkt verwertungsbezogene Forschung sich auch auf "schmutzige" außerwissenschaftliche Werte einlassen muss.

Und dann gibt es da auch noch die Erkenntnistheoretiker und die Methodologen, die sich mit Vorliebe an der idealtypischen, an der "reinen", der nur auf die Wahrheit und nichts als die Wahrheit verpflichteten theoriebildenden und -testenden Wissenschaft orientieren und die daraus ihre Normen korrekten wissenschaftlichen Vorgehens ableiten. Deren strikte Anforderungen an Datenqualität, Messgenauigkeit, Repräsentativität und anspruchsvolle Analysemodelle aber werden von den direkt verwertungsorientierten Forschern chronisch verletzt, *müssen* chronisch verletzt werden. Darf es da verwundern, dass "richtige" Methodologen die empirische Evaluation eher als eine Spielwiese für wissenschaftliche "Schmuddelkinder" ansehen und sie daher möglichst meiden?

Warum dieser Ausflug in innerwissenschaftliche "Nestbeschmutzung"? Weil - wie zu zeigen sein wird - alles dies irgendwie von Bedeutung ist, wenn es um Evaluation geht und deshalb zumindest kurz gestreift werden muss.

¹ Überarbeiteter Text auf der Grundlage des Vortrags bei der Feierstunde zur Eröffnung des Centrums für Evaluation am 21.2.2003 in Saarbrücken

Und die Gesellschaft?

Auch die hat ihre Probleme – und nicht zu knapp! Deshalb fordert sie von Zeit zu Zeit Hilfe bei der Wissenschaft ein, und fühlt sich doch nicht immer so recht wohl dabei. Zugleich nämlich vernimmt man Klagen über die Gefahren zunehmender Verwissenschaftlichung des Alltags und darüber, dass wissenschaftliche Konstruktionen die soziale Welt aus einer verfremdenden praxisfernen Perspektive wissenschaftlicher Elfenbeintürme definieren und dass dadurch der Blick auf die "wirklichen sozialen Probleme" verstellt werden könnte - wie umgekehrt natürlich auch in der Wissenschaft davor gewarnt wird, dass die Wissenschaft bei zu starker Vereinnahmung durch die Politik und deren Informationsbedarf die „Thematisierungskompetenz“ aus der Hand gibt, so dass ihr am Ende nur noch die undankbare Aufgabe bleibt, „im Brunnen nach den dort bereits hineingefallenen Kindern zu suchen“, statt vorausschauend unerwünschte Entwicklungen verhindern zu helfen (so z.B. Häußermann/Siebel 1978, 493).

Sie sehen, es handelt sich in der Tat nicht um eine problemlose Beziehung, von der hier die Rede ist. Doch damit nicht genug. Schon in der Formulierung meines Themas "Evaluation in Wissenschaft und Gesellschaft" steckt eine Fußangel: "... *und* Gesellschaft"? Müßte es nicht richtiger heißen: "Wissenschaft *in der* Gesellschaft"?

Der Erfinder des Begriffs "Soziologie", Auguste Comte, hatte jedenfalls Mitte des 19. Jahrhunderts durchaus die Utopie einer "positiven" (gemeint: einer empirischer) Sozialwissenschaft, die den Maßstab für Entscheidungen darüber liefern könnte, was gesellschaftlich gut und was richtig ist. In diesem Falle bräuchte die Politik nicht mehr ein Kampfplatz widerstreitender Interessen zu sein, bräuchte sie nur noch die Rolle eines ausführenden Organs zu spielen. Zugespitzt: Wissenschaft als Legislative, Politik als Exekutive. Eine weltfremde Utopie? Immerhin wurde gut ein Jahrhundert später diese Utopie durchaus als Realmodell diskutiert: als technokratisches Modell der Politikberatung, auch als indikatorengestützte Planung und Politik. Auf Letzteres werde ich noch näher eingehen.

Ist also eine konzeptionelle Trennung der sozialen Welt in Gesellschaft *und* Wissenschaft gerechtfertigt? Darüber kann man mit Fug und Recht streiten. Für mein Thema jedenfalls ist diese Trennung *notwendig*. Und dabei befinde ich mich in guter Gesellschaft. So haben zum Beispiel am 27. Mai 1999 Vertreter des Stifterverbandes für die Deutsche Wissenschaft und führender Wissenschaftsorganisationen² ein Memorandum "Dialog Wissenschaft und Gesellschaft" unterzeichnet. Darin wurde sogar ein Aktionsprogramm zur Förderung von Initiativen begründet, die "die Verbesserung des Dialogs zwischen Wissenschaft und Gesellschaft zum Ziel haben".³

Auch das vertraute Bild vom "Elfenbeinturm Wissenschaft" ist ja nur aus der Idee eines autonomen, gesellschaftsunabhängigen Bereichs nachvollziehbar. Bezugnehmend auf dieses Bild haben in den letzten Jahren mehrere Tagungen das Thema aufgegriffen, wie dieser Elfenbeinturm wieder für die Öffentlichkeit begehbar gemacht werden könnte.⁴ Angestrebt wird eine Stärkung der Dialogfähigkeit der Wissenschaft mit anderen gesellschaftlichen Gruppen.

Wissenschaftler sollen:

- „sprachfähig sein, aber auch zuhören können,
- Bürgernähe auf Augenhöhe suchen, aber auch zu politischer Einschätzung und Beratung fähig sein,
- sich den Medien öffnen, aber auch den Umgang mit Medienschetle souverän beherrschen,

² Deutsche Forschungsgemeinschaft, Max-Planck-Gesellschaft, Hochschulrektorenkonferenz, Hermann von Helmholtz-Gemeinschaft, Fraunhofer-Gesellschaft, Wissenschaftsrat und die Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz.

³ Das Memorandum ist dokumentiert unter: www.berlinews.de/archiv/428.shtml.

⁴ organisiert z.B. vom Wissenschaftszentrum Berlin oder von der Robert-Bosch-Stiftung.

- ihre Arbeit und Ergebnisse offen und transparent darstellen, was z.B. auch die Offenlegung ökonomischer und anderer (Eigen-)Interessen beinhaltet.⁵

Mit „Wissenschaft“ ist hierbei zwar in erster Linie die disziplinentorientierte Grundlagenwissenschaft gemeint. Thematisiert wird damit aber ein grundlegendes Problem auch der Evaluation: die Experten-/Laienkommunikation. Auch Evaluation muss ja, um ihren Zweck erfüllen, um nutzbar werden zu können, zunächst einmal von *allen* Beteiligten und Betroffenen verstanden werden. Wegen ihres hohen Spezialisierungsgrades haben die Wissenschaften in ihren Teilgebieten jeweils eigene Sprachen entwickelt, die in der Regel für Nichtwissenschaftler nicht nur die wissenschaftlichen Inhalte undurchschaubar werden lassen, sondern auch ihre Methoden und Verfahren schwer zugänglich machen. Neben allen wissenschaftlichen Qualifikationen müssen daher kompetente Evaluatoren ganz offensichtlich zugleich auch alle oben genannten kommunikativen Qualifikationen aufweisen.

Die Inhomogenität des Systems Wissenschaft

Wie schon zu Beginn kurz angesprochen, ist es in unserem Zusammenhang zu kurz gegriffen, von "der Wissenschaft" zu sprechen. Die Wissenschaft – auch wenn man sie als autonomen, aus der Gesellschaft ausdifferenzierten Bereich betrachtet – ist alles andere als homogen. Das gilt nicht nur für die schon vorgenommene Einteilung nach dem Grad von Grundlagen- oder Anwendungsorientierung sowie deren durchaus unterschiedliche Wertbasis (Stichworte etwa: Wertneutralität versus Wertsensibilität versus parteiergreifende Wissenschaft wie z.B. Aktionsforschung). Mindestens ebenso bedeutsam ist die Trennung in unterschiedliche erkenntnistheoretische und methodologische Schulen. Es macht schon etwas aus, ob man sich zum erkenntnistheoretischen Konstruktivismus oder zum erkenntnistheoretischen Realismus bekennt, ob man forschungsmethodisch standardisierende (nach Möglichkeit quantifizierende) oder eher situationsflexible, qualitative Strategien bevorzugt. Evaluatoren sind auch hier wieder mit Vorliebe "zwischen den Stühlen" zu finden: Sie haben es eigentlich durchweg mit sozialen Konstruktionen zu tun, bekennen sich aber überwiegend zum erkenntnistheoretischen Realismus. Und methodisch ist für Evaluationen von ihrem Gegenstand her meist ein Methoden-Mix, eine Mischung aus qualitativ und quantitativ angemessen, wogegen vom Adressaten ihrer Informationen oft den quantitativen Daten der höhere Grad an Glaubwürdigkeit zugemessen wird.

I.

Disziplinentorientierte versus anwendungsorientierte Sozialwissenschaft?

Kommen wir nun zur „angewandten Sozialforschung“ und damit einen deutlichen Schritt näher an das Thema „Evaluation“; denn: alles was für diesen Bereich wissenschaftlicher Betätigung gilt, das gilt auch für die Evaluationsforschung als deren Spezialfall.

Von angewandter Sozialforschung wird allgemein gesprochen, wenn ein erfahrungswissenschaftliches Untersuchungsvorhaben auf die unmittelbare Lösung praktischer, gesellschaftlich-politischer Probleme hin angelegt ist (Habermehl 1992). Praxisfelder finden sich in so unterschiedlichen Bereichen wie Politikberatung, Konsumforschung und Marketing, kirchliche Gemeindefarbeit, Hochschulentwicklungsprogramme und bei der Umsetzung sozialpolitischer Maßnahmen. Es geht darum, gezielt und systematisch Grundlagen für außerwissenschaftliche Entscheidungsprozesse bereitzustellen. (Clemens 2000, 215)

⁵ www.bosch-stiftung.de/a52_pbl_sub01.htm.

Als *zentrale Funktionen* disziplinorientierter Sozialforschung gelten: „Aufklärung“ (enlightenment) durch Veröffentlichung, Wissens*vermehrung*, Erklärung, Verallgemeinerung. Praxisorientierte Sozialforschung dagegen ist geprägt durch Aufgaben wie Wissens*übertragung* (Transfer) durch Beratung, durch Handlungsorientierung, Beeinflussung, Situationsbezogenheit.

Der *Einfluß auf „Praxis“ und „Politik“* geschieht bei disziplinorientierter Sozialforschung indirekt – Revision veralteter Auffassungen, Neuformulierung von Zielen und Prioritäten –, bei praxisorientierter Sozialforschung direkt – Erweiterung und Konkretisierung der Informationsbasis über Beteiligte und Betroffene („stakeholder“), maßnahmenbezogene Zielformulierung, Maßnahmenentwicklung und – implementation, Erfolgskontrolle.

Unterschiede finden wir auch bei der Art der Reaktionen auf die Forschung. Bei disziplinorientierter Sozialwissenschaft existiert kein direktes Feedback (außer bei Rezensionen von Veröffentlichungen), es geht ihr um die verwertungsneutrale Erweiterung des Wissensbestandes. Praxisorientierte Sozialforschung muss dagegen mit direktem Feedback durch ihre Nutzer rechnen; schließlich hat sie ja *bedarfsorientiert* Informationen bereit zu stellen und aufzubereiten.

Wer schon Evaluation betrieben hat, weiß, wieviel Zündstoff in jedem dieser für die angewandte Forschung aufgeführten Punkte steckt.⁶ Umso höhere Anforderungen an Fach- wie an Sozialkompetenz muss der Anwendungsforscher erfüllen, um in diesem „verminten“ Arbeitsfeld bestehen zu können!

Was macht Anwendungsforschung – und damit auch Evaluation – methodisch so besonders anspruchsvoll?

Von der Grundlagenwissenschaft unterscheidet sich die angewandte Sozialforschung durch eine Reihe spezifischer Schwierigkeiten.

Begründet liegt dies u.a. in einem zentralen Charakteristikum des „Gegenstands“ ihrer Forschung: der *Komplexität sozialer Sachverhalte*. Grundlagenforschung mit Zielen wie Theorietest oder Theorieentwicklung löst das Komplexitätsproblem durch konsequente *Vereinfachung der Untersuchungssituation* – bis hin zur Konstruktion völlig realitätsferner Laborsituationen, um den Einfluss einzelner Wirkfaktoren zu isolieren. Im Rahmen *anwendungsorientierter Forschung* verbietet sich ein solches Vorgehen. *Reale soziale Sachverhalte* sind kein der Vereinfachung zugängliches Experimentierfeld; sie sind und bleiben komplexe Realität. Eine *angemessene Beschreibung und Diagnose* kann in diesem Rahmen *immer auch nur komplex* ausfallen – oder sie ist zwangsläufig falsch und führt ebenso zwangsläufig zu falschen Schlussfolgerungen.

Eine Privilegierung der Grundlagenwissenschaft ist darüber hinaus deren erkenntnistheoretisch gewollte *Autonomie bei der Wahl ihres „Gegenstands“*. Sie kann (bis zu einem gewissen Grad) ihre Themen unter wissenschaftlichen Gesichtspunkten wählen und deren Relevanz unter Hinweis auf angestrebte „reine“ Erkenntnis begründen. Sie kann und soll sich bei allen zu treffenden Entscheidungen an wissenschafts-

⁶ Zwischen diesen beiden Polen – Grundlagenwissen / Anwendungspraxis – darf im übrigen die *„problemorientierte Sozialforschung“* nicht vergessen werden (ihre Ziele: Diagnose einer "problematisch" gewordenen sozialen Situation und ihres Zustandekommens, Entwicklung und/oder Präzisierung sozialwissenschaftlicher Theorien für diesen Problembereich). Wenn mit Recht von der Gefahr gesprochen werden kann, Wissenschaft könne als anwendungsorientierte Disziplin die "Thematisierungskompetenz" aus der Hand geben, dann gilt dies eigentlich eher für dieses Forschungsfeld als für die direkt anwendungsbezogene Evaluation. Es gilt zunehmend in dem Maße, wie Forschung auf Drittmittel angewiesen ist; denn dann können Förderungsanträge mit Aussicht auf Erfolg nur noch gestellt werden, wenn sie eine der Problemdefinitionen aufgreifen, die die geldgebenden Institutionen vorgegeben haben.

immanenten Normen – und nur an diesen – orientieren (insbesondere Erkenntnisfortschritt als Selbstzweck, Gültigkeit, Zuverlässigkeit, Generalisierbarkeit der Resultate); außerwissenschaftliche Interessen werden in den Entdeckungs- und Verwertungskontext ausgegliedert. Das ist bequem! Diese Entlastung von gesellschaftlicher Verwertungs- und Nützlichkeitsargumentation ist wesentlicher Bestandteil der wissenschaftlichen Legitimation.

Anwendungsorientierte Forschung muss dagegen nicht nur die Definition und Präzisierung ihrer Fragestellungen an außerwissenschaftlichen Erkenntnisinteressen und Verwertungskontexten ausrichten: dies ist ihr zentrales Definitions- und Legitimationskriterium. Vielmehr darf auch bei der methodologischen – also empirisch-wissenschaftlichen – Begründung jeder Detailentscheidung innerhalb des Forschungsablaufs die unmittelbare Verwertbarkeit der erzielbaren Resultate nicht aus dem Blick geraten. Der „Elfenbeinturm der Wissenschaft“ ist für diese Forschungsrichtung kein möglicher Zufluchtsort. Ganz im Gegenteil werden wissenschaftsexterne Zweckmäßigkeitsüberlegungen aus dem Verwertungskontext „internalisiert“ und geraten so auf die gleiche Stufe wie die methodologischen Normen der Gültigkeit und Zuverlässigkeit. Dennoch dürfen auch letztere nicht vernachlässigt werden, denn falsche – also empirisch nicht haltbare – Informationen sind genauso wertlos, wie dies pragmatisch unbrauchbare Resultate wären. Dieser Spagat ist schwierig!

Noch ein Privileg der Grundlagenforschung: Sie kann – zumindest vom Prinzip her – unter *Entlastung von aktuellem Zeitdruck* arbeiten. Neue Erkenntnisse – „Entdeckungen“ – lassen sich schließlich nicht mit dem Terminkalender erzwingen. Benötigt ein Forschungsprogramm mehr Zeit, dann muß diese eingeräumt werden. Und falls dennoch das vorgesehene Arbeitsprogramm nicht geschafft wurde und bei zeitlich befristeten Projektressourcen nicht eingelöste Desiderate unter „weiterer Forschungsbedarf“ ausgewiesen werden, dann entwertet dies jedenfalls nicht das schon Erreichte. Anders bei anwendungsorientierter Forschung. Hier haben die Ergebnisse im zuvor definierten Umfang zeitpunktgenau vorzuliegen. Verspätet vorgelegte Befunde sind als vorgesehene Grundlage von Beratungen oder zur Absicherung von Entscheidungen nicht mehr verwertbar, sobald der Entscheidungstermin verstrichen ist.

Last, but not least: *Grundlagenforschung darf sich „irren“*. Damit ist gemeint: Hypothesen, die als Ausgangspunkt gewählt werden, dürfen sich im Verlauf der Forschung als falsch erweisen. Es gilt sogar als besonders wünschenswert, die Suche nach neuer Erkenntnis von „gewagten“ Annahmen aus aufzunehmen. Deren informationsreiches Scheitern ist nicht selten der Startpunkt für grundlegende Erkenntnisse, die eine neue Forschungslinie begründen. Diesen Luxus darf sich angewandte Sozialwissenschaft nicht leisten. Bei der Konzipierung des Forschungsdesigns ist große Sorgfalt darauf zu verwenden, dass die zugrunde liegenden Annahmen und Hypothesen einen hohen Grad empirischer Bewährung aufweisen und dass der Prozeß der Gewinnung, Auswertung und Interpretation aller Informationen methodisch abgesichert und mit begleitender Qualitätskontrolle verläuft. Jede falsche Schlußfolgerung im Verwertungskontext, die wegen fehlerhafter Forschungsdaten gezogen wird, hat Konsequenzen für einen nicht absehbaren Kreis von Betroffenen. Absolut korrekte Forschung und ein hohes Maß an Verantwortungsbewusstsein ist in solchem Kontext allein schon eine Forderung der Wissenschaftsethik.

Anwendungsforschung – ein blinder Fleck in der universitären Ausbildung

Es dürfte sich bereits nach diesen – wenigen ausgewählten – Argumenten nahezu von selbst verstehen, dass für angewandte Forschung ein hoher Standard an Kompetenz und Professionalität – methodische und theoretische ebenso wie praktische – benötigt wird (Clemens 2000). Sie eignet sich – anders als universitäre Disziplinforschung – wenig zur „Spielwiese“, um zunächst noch „das Forschen zu lernen“. Werden in Projekten mit der Ziel der Gewinnung zweckfreier Erkenntnis Fehler begangen, dann geschieht dies ohne gravierende Nachteile für Dritte (allenfalls werden knappe Ressourcen verschwendet), und selbst

dann können die Beteiligten daraus immerhin noch Erfahrungen gewinnen, die für Nachfolgeprojekte nutzbar sind.

Entgegen dieser Einschätzung können wir in der universitären Forschungsbildung eine überraschende Diskrepanz beobachten: Die Mehrzahl der faktisch durchgeführten Forschungen ist anwendungsorientierte Forschung (zumindest wenn man seinen Blick nicht auf den Elfenbeinturm der „rein universitären Forschung“ beschränkt). Dennoch orientiert sich die Methodologie der Forschung ebenso wie die Ausbildung der Studierenden nahezu ausschließlich an der disziplinentestenden Grundlagenforschung, und dies mit besonderer Vorliebe an der theorienehestenden Grundlagenforschung. Das Angebot an anwendungsorientierter Weiterbildung, wie es vom CEval jetzt projiziert wird, stößt also nicht lediglich in eine Marktlücke, es ist vielmehr sachlich zwingend geboten. Die explizite Professionalisierung der Anwendungsforschung, ihre Ausstattung mit einem auf ihre besonderen Notwendigkeiten zugeschnittenen epistemologischen Kanon und mit einem ergänzenden Katalog von Methoden und Qualifikationen könnte durchaus ein lohnendes Ziel sein, das sich das CEval „auf seine Fahnen schreibt“. Angewandte Sozialforschung ist weit mehr als die bloße Anwendung sozialwissenschaftlicher Disziplinärforschung (van de Vall 1993 spricht hier von einem „professionellen Paradigma“ mit spezifischen epistemologischen, implementären und strategischen „Parametern“). Auch in anderen Disziplinen (z.B. Medizin, Ingenieurwesen, Psychologie) haben Wissenschaft und Praxis getrennte methodologische, theoretische und normative Paradigmen entwickelt (s. auch Eberlein/Obermeier 1985).

II.

Der Informationsbedarf von Politik und Planung

Ich komme zur nächsten Perspektive in meiner Argumentation: Steuerung der Gesellschaft durch Planung und Erfolgskontrolle. Zunächst auch hier ein paar Vorüberlegungen.

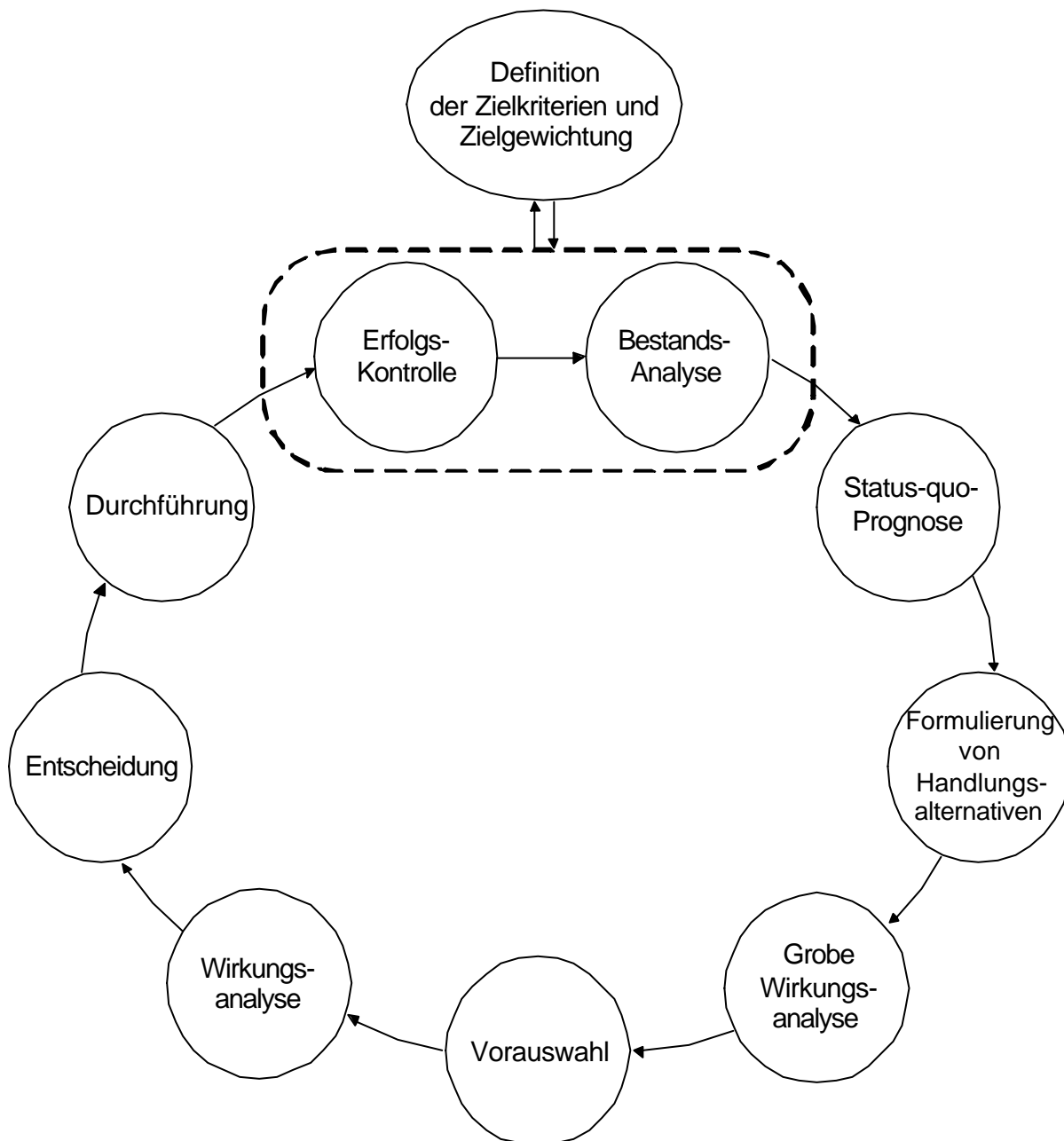
„Zielorientiertes politisches Handeln braucht Planung, Planung braucht Information“ – so kann man eine Maxime der politischen und sozialwissenschaftlichen Diskussion in den 1970er Jahren zusammenfassen. Planung wurde als „anthropologisches Grundmuster“ auch für individuelle Akteure entdeckt, geplantes Handeln als Handeln auf höherer Rationalitätsstufe von rein reaktivem Verhalten abgegrenzt und zum konstituierenden Merkmal für *menschliche* Tätigkeit per se erhoben (s. Werner 1975, Kap. 1). Planung ist danach zu verstehen als (noch) nicht Handeln, sondern *probeweises* Umgehen mit einer symbolischen Repräsentation der Realität, als (vorwegnehmende) gedankliche Manipulation der Realität auf der Ebene symbolischer Repräsentation und (informationsgestützter) Entscheidung zwischen Alternativen. Existiert nun eine Vielzahl planend handelnder Akteure (Gesellschaft), ergibt sich zwangsläufig die Notwendigkeit von Koordination. Dies geschieht entweder (im Marktmodell) direkt als Prozess-Steuerung: Jedes Individuum ist zugleich planender und handelnder Akteur, Koordination und Abstimmung erfolgen durch den „Markt“. Oder die Koordination erfolgt (im Planungsmodell) stellvertretend für die individuellen Akteure durch dazu legitimierte Repräsentanten.

Bei dieser Ausweitung des Konzepts von Probehandeln auf die gesellschaftliche Ebene wird die Identität der Akteure aufgegeben. Sie differenzieren sich aus in Planungssubjekt, Planungsobjekt und Aktionssubjekt. Als Konsequenz ergeben sich mehrere Probleme:

- Legitimierung der Planungs- und Aktionssubjekte (verbunden mit der Forderung nach Transparenz und öffentlicher Kontrolle);
- Problem der Zieldefinition (Forderung nach Partizipation zur Berücksichtigung der „Bedürfnisse der Betroffenen“, nach Transparenz der Entscheidungsprozesse);
- Problem der Alternativenbewertung bei konfligierenden Zielen und Bedürfnissen.

„Information“ wird jetzt zur zentralen Politik-Ressource; genauer: empirische Information in Form von (gesellschaftlichen) „Indikatoren“ als Instrument symbolischer (= datenmäßiger) Repräsentation planungs- und handlungsrelevanter Realität. Es versteht sich von selbst, dass in einem Konzept informationsgestützter „Steuerung durch Planung“ die Informationslieferanten, also die empirischen Sozialwissenschaften, erheblich an Bedeutung gewinnen. Ihren Ausdruck fanden solche Vorstellungen in Politikmodellen wie dem folgenden:

Abbildung 1: 10 Planungsschritte



Schema des kontinuierlichen Planungsprozesses
aus: J. Eekhoff, R. Muthmann u.a.: Methoden und Möglichkeiten der Erfolgskontrolle, Bonn-Bad Godesberg 1977, S. 18

Die hier vorgenommene Einteilung in zehn Planungsschritte ist natürlich willkürlich, weil einige nicht klar abzugrenzen sind und andere weiter unterteilt werden können. Sie verdeutlichen jedoch die Grundzüge des seinerzeitigen Denkens, das einen kontinuierlichen, linear verlaufenden Politikprozess unterstellte, beginnend

mit der grundlegenden Zielbestimmung über die empirische Feststellung von Handlungsbedarf, die Entwicklung von gedanklichen Handlungskonzepten bis zu deren Implementation und nachträglicher Erfolgskontrolle. Ich zitiere dazu den Autor des obigen Modells:

- „(1) Die Definition der Zielkriterien ist Voraussetzung für alle folgenden Planungsschritte. ...
- (2) In der Bestandsanalyse werden die Werte der Zielkriterien, d.h. die gegenwärtig bestehenden Zielerreichungsgrade, ermittelt, und es wird untersucht, wie es zur augenblicklichen Situation gekommen ist. ...
- (3) In der Status-quo-Prognose werden die Zielerreichungsgrade prognostiziert, die sich vermutlich ergeben, wenn keine zusätzlichen Maßnahmen ergriffen werden.
- (4) Auf dieser Planungsstufe werden die Handlungsmöglichkeiten ... durch eine Vielzahl von Handlungsalternativen dargestellt, wobei eine erste Vorauswahl durch die Planer nicht zu umgehen ist. ...
- (5) Durch eine grobe Wirkungsanalyse werden die wichtigsten Zielerträge der Maßnahmen geschätzt, um die Vor- und Nachteile ... ungefähr gegenüberstellen zu können.
- (6) Da nicht alle grob skizzierten Handlungsalternativen genau untersucht und detailliert geplant werden können, muß ihre Anzahl durch die Vorauswahl eingeschränkt werden. ...
- (7) Die in die engere Wahl genommenen Handlungsalternativen werden einer sorgfältigen Wirkungsanalyse unterzogen. Nach dem With-and-without-Prinzip werden die Abweichungen der Zielerreichungsgrade von der Status-quo-Situation prognostiziert und in einer Zielertragsmatrix zusammengestellt. Sowohl die Status-quo-Prognose als auch die Wirkungsprognose werden später in der Erfolgskontrolle überprüft.
- (8) Die Politiker entscheiden über das durchzuführende Handlungsprogramm, indem sie aufgrund der erwarteten Zielwirkungen die optimalen Maßnahmen direkt auswählen oder indem sie nach bestimmten Entscheidungsverfahren ihre Zielgewichte mitteilen und damit indirekt das Handlungsprogramm bestimmen.
- (9) Erst mit der Durchführung der beschlossenen Maßnahmen wird die angestrebte Veränderung der Zielerreichungsgrade in Gang gesetzt. ...
- (10) Die Erfolgskontrolle setzt für neue Maßnahmen praktisch parallel zur Durchführung ein. Sie geht aber über die Kontrolle der in einem Planungszyklus beschlossenen Maßnahmen hinaus und erfaßt auch die Zielwirkungen der exogenen Faktoren und der sonstigen Aktivitäten...“ (Eekhoff u.a. 1977, 17)

In diesem Modell wird – wie leicht zu erkennen – nahezu idealtypisch das Denkschema einer "verwissenschaftlichten Praxis" aufgegriffen und realisiert. Würde es funktionieren, wäre die Utopie von Comte Realität geworden.

Die Rolle der Wissenschaft: Beitrag zur (angestrebten) Planungsrationaltät

Die Funktion der Wissenschaft in diesem Politikmodell beginnt bei der Definition von Zielen, Zielgewichtungen und Zielerreichungskriterien. Entwickelt werden „komplette“, d.h. alle Politikfelder umfassende, bis in Feinziele gegliederte und operationalisierte Zielsysteme, die sich am gesellschaftlichen Oberziel „Lebensqualität“ orientieren. Aufwändige Forschungsprojekte, z.B. mit weltweiter Vorbildfunktion in Japan entwickelt, von der OECD finanziert, in Deutschland im Rahmen eines Sonderforschungsbereichs an den Universitäten Frankfurt und Mannheim (SPES - Sozialpolitisches Entscheidungs- und Indikatorensystem) vorangetrieben, entwerfen und begründen umfassende Zielkataloge. In ihnen kann die Politik ablesen, was als normativer Konsens gelten kann, was „gewollt werden soll“. Die Politik würde – bezöge sie sich darauf – aus dem Tauziehen divergierender Interessen entlassen.

Ob politischer Handlungsbedarf besteht, ergäbe sich ebenfalls nicht daraus, welche Interessengruppen die stärkste Durchsetzungskraft haben. Besteht die begründete Vermutung, dass in einem Politikfeld die Ist- von der Sollsituation abweicht, liefert eine zielbezogene Bestandsanalyse die gewünschte Auskunft (im Idealfall muss diese Bestandsanalyse nicht einmal ad hoc durchgeführt werden, sondern es existieren auf die

ressortspezifischen Zielkataloge abgestimmte Systeme empirischer Daten: „soziale Indikatoren“). Weisen zudem Status-quo-Prognosen (ebenfalls basierend auf der – hypothetischen – Datenbank mit Indikatoren-Zeitreihen) darauf hin, dass die autonome, ohne politische Eingriffe zu erwartende Entwicklung nicht in der gewünschten Richtung verläuft, dann ist die Entscheidung gefallen: Es besteht Handlungsbedarf.

In diesem Fall entwerfen die Planer mehrere Programmalternativen, deren vermutliche Wirksamkeit unter Rückgriff auf wissenschaftlich bereitgestellte Entscheidungs- und Beurteilungsmodelle abgeschätzt und aus denen eine Vorauswahl getroffen wird (auch hier wird nicht "ideologisch", sondern „sachrational“ begründet entschieden). Danach folgt eine ins Detail gehende Wirkungsanalyse und -prognose (unter Anwendung z.B. des methodologischen Modells der Technikfolgenabschätzung) und schließlich die „objektive“ Entscheidung zur Implementation der best-geeigneten Alternative.

Nach der Implementation sowie einer programmspezifisch angemessenen Durchführungsphase folgt abschließend die empirische Erfolgskontrolle (unter Einsatz u.a. des Instrumentariums sozialwissenschaftlicher Wirkungsforschung). Diese liefert die empirischen Informationen zur Beurteilung der Eignung der realisierten Politik und ist ggf. – falls eine hinreichende Annäherung an die Zielsituation (noch) nicht erreicht werden konnte – als aktualisierte Bestandsanalyse der Ausgangspunkt des nächsten Planungszyklus.

Das „Versagen“ der Wissenschaft

Wie wir alle wissen, funktionierte dieses Modell nicht. Die von der Politik vereinnahmte Wissenschaft hat auf der ganzen Linie „versagt“:

- Die Zielkataloge gaukelten ein in sich stimmiges und widerspruchsfreies gesamtgesellschaftliches Zielsystem vor, das in dieser Form keine Entsprechung in der Realität hatte. Der Regelfall ist in der gesellschaftlichen Realität gerade nicht der sachrationale Konsens über politische Ziele; vielmehr existieren miteinander wettstreitende *alternative* Zielsysteme. Diese sind darüber hinaus niemals völlig in sich widerspruchsfrei; es bestehen auch *innerhalb* der jeweiligen normativen Systeme Zielkonkurrenzen und Zielkonflikte.
- Die Zielkataloge konnten Wertewandel nicht antizipieren. Festgeschriebene Zielsysteme selbst in den Bereichen, in denen sie zunächst auf weitgehendem Konsens beruhten, waren plötzlich nicht länger konsensfähig.
- Die soziale Realität erwies sich als zu komplex, als dass sie in hinreichend differenzierten Systemen sozialer Indikatoren für alle Politikfelder abgebildet werden könnte. Selbst im Falle ihrer (teilweisen) Realisierbarkeit bildeten sie im Idealfall immer „nur“ das jeweils zugrunde liegende Zielsystem ab. Bei dessen Wandel aber wurden sie für den ihnen zugeschriebenen Zweck (u.a. „Frühwarnfunktion“ für das Erkennen von politischem Handlungsbedarf) unbrauchbar.
- Als eine prinzipiell unlösbare Aufgabe stellte sich die Prognoseproblematik heraus. Prognosen wurden von dem Modell kontinuierlicher Planung an drei Stellen eingefordert: als Status-quo-Prognose bei der Feststellung von Handlungsbedarf, als grobe Wirkungsabschätzung für die Vorauswahl zwischen den Planungsalternativskizzen sowie als aufwändige Technikfolgenabschätzung vor der endgültigen Entscheidung für die Implementation des best-geeigneten Programms.
- Schon die Status-quo-Prognoserechnungen waren in zweierlei Hinsicht fehleranfällig: Jedes Rechenmodell (und sei es auch noch so komplex) ist notwendigerweise eine drastische Reduktion der wesentlich komplexeren sozialen Realität. Aber das Fehlen schon einer einzigen, scheinbar irrelevanten Variablen kann aufgrund nicht erkannter Wechselwirkungen mit anderen Variablen zu erheblichen Fehlprognosen führen. Außerdem muss die Status-quo-Prognoserechnung die Rahmenbedingungen, in

die das zu prognostizierende Problemfeld eingebettet ist, als in Struktur und Ausprägung stabil bleibend unterstellen (c.p.-Bedingung), oder deren Veränderung muss korrekt anzipiert werden. Jede nicht antizipierte Veränderung des Umweltrahmens aber macht die Aussagen der Modellrechnung als „objektive“ Entscheidungsgrundlage wertlos.

- Erst recht gelten die genannten Probleme bei Wirkungsanalysen für hypothetische Realitäten, die sich in Abhängigkeit von den einzelnen Planungsalternativen einstellen würden (beim Planungsschritt ‚grobe Wirkungsabschätzung‘ ebenso wie bei der differenzierten Folgenabschätzung).

Nicht zuletzt wurde in diesem Planungsmodell auch noch die Implementationsphase fälschlicherweise als relativ unproblematisch angesehen. Als implizite Prämisse lässt sich rekonstruieren: ‚Was beschlossen ist, wird von einer rational handelnden Verwaltung exakt in der vorgesehenen Weise ausgeführt.‘ Diese Fehlannahme allerdings war nicht so sehr ein Versagen der Wissenschaft als ein Denkfehler im zugrunde liegenden Politikmodell, das hier von dem „radikal rationalistischen“ (Vedung) Wunschdenken geleitet war, die Exekutive sei tatsächlich lediglich ausführende Instanz ohne eigene Politikziele: „In meiner Interpretation betont der radikale Rationalismus die zukunftsorientierten, entscheidungsfindenden Phasen im öffentlichen Interventionsprozeß. Sein wesentliches Merkmal war die kluge und synoptische Gestaltung der Programme durch eine zentrale Planungsmaschinerie vor ihrer Umsetzung; Ausführung und tatsächliche Zielerfüllung wurden als relativ unproblematisch erachtet“ (Vedung 1999, 24).

Die Alternative: Experimentelle Politik

Die Alternative zum Modell „kontinuierlicher Planungsprozess“ musste jedoch nicht lange gesucht werden; sie wurde nahezu parallel praktiziert: Während die Routinepolitik in den einzelnen Ressorts weiterhin indikatoren gestützt betrieben wird, versucht man es für innovative Vorhaben mit „experimenteller Politik“ (für ein Resümee s. Wollmann 1980). Das heißt: Die Wissenschaft wird aus der Politik als arbeitsteilig voll mitwirkender Bestandteil entlassen. Stattdessen „importiert“ die Politik die *Entscheidungslogik* des „Königswegs“ der Grundlagenforschung und Wirkungsanalyse: die kontrollierte Durchführung und Beobachtung von Experimenten: Bevor eine neue Programmidee flächendeckend implementiert wird, führt man sie *versuchsweise* in abgegrenzten Gebieten (in Feldversuchen oder Pilotprojekten) durch, und zwar so, dass sie „rückholbar“ bleibt. Das heißt, wenn sich die Idee als nicht erfolgreich erweist, kann sie wieder aufgegeben werden. Stehen mehrere Programmalternativen zur Wahl, können sie vergleichend erprobt werden, um später die bestgeeignete endgültig einzuführen. Beispiele für groß angelegte Pilotvorhaben waren die Erprobung des Schulkonzeptes Gesamtschule, Feldversuche zu Bildschirmtext und zum Kabelfernsehen, gegenwärtig das städtebauliche Bund-Länder-Programm „Soziale Stadt“.

Die Pilotprojekte werden jeweils wissenschaftlich begleitet. Aufgaben der wissenschaftlichen Begleitung sind: Monitoring und Analyse der Implementation, Messung der bewirkten Veränderungen im Eingriffsfeld des Programms, Zurechnung der Wirkungen zu den eingesetzten Maßnahmen, Erfolgsbeurteilung durch Vergleich der Ergebnisse mit den Programmzielen, Identifikation und Analyse ungeplanter Effekte. Dabei kann die empirische Wissenschaft auf die schon vorher entwickelte Methodologie der Programmforschung zurückgreifen. Sie wird also in der Funktion als angewandte Sozialforschung tätig und liefert empirisch abgesicherte Informationen, sie ist aber nicht mehr *direkt* eingebunden in den politischen Planungsprozess.

Vom schwierigen Verhältnis zwischen Wissenschaft und Politik

Diese wechselseitige Bezugnahme von Wissenschaft und Politik aufeinander kann einerseits als eine Art Symbiose zum Nutzen beider Seiten begrüßt werden. Es gibt aber auch grundsätzliche Kritik, die warnt, hier verliere die Wissenschaft „einerseits ihre Sonderstellung in der Gesellschaft als die Instanz, die letztlich

gesichertes Wissen autoritativ verkünden kann. Andererseits wird das Prinzip der Forschung, des hypothetischen und experimentellen, lernenden Umgangs mit Information in der Gesellschaft zum allgemeinen Handlungsmodus.“ (von Rauchhaupt 2001)

In der Tat kann mit Recht die Frage aufgeworfen werden, ob Politik und Wissenschaft überhaupt für eine „Symbiose“ hinreichend miteinander kompatibel sind. Politiker und Wissenschaftler handeln in ganz verschiedenen Handlungssystemen und unterliegen ganz unterschiedlichen Interesseneinflüssen und Bedürfniszwängen. „So muß der Politiker in seinem Handeln vor allem solche Bedingungen berücksichtigen, die am sichersten politischen Erfolg versprechen. ... Der Politiker geht somit sehr betont von den Bedürfnissen einer politischen Praxis aus, die im allgemeinen auf schnelles Handeln ausgerichtet und an Praktikabilitätsmaximen, z.B. auch an der Notwendigkeit zum Kompromiß, orientiert ist. ... Der Wissenschaftler bringt seinerseits erfahrungsgemäß in aller Regel Interessen in den Beratungsprozeß ein, die nicht notwendig auf den eigentlichen Zweck des Beratungsgeschäfts beschränkt sind.“ (Schneider 1985, 25) Das erinnert an die alte Auseinandersetzung um das technokratische versus dezisionistische Modell der Politikberatung. „Während im *technokratischen Modell* das Expertenwissen des Beraters als direkt handlungsleitend und eine Problemlösung als Ergebnis der analysierten Sachzwänge gilt – und nicht von politischen und gesellschaftlichen Prozessen der Willensbildung –, verbleibt die Entscheidung im *dezisionistischen Modell* eindeutig beim Politiker, der Ziele und Wege der Beratungstätigkeit bestimmt“ (Clemens 2000, 219f.).

Andererseits verwischen die Unterschiede in dem Maße, wie wissenschaftliche Expertise auch in das politische Alltagsgeschäft einbezogen wird, um damit Entscheidungen abzusichern, „sei es, weil die Probleme zu speziell sind, um durch bloße Staatskunst gelöst zu werden, oder nur, weil man sich mit einem akademisch gesiegelten Gutachten aus der Verantwortung stehlen kann. Die Wissenschaftler jedoch werden dabei leicht in politische Grabenkämpfe hineingezogen, vor allem dort, wo die Sachlage auch wissenschaftlich ungeklärt ist. ... Von außen sind politische und wissenschaftliche Kontroversen bald kaum noch zu trennen“ (von Rauchhaupt, 2001).

Eine solche Entwicklung kann für beide Seiten nicht wünschenswert sein: Wenn darunter die Glaubwürdigkeit der Wissenschaft als Instanz für unparteiische, für „wahre“ Informationen leidet, ist dies auch zum Schaden der Politik, die sich doch durch glaubwürdige Informationen legitimieren will.

III.

Evaluation als angewandte Sozialwissenschaft

Kommen wir nun endlich zum zentralen Punkt des Vortragsthemas: „Evaluation“. Zwar ist das meiste bereits in den beiden vorangegangenen Abschnitten gesagt worden, denn Evaluation ist auch – und vor allem – ein Spezialfall angewandter Sozialwissenschaft. Der Begriff ist jedoch in letzter Zeit in so viele Diskussionszusammenhänge eingedrungen und hat sich dabei zugleich so weit „veralltäglich“, dass ich es dennoch nicht bei einem einfachen Verweis auf das bisher Gesagte bewenden lassen kann.

„Evaluation“ ist zu einem äußerst unscharfen Modewort geworden und wird von manchem lediglich als „wohlklingendes“ Fremdwort für jede Form von Bewertung oder Beurteilung verwendet. Doch selbst wenn der Begriff klar definiert ist, bleibt eine gewisse Unschärfe, da dasselbe sprachliche Zeichen für unterschiedliche Typen von *Referenzobjekten* stehen kann (und steht). Nicht selten ist dies der Ausgangspunkt für Missverständnisse und Verwirrungen.

- Eine erste Gruppe von Referenzobjekten ist auf der symbolischen und gedanklichen Ebene angesiedelt. „Evaluation“ steht für ein *spezifisches wissenschaftliches Denkmodell*: ein nachprüfbares Verfahren

des Bewertens. Dieses Verständnis war bisher implizit gemeint, wenn ich von Evaluation als Spezialfall angewandter Sozialforschung gesprochen habe.

- Die zweite Begriffsebene bezieht sich auf ein *spezifisches Handeln*, einen Prozess: auf *zielorientiertes Informationsmanagement*. Um diese (Handlungs-)Ebene ging es in erster Linie, wenn aus der politischen Perspektive argumentiert und z.B. von der Tendenz zu einer Symbiose von (wissenschaftlich angeleiteter) Politik und (an politischem Handeln orientierter) Wissenschaft die Rede war. Im allgemeinen Sinne gilt hier als Evaluation jede methodisch kontrollierte, verwertungs- und bewertungsorientierte Form des Sammelns, Auswertens und Verwertens von Informationen. Dabei ist es müßig, darüber zu streiten, ob das Erheben rein deskriptiver Daten über einen zu bewertenden Sachverhalt „schon“ und das Ziehen von Schlussfolgerungen und Konsequenzen für diesen Sachverhalt „noch“ zur Evaluation zählt.
- Und schließlich bezeichnet „Evaluation“ auch noch etwas Punktuelleres: das *Resultat* des Evaluationsprozesses, die Dokumentation der Wertaussagen in einem Evaluationsbericht oder – gutachten.

Ähnlich vielfältig (und meist unbestimmt bleibend) ist die *Semantik* des Begriffs. Das ist nicht überraschend, da wir auf „Evaluation“ und „Evaluierung“ in den verschiedensten Diskussionskontexten stoßen – im Alltag ebenso wie in der Politik, in der Methodologie empirischer Wissenschaft ebenso wie im (spezifischeren) Zusammenhang der Umfrageforschung – und da selbstverständlich unterschiedliche Kontexte sich unterschiedliche Begriffsbedeutungen schaffen.

- Am wenigsten spezifisch ist der *alltägliche Sprachgebrauch*: „Evaluation“ bedeutet nichts weiter, als dass *irgend jemand irgend etwas nach irgendwelchen Kriterien in irgendeiner Weise bewertet*. Derselbe Sachverhalt kann von verschiedenen Individuen sehr unterschiedlich bis gegensätzlich eingeschätzt und beurteilt werden.
- In *politischen Argumentationen* sind die Begriffsverwendungen zwar wesentlich spezifischer, unglücklicherweise aber auch außerordentlich vielfältig. Die Bezeichnung gilt für *Effizienzmessungen* in ökonomischen Zusammenhängen ebenso wie für die von Sachverständigen vorgenommene *Analyse* der Funktionsfähigkeit von Organisationen (etwa bei der Evaluation wissenschaftlicher Einrichtungen). Auch die *beratende und moderierende Beteiligung* im Prozess der Entwicklung von Handlungsprogrammen mit dem Ziel ihrer Optimierung wird von diesem Begriff erfasst („formative“ oder „responsive“ Evaluation).
- In der *empirischen Methodologie* meint „Evaluation“ hingegen das *Design* für einen spezifischen Typ von Sozialforschung, der die Beschaffung von Informationen über Verlauf und Resultate eines (Handlungs- und Maßnahmen-), „Programms“ mit explizit formulierten Zielen und Instrumenten zum Gegenstand hat. Evaluationsziele sind die „wissenschaftliche Begleitung“ der Programm-Implementation und/oder die „Erfolgskontrolle“ und „Wirkungsanalyse“. Der Ansatz ist im Idealfall experimentell oder quasi-experimentell und strebt an, zu empirisch begründeten, intersubjektiv nachprüfbaren und somit „objektivierten“ *Bewertungen* zu gelangen.
- Schließlich wird auch im Zusammenhang „gewöhnlicher“ *Umfrageforschung* von Evaluation gesprochen. Hier ist die Erhebung und Auswertung bewertender (also „evaluierender“) Aussagen von Befragten gemeint, die in einem angebbaren Verhältnis zu dem zu evaluierenden Gegenstand stehen (etwa Kunden/Klienten, Betroffene, Teilnehmer von Bildungsveranstaltungen). Ermittelt werden durch „Evaluations“-Umfragen *subjektive Werturteile* anhand explizit vorgegebener spezifischer Kriterien ebenso wie allgemeinere *Zufriedenheits- oder Unzufriedenheitsäußerungen* oder auch *Akzeptanzinformationen*. Ein spezifisches Evaluationsdesign existiert in diesem Fall nicht. Ins Auge

fällt stattdessen die Analogie zur Meinungsforschung, mit dem einzigen Unterschied, dass nicht Meinungen, sondern Bewertungen und/oder Zufriedenheitseinschätzungen abgefragt werden.

Gemeinsam ist allen diesen Verwendungen, dass – im *Unterschied zum alltagssprachlichen Verständnis* – nicht „irgend etwas“ evaluiert wird, sondern dass spezifizierte Sachverhalte, Programme, Maßnahmen, manchmal auch ganze Organisationen Gegenstand der Betrachtung sind. Zweitens nimmt nicht „irgend jemand“ die Evaluation vor, sondern es sind Personen, die dazu in besonderer Weise befähigt erscheinen: wissenschaftliche „Sachverständige“, methodische oder durch Praxiserfahrungen ausgewiesene „Experten“, konkret „Betroffene“. Drittens kommt das Urteil nicht nach „irgend welchen“ Kriterien zustande, sondern diese müssen *explizit* auf den zu bewertenden Sachverhalt bezogen sein. Und schließlich darf bei einer systematischen Evaluation nicht „irgendwie“ vorgegangen werden, sondern das Verfahren ist zu „objektivieren“, d.h. im Detail zu planen und in einem „Evaluationsdesign“ verbindlich für alle Beteiligten festzulegen.

Das wissenschaftliche Design: Programmforschung

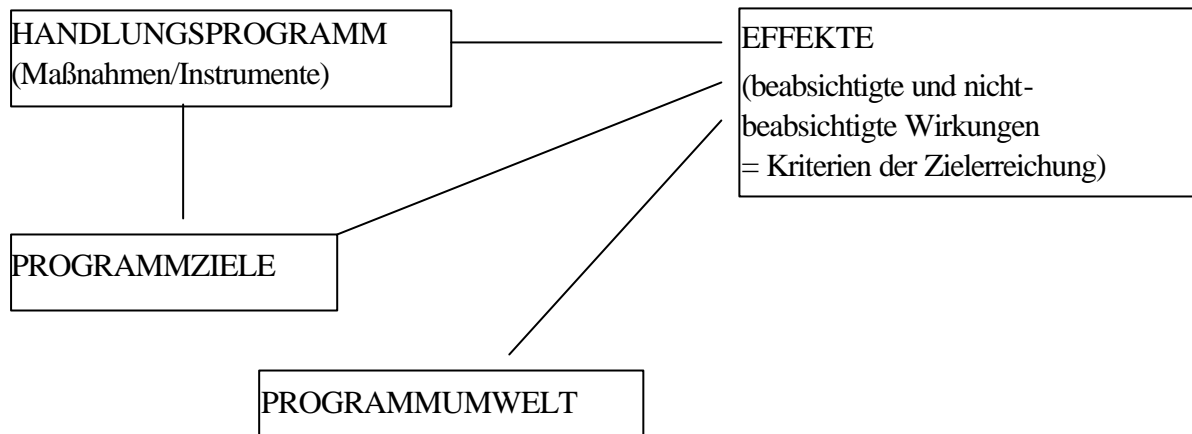
Es mag erstaunen, dass trotz dieser offensichtlichen Vielfalt der „öffentliche Blick“ auf das Konzept „Evaluation“ eher eindimensional bleibt. Wenn nämlich die Funktion von Evaluation wirklich zum Gegenstand des wissenschaftlichen und/oder gesellschaftlichen Diskurses wird, dann bezieht sich die Argumentation (meist implizit, seltener explizit) weit überwiegend nur auf das *Forschungsmodell der Evaluation* in seiner spezifischen Form als Programmforschung – vermutlich eine Spätfolge ihrer Einbindung in das oben skizzierte Modell informationsgestützter Planung und ihres damit gewonnenen Gewichts. Problematisch ist diese Engführung in dreifacher Hinsicht:

1. Das Dimensionsmodell dieses Designs wirkt einfach und erscheint unmittelbar plausibel.
2. Es ist – fast immer – empirisch nicht einlösbar (dies werde ich im Folgenden näher erläutern). Folglich wird
3. hinter der (Programmforschungs-),„Fassade“ faktisch ganz anders verfahren.

In der *Methodologie empirischer Sozialforschung* bezeichnet Evaluation das *Design* für einen spezifischen Forschungstyp, durchgeführt im idealtypischen Fall als Feldexperiment mit Kontrollgruppen. Wo die Voraussetzungen dafür nicht in vollem Maße erfüllt sind – und das ist überwiegend der Fall –, behilft man sich mit „Ersatzlösungen“ für diejenigen Designkomponenten, die nicht idealtypisch realisiert werden können, wobei aber die generelle Orientierung an der Vorgehens- und Argumentationslogik des Experiments erhalten bleibt.

Gegenstand der Evaluation können unterschiedlichste Vorhaben in unterschiedlichsten Phasen ihrer Realisierung sein (von der Entwurfsphase eines Entwicklungsprojekts bis zur nachträglichen Überprüfung eines seit längerem in der Praxis „bewährten“ Programms. Bei aller denkbaren Vielfalt bleibt dennoch allen Vorhaben gemeinsam, dass sie (mindestens) drei interdependente Dimensionen aufweisen – nämlich Ziele, Maßnahmenprogramm, Effekte – und dass sie (anders als in einem Forschungslabor) von Umgebungseinflüssen nicht abgeschirmt werden können.

Abbildung 2: Programm-Evaluation



Die drei in der Abbildung dargestellten Programmdimensionen (Ziele – Maßnahmen – Effekte) können jeweils mehr oder weniger konkret oder abstrakt, mehr oder weniger festliegend oder variabel, mehr oder weniger ausformuliert oder nur implizit, mehr oder weniger offiziell oder informell sein. In jedem Fall aber orientieren die Beteiligten in dem zu evaluierenden Programm ihr Argumentieren und Handeln daran. Mit diesen drei Dimensionen muß sich daher auch jede Evaluation auseinandersetzen: Ungenaue Formulierungen von Zielen und Maßnahmen sind zu präzisieren und zu operationalisieren, implizit gelassene zu rekonstruieren, ungeordnete Ziele sind in einem Zielsystem zu ordnen, Zielkonflikte herauszuarbeiten. Ziele sind von Maßnahmen (als Instrumente zu deren Erreichung) abzugrenzen. Die Art und Weise der vorgesehenen Realisierung (Implementation) ist zu berücksichtigen und ggf. zu konkretisieren. Schließlich ist zu klären, was das Handlungsprogramm im Detail bewirken soll (und *darüber hinaus* bewirken kann): Welche Veränderungen müssen in welcher Frist an welcher Stelle auftreten, damit die Ziele als erreicht gelten? Auf welche Weise können sie festgestellt und gemessen werden? Wie können feststellbare Veränderungen als Wirkungen des Programms identifiziert und gegenüber anderen Einflüssen abgegrenzt werden?

Das Modell scheint in idealer Weise dem Informationsbedarf des oben skizzierten Politikmodells zu entsprechen. Planung und Politik (Legislative ebenso wie Exekutive) benötigen Informationen über den Stand und die Veränderung der Zielerreichungsgrade aus der *Ergebniskontrolle*, über die eingesetzten Mittel und den Ablauf der Maßnahmen aus der *Durchführungskontrolle* und über mitwirkende exogene Faktoren aus der *Bedingungskontrolle*. Außerdem besteht ein Bedarf an Daten über die Zielerträge der eingesetzten Maßnahmen – *Wirkungskontrolle* –, an Methoden für das Zusammenwiegen der Zielerträge – *Bewertung* –, an Aussagen über ausgelassene Chancen – *Kontrolle der Möglichkeitsgrenzen* – sowie an Maßstäben, mit denen die *Leistung der Verwaltung* beurteilt werden kann (Eekhoff u.a. 1977, 11 ff.).

Die empirische Forschungsmethodik stellt *im Prinzip* das Instrumentarium zur Einlösung dieser Informationsbedarfe zur Verfügung. Für die Ergebnis-, Durchführungs- und Bedingungskontrolle sind „lediglich“ gültige und zuverlässige deskriptive Messungen der Zielgrößen in dem durch das Programm definierten Wirkungsfeld (Monitoring des Outputs) sowie der durchgeführten Maßnahmen (Monitoring der Instrumentvariablen, des Inputs) und der zugleich gegebenen externen Einflüsse (Monitoring der exogenen Variablen) erforderlich. Schwieriger ist zwar die Aufgabe der Zurechnung der Veränderungen zu den durchgeführten Maßnahmen, also die Feststellung der quantitativen und qualitativen Zielerträge. Aber auch hierfür bietet sich ein wissenschaftlich akzeptiertes Design an: das der experimentellen (bzw. – wo nicht alle Einflussgrößen methodisch kontrollierbar sind – der quasi-experimentellen) Untersuchungsanordnung.

Selbst die eigentliche *Erfolgsbewertung* wirft begründungslogisch keine besonderen Probleme auf: Das Programm selbst stellt die Kriterien (Zieldimensionen) und Standards (zu erreichende Sollwerte) bereit, so dass die Bewertung sich auf einen Vergleich von gemessenen mit den angestrebten Zielerträgen sowie auf deren Gewichtung für ein Gesamturteil reduziert.

Schwieriger gestaltet sich die von *Eekhoff* geforderte Kontrolle auch der „Möglichkeitsgrenzen“. So sehr sein Argument („Im Interesse einer effektiven Entwicklungssteuerung sollte sich das Augenmerk nicht nur auf die tatsächlichen Wirkungen abgeschlossener und laufender Programme, sondern auch auf die nicht genutzten Möglichkeiten richten“; Eekhoff, a.a.O.) einleuchtet, wären für eine diesbezügliche, intersubjektiv nachvollziehbare Evaluationsaussage hypothetische Situationszustände (Szenarien) heranzuziehen, deren Realitätsgehalt immer fraglich bleiben müsste. Und ob – schließlich – die Evaluation so umfassenden Zugang auch zu internen Verwaltungsvollzugsdaten erhalte, dass deren Effizienz „in allen Teilbereichen bis hinunter zu den einzelnen Bediensteten“ (a.a.O.) einschätzbar würde, darf wohl bezweifelt werden.

Die beiden letztgenannten Punkte einmal außer Acht gelassen (derart umfassende Evaluationen wurden meines Wissens ohnehin niemals ernsthaft in Angriff genommen⁷), wirkt das skizzierte Modell einer kausalanalytisch angeleiteten Programmevaluations- und Wirkungsforschung in sich schlüssig und einleuchtend und scheint nur noch einer weiteren Differenzierung hinsichtlich der Methodik zu bedürfen (worauf ich in diesem Vortrag verzichte). Bei näherem Hinsehen allerdings wird erkennbar, daß es von außerordentlich anspruchsvollen Voraussetzungen über den Gegenstand der Untersuchung ebenso wie bei den programmdurchführenden Instanzen und der Evaluation selbst ausgeht. Diese Voraussetzungen mögen zwar bei Vorhaben der Grundlagenforschung (vereinzelt) gegeben sein, sind jedoch in Programmforschungsprojekten wenig realitätsnah. Drei dieser meist implizit gelassenen Voraussetzungen sind besonders hervorzuheben, da deren Erfüllung eine wesentliche Bedingung dafür ist, das methodologische Forschungsprogramm empirischer Kausalanalysen überhaupt anwenden zu können:

- *Vor* der Entwicklung des Forschungsdesigns muss Klarheit über die Untersuchungsziele – bezogen auf einen definierbaren und empirisch abgrenzbaren Untersuchungsgegenstand – bestehen. Für die Dauer der Datenerhebung dürfen sich weder die Untersuchungsziele noch die wesentlichen Randbedingungen des Untersuchungsgegenstandes in unvorhersehbarer Weise ändern.
- *Vor* der Entwicklung des Forschungsdesigns müssen begründete Vermutungen (Hypothesen) über die Struktur des Gegenstandes wie auch über Zusammenhänge und Beziehungen zwischen dessen wesentlichen Elementen existieren, nach Möglichkeit in Form empirisch bewährter Theorien. Erst auf ihrer Basis kann ein Gültigkeit beanspruchendes Indikatorenmodell konstruiert, können geeignete Meßinstrumente entwickelt, kann über problemangemessene Auswertungsverfahren entschieden werden.
- Der Forscher muss die Kontrolle über den Forschungsablauf haben, um die (interne und externe) Gültigkeit der Resultate weitestgehend sicherzustellen.

Im Normalfall der Begleitforschung zu Programm-Implementationen oder gar zu Modellversuchen neuer Techniken, neuer Schulformen, zur Erprobung alternativer Curricula oder Lernformen u.ä. ist keine einzige dieser Bedingungen voll erfüllt. Die Untersuchungssituation weist vielmehr in dieser Hinsicht erhebliche „Mängel“ auf. Selbst in günstigsten Fällen wird bei weitem nicht das erfüllt sein, was Wollmann/Hellstern als konkrete „politische Programme“ definieren, nämlich „komplexe Aktionsmodelle . . . , die auf die Erreichung bestimmter Politikziele gerichtet sind, auf bestimmten Handlungsstrategien beruhen und für deren Abwicklung bestimmte finanzielle, personelle und sonstige administrative Ressourcen (Richtlinien,

⁷ Das bedeutet allerdings nicht, dass nicht auch in neueren Veröffentlichungen Konzepte für ähnlich umfassende Programmevaluationen vorgestellt würden, so z.B. der „Acht-Fragen-Ansatz zur Politikevaluation“ von Vedung (1999, 85ff.).

„flankierende Maßnahmen‘ usw.) bereitgestellt werden“ (1978, 7). Nimmt man diese Definition beim Wort, müssten die Kenntnisse, die man sich durch die Forschung erhofft, bei den Akteuren bereits weitestgehend vorhanden sein. Ähnliches gilt hinsichtlich des vorab notwendigen empirisch abgesicherten Wissens über die interne Struktur des Untersuchungsgegenstandes, um ein angemessenes Design entwickeln und methodisch begründen zu können. Und hinsichtlich des dritten o.g. Punktes gilt quasi per Definition, dass die Anforderungen des zu erprobenden Programms Vorrang vor den Anforderungen sozialwissenschaftlicher Forschung haben (müssen). Eine Evaluation, die die Bedingungen der Programmdurchführung entsprechend den Anforderungen der Forschung mitgestalten könnte, würde lediglich Forschungsartefakte produzieren – denn sie würde ein Programm begleiten und evaluieren, das ohne die Einfluss nehmende Forschung so nicht existiert hätte und auch im Nachhinein so nicht existieren würde (ausführlicher dazu Kromrey 1988).

Fazit: Das Evaluationsdesign der Programmforschung weckt durch seine plausibel scheinende Struktur Erwartungen, die die Empirie nicht einlösen kann. Die Forscher im Evaluationsfeld müssen daher nach Ersatzlösungen suchen (ex-post-facto-Analysen statt Experimentaldesign, Indikatorenzeitreihen, von den subjektiven Handlungstheorien der Akteure angeleitete Deskriptionen, Fragebogenerhebungen bei Beteiligten und/oder Betroffenen u.ä.). Das Resultat sind vergleichsweise „weiche“ Daten, die entweder (im günstigeren Fall) bei den Adressaten der Evaluation Enttäuschung über die „Irrelevanz der Forschung“ hervorrufen oder (im ungünstigen Fall) von den Adressaten dennoch als „wissenschaftlich belegte Fakten“ für bare Münze genommen werden und zu Fehlentscheidungen Anlass geben können.

Damit dies nicht geschieht – damit also weder nicht einlösbare Hoffnungen geweckt werden noch die Gefahr von Fehlentscheidungen provoziert wird –, sollte das Design der Programmforschung abgelöst werden durch ein Modell, das die verschiedenen Funktionen (Programmentwicklung und/oder – implementation, empirische Informationsgewinnung und –aufbereitung, Evaluierung, Entscheidung über Schlussfolgerungen/Konsequenzen) klar voneinander abgrenzt und in dem die beteiligten Akteure (Entwickler/Durchführende, empirische Wissenschaftler, Adressaten und Betroffene) mit ihren Interessen und Kompetenzen erkennbar werden. Funktionen, die die Wissenschaft entsprechend ihrer Kompetenz in diesen Kontext einbringen kann, sind: zielgenaue Beschaffung empirischer Informationen als Grundlage von Entscheidungen und Evaluierungen, Interpretation der Informationen durch zielgruppen-adäquate Erläuterung und Beratung, ggf. Moderation und Mediation zwischen divergierenden Interessen wie auch zwischen verschiedenen Situationsdeutungen der Beteiligten. Die Funktion des Evaluierens (des Fällens von bewertenden Urteilen) habe ich bewusst in dieser Funktionszuschreibung nicht aufgeführt.

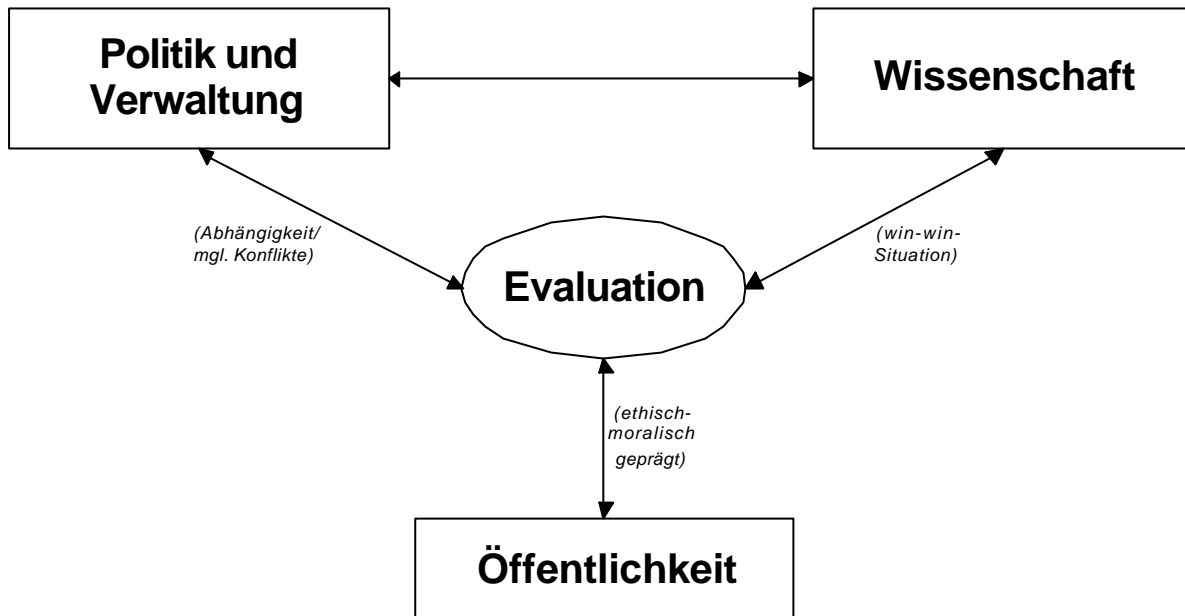
IV.

Interaktion Evaluation – Politik – Wissenschaft – Öffentlichkeit

Um noch einmal ausdrücklich zu meinem Thema zurückzukommen: Ich habe bisher aus verschiedenen Perspektiven versucht, „Evaluation“ als ein Konzept zu beleuchten, das in unauflöslicher Weise eingebunden ist in den Kontext von Wissenschaft einerseits sowie Politik, Planung und Verwaltung andererseits, so unauflöslich, dass kaum noch entscheidbar ist, wo denn „die Evaluation“ eigentlich zu Hause ist. Das Verhältnis der Evaluation zu ihren Umwelten ist dadurch – auch das habe ich versucht zu illustrieren – alles andere als unproblematisch. Dennoch habe ich bisher nicht das ganze Ausmaß des Problemnetzwerks herausgearbeitet, sondern bewusst vereinfacht. Wenn von „Gesellschaft“ die Rede ist, darf natürlich die „Öffentlichkeit“ nicht vergessen werden. Und hier öffnet sich für die Evaluation wieder ein zusätzliches Konfliktfeld: Während die Wissenschaft die Orientierung an möglichst präzisen, streng methodisch abgesicherten Resultaten verlangt, hat die Politik Bedarf an schnellen, termingenauen, möglichst „einfachen“ Informationen; während die Politik vertrauensvolle (und auch vertrauliche) Information und Beratung wünscht, möchte die Öffentlichkeit möglichst frühzeitig und möglichst vollständig informiert werden, gerade um auch das politische Entscheiden und Handeln kontrollieren zu können. Die folgende

Abbildung, die ich dem instruktiven Aufsatz von Th. Widmer (2000) – „Wenn Wissenschaft zur praktischen Kunst wird“ – entnommen habe, visualisiert anschaulich den Ort der Evaluation „zwischen allen Stühlen“; sie eignet sich daher ausgezeichnet zur Strukturierung eines abschließenden Resümees.

Abbildung 3: *Evaluation in Wissenschaft und Gesellschaft*



(nach Th. Widmer, 2000: Qualität der Evaluation, S. 82)

Evaluation – Politik/Verwaltung

Politik und Verwaltung sind, darauf weist Widmer hin, am Evaluationsprozess zumeist in verschiedenen Rollen beteiligt. Sie sind häufig Auftraggeber einer Evaluation, aber oft auch wichtige Informationsquelle und zugleich auch noch meist ihr primärer Adressat. „Nicht zuletzt sind Politik/Verwaltung auch die intendierten Nutzer einer Evaluation. ... Die Ansprüche, die seitens Politik/Verwaltung an die Evaluation gerichtet werden, lassen sich mit den Stichworten umschreiben: Nützlichkeit, Zeitgerechtigkeit, politische Opportunität, Anschlußfähigkeit, Sparsamkeit, Praktikabilität und Korrektheit“ (Widmer 2000, 83). Evaluation muss, wenn sie Auftragnehmerin ist, ein selbstverständliches Interesse daran haben, die Erwartungen des Auftraggebers zu erfüllen, ist sie doch – je nach Organisationsform – finanziell mehr oder weniger vom Auftraggeber abhängig. Mit anderen Worten: „Das Verhältnis zwischen Politik/Verwaltung und Evaluation ist somit schwergewichtig ökonomisch und juristisch geprägt“ (Widmer, a.a.O.)

Evaluation - Wissenschaft

Die spezifische Struktur des Verhältnisses von Wissenschaft und Evaluation habe ich schon mehrfach angesprochen und dabei besonders die „Problemzonen“ berührt. Widmer weist dagegen auf Aspekte hin, die die Wechselbeziehung eigentlich eher zu einer „problemfreien Zone“ werden lassen sollte. Einerseits profitiert nämlich Wissenschaft von der Evaluation in zweierlei Hinsicht: „Die Evaluation ist in der Lage:

- Erstens, die gesamtgesellschaftliche Legitimation des Wissenschaftssystems zu unterstützen. Dies geschieht insbesondere dann, wenn sich die Evaluation als praxisrelevant erweist.
- Zweitens, für wissenschaftliche Erkenntnisse wie Theorien und Methoden empirische Evidenz zu produzieren. Gleichzeitig kann die Evaluation die Fachdisziplinen auch mit zusätzlichem inhaltlichem Wissen alimentieren (...)

Die Evaluation wiederum ist auf die Wissenschaft als Lieferantin wissenschaftlicher Wissensbestände angewiesen. Die Evaluation kann nur mit Wissenschaft bestehen, da letztere zu ihrer gesellschaftlichen Reputation beiträgt. Ohne die ‚reine‘ Wissenschaft würde die Evaluation ihre Daseinsberechtigung als wissenschaftliche Dienstleistung verlieren. Das Verhältnis zwischen Wissenschaft und Evaluation ist damit hauptsächlich durch den Austausch von Wissen und Reputation geprägt“ (Widmer 2000, 83f.) – und stellt sich aus dieser Perspektive somit als geradezu prototypische Win-win-Situation dar. „Missionieren“ wir also in unserer Disziplin, auf dass diese Kenntnis größere Verbreitung findet!

Evaluation – Öffentlichkeit

Eine deutlich andere Qualität weist das Verhältnis zwischen Öffentlichkeit und Evaluation auf. Eine direkte gegenseitige Abhängigkeit – wie in den beiden anderen Beziehungsstrukturen – besteht hier nicht. Da die Evaluator – so können wir bei Widmer lesen – zumindest vordergründig nicht auf eine Interaktion mit der Öffentlichkeit angewiesen scheint, könnte es sich geradezu anbieten, auf Interaktionen mit ihr gänzlich zu verzichten und sich primär in den für sie „lebensnotwendigen“ Beziehungen mit Politik und Wissenschaft zu engagieren (Widmer 200, 84). Es läge also nahe, die Informations- und Aufklärungsbedürfnisse der Öffentlichkeit zu vernachlässigen und sich so im Dienste von Politik/Verwaltung und/oder Wissenschaft instrumentalisieren zu lassen. Ein solches Verhalten wäre allerdings nicht nur kurzsichtig, sondern auch aus der Sicht einer Anwendungsforschung, die sich auf das normative Postulat einer nicht „wertfreien“, sondern „wertsensiblen“ Wissenschaft beruft (s. Eberlein/Obermeier 1985), nicht vertretbar. Ähnlich argumentiert Widmer: „Ein demokratisch verantwortungsvoller Evaluationsprozeß erfordert nicht nur, daß der resultierende Schlußbericht zuhanden der Öffentlichkeit publik gemacht wird. Der Evaluationsprozeß sollte von Beginn weg die Anliegen der Beteiligten & Betroffenen berücksichtigen“ (a.a.O.).

V.

Offene Fragen

Mit diesem Plädoyer bin ich eigentlich am Ende meines Vortrags angelangt. Kurz hinweisen möchte ich aber dennoch auf ein paar offen gebliebene Fragen, die durchaus zu meinem Thema zu zählen wären, die aber draußen bleiben mussten.

Nicht angesprochen wurde beispielsweise (indem ich mich immer wieder auf das Evaluationsmodell "Programmforschung" zurückgezogen habe) ein ganze Perspektive: Evaluation durch Befragung. Nicht angesprochen wurde auch (indem ich mich als „Gegenstände“ der Evaluation ohne weitere Spezifizierung abstrakt auf „Programme“ bezogen habe) ein ganzer Gegenstandsbereich: Evaluation der Wissenschaft selbst.

Bisher habe ich Evaluation im Rahmen angewandter Sozialforschung in etwa so behandelt:

- *Hier* ist die Wissenschaft, und aus ihr hat sich der Verfahrenstyp Evaluation herausdifferenziert;
- *dort* ist die Gesellschaft, und in ihr gibt es "Gegenstände" (Programme, Einrichtungen), die auf der Basis empirisch-wissenschaftlicher Informationen bewertet werden sollen, im Idealfall unter Bezugnahme auf die expliziten Ziele des Programms oder der Einrichtung, ggf. ergänzt um die Ziele der Adressaten.

Nun ist aber in letzter Zeit die Wissenschaft selbst zum *Gegenstand* von Evaluation geworden; das heißt: Wissenschaft wird plötzlich und unerwartet als ein gesellschaftlicher Bestandteil neben anderen angesehen. Sich damit abzufinden, ist gar nicht so einfach!

Aber wie sollte das – das Evaluieren von Wissenschaft mit einem Konzept aus der Wissenschaft – denn gehen nach dem bisher Ausgeführten?

- Wer evaluiert? (die Wissenschaftler, „die Gesellschaft“ z.B. in Form des Marktes, die Politiker?)
- Wer beschafft die Informationen, die evaluations-relevanten Daten?
- Auf welche und auf wessen Ziele wird Bezug genommen? (auf die Ziele der Gesellschaft, repräsentiert durch die jeweilige politische Mehrheit? auf die Ziele der Wissenschaft selbst?)
- Was sind überhaupt die Zielerreichungskriterien? Wie definiert man „Erfolg“ von Wissenschaft?

Und was passiert eigentlich, wenn im Hochschulbereich bei der Evaluation von Lehre und Studium überwiegend nur Studierende befragt werden? Welche Kriterien wenden diese Evaluatoren an? Warum evaluieren dieselben Personen, wenn sie nicht mehr im Studium, sondern im Beruf sind, denselben Gegenstand anders? Was folgt daraus?

Was passiert, wenn statt dessen peers (Wissenschaftler desselben Fachs) die Evaluationen vornehmen? Ist diese Selbstevaluation (Subjekt und Objekt der Bewertung sind identisch) überhaupt Ernst zu nehmen? Man stelle sich vor: in der Justiz würde die Regelung eingeführt, dass die Beklagten das Urteil über ihre Tat selbst sprächen.

Die Liste der Fragen ließe sich leicht verlängert. Ich will dazu heute keine Antworten suchen. Stattdessen komme ich nun doch zum Schluss:

Der Anlass, aus dem ich zu diesem Vortrag eingeladen worden bin - Eröffnung des Zentrums für Evaluation -, dieser Anlass war mir Motiv für mein Anliegen, zu skizzieren und zu begründen, dass diese Einrichtung *nicht* eine Nischenfunktion wahrnimmt, dass sie keine Orchidee zur Verzierung der akademischen Disziplin Soziologie mit einem Hauch von Praxisrelevanz ist, sondern dass sie unauflösbar verbunden ist mit allen Teilkulturen der Wissenschaft, und dass sie ebenso unauflösbar verbunden ist mit allen Sektoren von Gesellschaft und Politik.

Dass ich hierbei besonders auf die jeweiligen Problemfelder aufmerksam gemacht habe, soll die Gründer nicht entmutigen, sondern soll ihr Bewusstsein stärken nicht nur für die Daseinsberechtigung, sondern für die *Notwendigkeit* einer solchen Einrichtung. Es gibt viel zu tun. Also: packen Sie's an!

Quellen

Clemens, Wolfgang, 2000: Angewandte Sozialforschung und Politikberatung. In: Clemens/Strübing 2000, 211-232

Clemens, Wolfgang; Strübing, Jörg (Hg.), 2000: Empirische Sozialforschung und gesellschaftliche Praxis. Bedingungen und Formen angewandter Forschung in den Sozialwissenschaften, Opladen: Leske+Budrich

Eberlein, G.; Obermeier, O.P., 1985: Wertsensible Wissenschaft als pragmatisches Alternativparadigma. Zur Neubegründung „reiner“ und angewandter Sozialwissenschaften. In: Klages 1985, 10-34

Eekhoff, Johann; Muthmann, Rainer; Sievert, Olaf; Werth, Gerhard; Zahl, Jost, 1977: Methoden und Möglichkeiten der Erfolgskontrolle städtischer Entwicklungsmaßnahmen, Bonn-Bad Godesberg: Schriftenreihe "Städtebauliche Forschung" 03.060

Habermehl, Werner, 1992: Angewandte Sozialforschung. München, Wien: Oldenbourg

Häußermann, Hartmut; Siebel, W., 1978: Thesen zur Soziologie der Stadt. In: Leviathan 6, Heft 4, 484-500

Hornbostel, Stefan, 1997: Wissenschaftsindikatoren. Bewertungen in der Wissenschaft, Opladen: WV

- Klages, Helmut (Hg.), 1985: Arbeitsperspektiven angewandter Sozialwissenschaft, Opladen: WV
- Kromrey, Helmut, 1988: Akzeptanz- und Begleitforschung. In: Massacommunicatie (Nijmegen), Jg. 16, Nr. 3, 221-242
- Kromrey, Helmut, 1999: Von den Problemen anwendungsorientierter Sozialforschung und den Gefahren methodischer Halbbildung. In: Sozialwissenschaften und Berufspraxis, Jg. 22, H. 1, 40-57
- Memorandum "Dialog Wissenschaft und Gesellschaft", 1999: www.berlinews.de/archiv/428.shtml
- Rauchhaupt, Ulf von, 2001: Einsinkende Altbauten. Verliert die Wissenschaft ihre Autorität? Die "Stunde der Wahrheit" des Soziologen Peter Weingart. Rezension in: DIE ZEIT 13/2001, Kultur)
- Schneider, Norbert, 1985: Politik und Wissenschaft - ein fruchtbares Spannungsverhältnis. In: Max Kaase, Manfred Küchler (Hg.): Herausforderungen der Empirischen Sozialforschung, Mannheim: ZUMA 1985, 24-32
- Stockmann, Reinhard (Hrsg.), 2000: Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder, Opladen: Leske+Budrich
- van de Vall, Marc, 1993: Angewandte Sozialforschung. Begleitung, Evaluierung und Verbesserung sozialpolitischer Maßnahmen, Weinheim, München: Juventa
- Vedung, Evert, 1999: Evaluation im öffentlichen Sektor, Wien, Köln, Graz: Böhlau
- Werner, Rudolf 1975: Soziale Indikatoren und politische Planung. Einführung in Anwendungen der Makrosoziologie, Reinbek: rororo
- Widmer, Thomas, 2000: Qualität der Evaluation - Wenn Wissenschaft zur praktischen Kunst wird. In: Stockmann 2000, 77-102
- Wiebe, Burckhard (Hg.), 1988: Zwischen Marktplatz und Elfenbeinturm - Gesellschaft und Sozialwissenschaften, Berlin
- Wollmann, Hellmut; Hellstern, G.-M., 1978: Sanierungsmaßnahmen. Städtebauliche und stadtstrukturelle Wirkungen, Bonn-Bad Godesberg: Schriftenreihe „Städtebauliche Forschung“, 02.012
- Wollmann, Hellmut, 1980: Implementationsforschung - eine Chance für kritische Verwaltungsforschung? In: H. Wollmann (Hg.): Politik im Dickicht der Bürokratie, Opladen (Leviathan Sonderheft 3), 9-48

Der Autor:

Prof. Dr. Helmut Kromrey, Institut für Soziologie, Freie Universität Berlin

Garystr. 55, 14195 Berlin, Tel. 030 / 83857618

eMail: kromrey@zedat.fu-berlin.de; URL: <http://userpage.fu-berlin.de/~kromrey/>