

Jahrestagung 2001:

"Welche Evaluationskultur braucht die Eidgenossenschaft?"

Congrès annuel 2001:

"Quelle culture d'évaluation pour la Suisse?"

Tagungsbeiträge

Philippe Mastronardi:

Die Tragweite von Art. 170 BV aus juristischer Sicht

Artikel 170 BV lautet: Die Bundesversammlung sorgt dafür, dass die Massnahmen des Bundes auf ihre Wirksamkeit überprüft werden. Der Verfassungsauftrag zur Wirkungsprüfung bedeutet zunächst, dass die Methodik der Evaluation nicht mehr allein eine Frage der disziplinären Logik ist. Die Evaluatoren erfüllen mit ihrer Arbeit eine Verfassungsaufgabe. Das können sie nur im Verbund mit den Politikern einerseits, den Juristen andererseits. Die Evaluation staatlichen Handelns wird zu einer Rechtsfrage. Denn im demokratischen Rechtsstaat sind alle wichtigen Ziele durch den Gesetzgeber festzulegen. Die Zielanalyse, welche den Massstab für die Wirksamkeit abgibt, besteht daher zunächst aus einer Gesetzesinterpretation. Dafür ist die Zusammenarbeit des Evaluators mit dem Juristen erforderlich. Die Festlegung von Kriterien der Wirksamkeit ist eine interdisziplinäre Aufgabe.

Der Verfassungsauftrag an die Bundesversammlung enthält sodann drei Elemente:

1. Die Evaluation wird der Bundesversammlung anvertraut, weil diese die demokratisch legitimierte Gesamtverantwortung für das Handeln des Bundes zu tragen hat;
2. die Bundesversammlung trägt dafür die umfassende Verantwortung, d.h., sie bestimmt Gegenstand, Umfang und Inhalt der Evaluation im Bund;
3. andererseits muss sie die Aufgabe nicht selber erfüllen, sondern nur dafür sorgen, dass sie erfüllt wird. Sie kann dafür geeignete Institutionen schaffen.

Die Wirkungsprüfung soll demnach ein Instrument in der Hand der obersten demokratischen Behörde des Bundes sein. Das Urteil über die Wirksamkeit der staatlichen Massnahmen ist letztlich nicht Sache der Experten, sondern der Politik. Die Legitimation des Staates über seine Wirksamkeit wird der demokratischen Legitimation untergeordnet.

Schliesslich erschöpft sich die Bedeutung von Art. 170 BV nicht im Evaluationsauftrag. Die Bestimmung geht von der weittragenden Voraussetzung aus, dass der Staat sich über die Wirkungen seines Handelns legitimiert. Der Auftrag zur Wirkungsprüfung ist nur sinnvoll, wenn die Outputlegitimation (Nützlichkeit) als verfassungsrechtlich anerkanntes Ziel neben die Inputlegitimation (demokratische und rechtsstaatliche Legalität) tritt.

Die Bestimmung errichtet implizit das Verfassungsprinzip der Effektivität staatlichen Handelns.

Die Beschlüsse der Behörden und der Vollzug durch die Verwaltung sollen nicht nur regelkonform, sondern auch im weitesten Sinne nützlich sein. Der Staat ist gehalten, seine Kompetenzen so zu nutzen, dass die von der Rechtsordnung gesteckten Ziele auch verwirklicht werden. Die Bundeskompetenzen und die bundesinternen Zuständigkeiten sind im Lichte dieses Wirkungsauftrags zu interpretieren.

Effektivität schliesst auch Effizienz ein. Sind die Strukturen, Verfahren und Handlungsweisen der Verwaltung ineffizient, so beeinträchtigt dies indirekt auch die Effektivität der staatlichen Massnahmen. Die Effizienz wird dadurch aber noch nicht zu einem Verfassungsprinzip. Vielmehr steht sie in einem dienenden Verhältnis zur Effektivität. Der Staat soll seine Wirkungsziele auf möglichst wirtschaftliche Weise erreichen. Die Frage nach der Effizienz hat somit Teil an der Pluralität politischer und rechtlicher Ziele staatlichen Handelns. Evaluationen dürfen sich daher nie auf blosse Effizienzfragen beschränken, sondern müssen die Wirkungsebene einschliessen. Ihre Fragestellungen müssen den Pluralismus der staatlichen Ziele einfangen.

Claudia Kaufmann:

« Quelle culture d'évaluation pour la Suisse ? »

Objectifs et problématique de la mise en oeuvre de l'art. 170 Cst. du point de vue de l'administration fédérale

- L'évaluation constitue un des éléments qui concourent à la légitimité de l'Etat prestataire de biens et de services.
- Cette fonction de l'Etat se base sur une législation dont le contrôle de la légalité (régularité) ne suffit plus. Elle demande une vérification par les effets.
- Les objectifs et les mesures d'une telle législation demandent une participation (une adhésion) de plusieurs partenaires.
- Les mesures agissent dans un contexte dynamique et le font évoluer.
- Chaque politique sectorielle a des effets sur d'autres domaines politiques (cohérence des activités étatiques).
- L'évaluation est un instrument de l'Etat démocratique: elle renforce la transparence des activités étatiques, elle est un médium entre la puissance étatique et les citoyens-nes.
- Elle est un enjeu politique.

- L'évaluation est d'abord une « préoccupation » de l'administration avant d'être une méthode ou avant de savoir sur quoi une évaluation doit porter.
- Même « préoccupation » que le Parlement ou les scientifiques ?

Oui, mais à partir de questions différentes (l'administration garante de l'intérêt général et de la mise en oeuvre des décisions du Parlement pouvoir discrétionnaire de l'administration)

Emmanuel Sangra:

Quels enseignements peut-on retenir au niveau fédéral de l'expérience genevoise?

L'exposé présente le modèle de la commission extraparlamentaire en tant qu'instance d'évaluation. Ce modèle est expérimenté depuis 1995 par le canton de Genève.

Assistée par un secrétariat professionnel, la commission externe d'évaluation des politiques publiques choisit ses propres thèmes d'évaluation, mais il lui arrive également d'être mandatée par le Conseil d'Etat ou par des commissions parlementaires. Ses résultats sont publiés. La commission est composée de seize membres qui sont choisis parmi la diversité culturelle et sociale du canton et font autorité dans le domaine de la gestion économique et politique. Ils n'appartiennent pas à l'administration.

Sous cette forme, la commission peut être considérée comme une instance pluraliste, mais qui n'intègre pas les acteurs mettant en oeuvre la politique évaluée. En tant qu'instance d'évaluation, elle allie ainsi plusieurs avantages en ce qui concerne l'indépendance et la qualité des résultats:

- indépendance dans le choix du thème, notamment possibilité de choisir un sujet que le gouvernement ou le parlement ne souhaitent pas forcément voir examiné,
- indépendance dans la fixation des questions d'évaluation, car le champ de l'évaluation n'est pas prédéfini par les acteurs concernés,
- indépendance dans le processus d'évaluation dont profitent également les éventuels mandataires de la commission,
- réunion des compétences apportées par les membres en matière de gestion publique et privée, ce qui contribue à la qualité des résultats, mais également à la pertinence des questions posées au départ de l'évaluation,
- réunion de sensibilités variées d'un point de vue culturel et social ce qui permet de mieux intégrer le contexte socio-politique d'une évaluation et qui renforce l'acceptabilité des résultats.

Sous cette forme et pour autant qu'elle ait suffisamment de moyens, une commission apparaît bien adaptée pour effectuer des évaluations récapitulatives « summative evaluations » qui permettent à des personnes étrangères au contexte du dispositif (autorités, usagers, médias, grand public, etc.) de mieux comprendre les effets d'une politique et, ainsi, de se faire une opinion sur des faits démontrés. Ce modèle permet de limiter au maximum les risques d'instrumentalisation de l'évaluation, que cela soit par l'autorité ou par l'administration en vertu de sa connaissance du terrain.

Dans le cadre de la réalisation de l'art. 170 de la constitution fédérale, il est essentiel de bien distinguer l'évaluation récapitulative, qui doit être confiée voire réalisée par un organe externe à l'administration, de l'évaluation endoformative, décidée par l'administration en vertu de sa propre logique d'action. Le but d'une évaluation conditionne ce qu'elle va apporter. Si l'aspect démocratique prédomine, une définition et une réalisation du mandat externe à l'administration est nécessaire. L'instance d'évaluation doit alors être en mesure de poser les questions d'évaluation pertinentes et de produire des résultats de qualité susceptibles d'être appliqués.

Rapports publiés par la commission externe d'évaluation des politiques publiques:

- Politique sociale du logement. Evaluation de l'encouragement à la construction selon la loi générale sur le logement, janvier 1997
- Politique de formation des personnes actives non qualifiées. Evaluation de la mise en oeuvre à Genève de l'article 41 de la loi fédérale sur la formation professionnelle, mai 1997
- L'Etat et ses contribuables. Evaluation des prestations de l'Administration fiscale cantonale aux yeux des contribuables, de leurs mandataires et de ses collaborateurs, septembre 1997
- Chômeurs en fin de droit. Evaluation de la politique cantonale d'emploi temporaire, septembre 1998
- Politique cantonale en matière de déductions fiscales: évaluation des déductions genevoises sous l'angle de leur impact financier, de leur vérification par l'administration et de l'égalité de traitement, décembre 1998
- Politique énergétique des Services industriels: évaluation de la mise en oeuvre des principes de la politique cantonale de l'énergie, mars 1999
- Politique cantonale d'éducation musicale: évaluation de l'impact des subventions aux écoles de musique, décembre 1999

- Subsidies en matière d'assurance-maladie: évaluation de la politique cantonale, février 2000
- Vitesse commerciale des TPG: évaluation des mesures d'accélération prises en tant que moyen de promotion des transports publics, octobre 2000

Groupe de travail 2:

Pour qui les évaluations doivent être mises en place?

(Présidence: Ernst Trachsler)

Res intellecta, in verbis simus facile

Schopenhauer, Eristische Dialektik

Si nous analysons les différentes étapes de l'évaluation, en débutant par le mandat en tant que tel et en débouchant sur la remise du rapport d'évaluation au mandant, il apparaît que la principale préoccupation de l'évaluateur est de savoir « pour qui les évaluations doivent être mises en place? ». Il est évident que le mandant lui-même se pose les mêmes questions, mais nous n'envisagerons ici que la perspective des mandataires, c'est-à-dire le point de vue des évaluateurs.

Trois questions principales sous-tendent finalement la problématique soulevée par le groupe de travail. C'est dans le cas des évaluations sommatives que ces trois questions se posent avec le plus d'acuité bien qu'il ne faille pas en sous-estimer l'importance dans les autres types d'évaluation.

Ainsi, quelle position prendre lorsque l'évaluateur qui agit, à travers son mandat, comme un catalyseur permettant à un projet (ou programme, stratégie etc.) évalué de s'améliorer, s'aperçoit d'un décalage entre les questions du mandat et les réels besoins des personnes faisant partie du projet (ou programme, stratégie etc.) que l'on évalue? Dans son souci légitime de valoriser l'évaluation, c'est-à-dire de lui donner une réelle utilité, se pose alors la question du choix de son bénéficiaire. Faut-il contenter les mandants ou bien répondre aux attentes provenant du « terrain »? Enfin se pose la question de la place de l'évaluateur dans ce contexte particulier.

Il apparaîtrait que, dans certains cas, le décalage qui existe entre les questions définies par le mandant et les attentes réelles des personnes travaillant sur le terrain serait le produit de la structure hiérarchique de l'administration, qui empêcherait au top management de cerner les véritables problèmes que rencontrent les personnes travaillant sur le terrain dans leur travail quotidien.

Toutefois une explication « culturelle » ou de pratique pourrait être envisagée. Aujourd'hui, dans l'administra-

tion, la culture de l'évaluation se résume généralement au double souci du respect du budget et de celui de la légalité. Or, une telle culture administrative ne peut que provoquer un décalage entre le top management et le « terrain », si ce n'est entraîner une véritable instrumentalisation de l'évaluation.

Afin de résoudre ce problème et d'arriver finalement au véritable objectif de l'évaluation, c'est-à-dire de faire évoluer, par l'apprentissage, un projet (ou programme, stratégie, etc.), il faut absolument que l'administration utilise l'outil d'évaluation avec une véritable culture de l'évaluation au sens, justement, d'outil d'apprentissage, et qu'elle fasse le deuil de ses vieux réflexes encore trop présents empêchant l'évaluation de déployer la globalité de ces effets.

Finalement, reste la question du positionnement de l'évaluateur. Force est de constater qu'il doit dire les faits, poser la situation et donner uniquement des recommandations intrinsèquement liées à ces faits. Quitte, alors, à renvoyer les questions « politiques » aux politiques.

Raphaël Audria

Collaborateur scientifique au Secrétariat général de la Ville de Genève

Arbeitsgruppe 3:

"Wer soll die Evaluationen durchführen?"

(Leitung: Kurt Grüter)

Ausgangslage

Der Auftrag der Arbeitsgruppe hat darin bestanden, Grundsätze, Leitlinien und Vorschläge zur Umsetzung von Art. 170 BV zu formulieren. Als Grundlage zur Diskussion wurde den TeilnehmerInnen Thesen abgegeben (siehe Beilage). Ein ausführlicherer Bericht der Arbeitsgruppe ist auf der Website der SEVAL zugänglich. Auf eine vertiefte Diskussion des Auftrags des Parlaments, institutionelle Vorschläge zur Umsetzung von Art. 170 BV vorzulegen (siehe Editorial von Prof. Mastronardi), ist mit Blick auf den Vorbehalt parlamentarischer Entscheidung und aufgrund der Abwesenheit einer parlamentarischen Vertretung in der Arbeitsgruppe verzichtet worden.

Hinsichtlich der rechtlichen Grundlagen zur Umsetzung von Art. 170 BV kann festgehalten werden:

- Art. 170 BV stellt keine Verpflichtung für irgendeine Stelle dar („Die Bundesversammlung sorgt dafür, dass die Massnahmen des Bundes auf ihre Wirksamkeit überprüft werden“).

- Bisher gibt es keinen Niederschlag im neuen Parlamentsgesetz. Überlegungen und Vorschläge der Aufsichtskommissionen liegen vor (siehe Editorial von Prof. Mastronardi).
- Es bestehen bereits andere Gesetze z.B. im Finanzbereich (Finanzkontroll-, Finanzhaushalts- und Subventionsgesetz), die als Grundlage einer Umsetzung des Verfassungsartikels dienen können.
- Zudem besteht eine Vielzahl von Evaluationsklauseln in Gesetzen. Deren Umsetzung ist bisher weitgehend an die Ämter delegiert.

Als zentrale Probleme der heutigen Situation sind genannt:

- Mängel einer effektiven Nutzung von Evaluationsergebnissen in der Gesetzgebung und im Vollzug;
- Fehlende Schwerpunktbildung in der Planung von Evaluationen auf allen Stufen des Bundes;
- Mangelnde Information bzw. Intransparenz der Durchführung von Evaluationen und der Nutzung von Evaluationsergebnissen;
- Unverbundenes Nebeneinander von Evaluationen und mangelnde Koordination zwischen verschiedenen Evaluationsstellen des Parlaments und der Bundesverwaltung;
- Probleme der Unabhängigkeit von Evaluationen als Mängel der Kooperation zwischen internen (Regierung und Verwaltung) und externen Evaluationsstellen (Parlament und Kontrollorgane); insbesondere Mängel der Unabhängigkeit bei Evaluationen mit gesetzlichem Auftrag (Evaluationsklauseln);
- Ungeklärtes Verhältnis zwischen Parlament, Kontrollorganen und Verwaltung bei der Umsetzung von Art. 170 BV (siehe Editorial von Prof. Mastronardi).

Ergebnisse der Diskussion:

Grundsätze, Leitlinien und Vorschläge zur Umsetzung von Art. 170 BV vorgeschlagen werden minimale Regelungen und Grundsätze auf Gesetzesstufe, eine Ausweitung bestehender Koordinationsverfahren sowie kulturelle, prozessuale und strukturelle Neuerungen. Auf Empfehlungen, wie sich das Parlament organisieren soll (siehe oben bzw. Editorial), wird verzichtet:

1. Evaluationen eines wirkungsorientierten Parlaments und einer wirkungsorientierten Verwaltung und Regierung dienen auf verschiedenen Stufen des Bundes unterschiedlichen operativen und strategischen, dezentralen und zentralen Informations-, Lern- und Kontrollbedürfnis-

sen. Aus dieser Perspektive sind entsprechende Freiheiten der Evaluation (pluralistische Evaluationskultur) bei der Umsetzung von Art. 170 BV zu gewährleisten.

2. Evaluationsklauseln sind im Rahmen der Gesetzgebung weiterhin zu verbreitern und zu verallgemeinern; In den Standards zu Botschaften des Bundesrats sollte ein Standard mit dem Inhalt, Evaluationsklauseln in Gesetzen vorzuschreiben aufgenommen werden;

3. Zwischen Evaluationen von Regierung und Management einerseits (Aufsicht, intern, formativ) sowie Parlament (GPK/PVK) und externen Kontrollorganen (EFK) andererseits (Oberaufsicht, extern, summativ) sollte die Koordination und Kooperation verstärkt werden;

4. Die Umsetzung von Art. 170 BV sollte sich an differenzierten Lösungen mit Blick auf unterschiedliche Stufen (Aufsicht, Oberaufsicht), funktionale Positionen (intern, extern) und Nutzungen (formativ, summativ) von Evaluationen orientieren;

5. Koordinations- und Kooperationsprozesse sollten auf der Grundlage von Pflichtenheften, differenziert nach unterschiedlichen Evaluationstypen und unterschiedlicher Relevanz von Evaluationen eingeführt werden; Die EFK kann in solchen Pflichtenheften als externe Beteiligte an Evaluationen der Verwaltung (z.B. mit der Funktion der Meta-Evaluation) beteiligt werden;

6. Evaluationen (Wirtschaftlichkeitsprüfungen) sollten seitens der EFK als externes Kontrollorgan im Rahmen der Oberaufsicht durchgeführt werden, insbesondere in überdepartementalen, risikoreichen, finanzrelevanten und wenig evaluierten Bereichen der Aufgabenerfüllung des Bundes; Meta-Evaluationen zur Überprüfung der Qualität und Nutzung von Evaluationen von Regierung und Verwaltung sollten durchgeführt werden;

7. Evaluationen seitens der GPK/PVK als externes Kontrollorgan sollten im Rahmen der Oberaufsicht des Parlaments durchgeführt werden;

8. Die Evaluationsplanung der Ämter und Departemente sollte in die bundesrätliche Legislaturplanung als zentrales Koordinations- und Legitimationsverfahren eingegliedert werden, dies, ohne die bestehende Pluralität der Evaluation(s)kultur im Sinne der grundsätzlichen Freiheit zur Evaluation zu beschränken;

9. Ein Koordinationsgesetz, das zu Koordinations- und Kooperationsprozessen im Bereich der Evaluation verpflichtet und das Instrument des Pflichtenhefts festlegt, sollte geschaffen werden;

10. An Evaluationspflichten im Rahmen des Finanzhaushalts- und Subventionsgesetzes sollte sich der Bund verstärkt orientieren;

11. Ein Kompetenzzentrum der Bundesverwaltung im operativen Bereich des Controlling, Coaching, fachliche Beratung, Herstellung von amts- und departementsübergreifender Information u.a.m. sollte geschaffen werden, dies zur Förderung und Unterstützung des Know-How und der Evaluationskultur (Übernahme von „Leadership“) in der Bundesverwaltung;

12. Die Evaluationsstandards der SEVAL sollten seitens der Evaluationsstellen des Bundes anerkannt werden. Die Durchführung von Evaluationen sollte sich an den entsprechenden Empfehlungen orientieren.

Arbeitsgruppe 4:

Welche Anforderungen stellt das Parlament an Evaluationen? (Leitung: Luzia Lehmann)

Die Parlamentarier haben ein Interesse zu begründen, dass gut ist, was sie tun, und wo es noch verbessert werden könnte. Dazu brauchen sie eine wissenschaftliche Untermauerung. Die kann nur von Evaluation kommen.

Wie die Evaluation gemacht wird, ist Sache der SpezialistInnen. Das Parlament will, dass das state of the art geschieht, kann und will dabei aber nicht mitreden.

Um die Qualität kontrollieren und garantieren zu können, müssen Standards vorhanden und generell anerkannt sein. Sie müssen aber sprachlich so abgefasst sein, dass sie für PolitikerInnen ohne weiteres verständlich sind. Kein Jargon.

Die seval-Standards können nur zum Teil erfüllt werden, vor allem was die Veröffentlichung von Evaluationsergebnissen angeht (setzt Genehmigung durch Kommission voraus). Aber das Parlament hat die Möglichkeit, die Qualität von Evaluationen durch zusätzliche Expertisen/Gutachten prüfen zu lassen (second opinion). Diese kann, muss aber nicht, aus der Verwaltung kommen.

Der Regelkanon der wissenschaftlichen Evaluation ist vorwiegend anhand kleiner, gezielter Themenbereiche entstanden. Für den Einsatz bei grösseren Themen (z.B. Gesetze) ist er nicht durchwegs optimal, noch auszubauen.

Die Aufgabe des Parlaments ist zu 80% erfüllt, wenn die Ziele der Evaluation klar definiert sind. Diese Zieldefinition passiert in einem Aushandlungsprozess zwischen politischer und wissenschaftlicher Rationalität.

Um das spezifische Know-how des Parlaments in diesem Bereich zu sichern und zu akkumulieren, ist die Einrichtung eines parlamentarischen Kompetenzzentrums denkbar, das Evaluationen in Gang setzt, verfolgt (Monitoring) und valorisiert.

Von seiten des Parlaments tritt eine Kommission als Auftraggeber von Evaluationen auf. Innerhalb der Kommission findet aber ein kollektiver Aushandlungs- und Entscheidungsprozess statt. Dessen Ergebnis ist in der Regel breit abgestützt (z.B. GPK NR 25 Mitglieder). In diesem Sinne ist ein parlamentarischer Evaluationsauftrag demokratisch legitimiert.

Ein Verwaltungsauftrag dagegen entspringt in der Regel dem Willen einer hierarchisch übergeordneten Stelle und kann deshalb tendenziell eher partikulären Interessen dienen: rationalisme pluraliste vs. rationalisme hiérarchique. Parlamentarische Meta-Evaluationen verwaltungsgesteuerter Evaluationen können partikularistische Verfälschungen nachweisen, ev. auch relativieren.

Im Bereich von Regierung und Verwaltung gibt es Evaluationen, sie folgen aber vielen verschiedenen Modellen und "Kulturen".

Theo Haldemann

Führen mit Evaluationsergebnissen

Podiumsdiskussion unter Leitung von Sylvia Furrer

Die vier Regierungs- und Parlamentsmitglieder aus den Kantonen Luzern, Solothurn und Wallis haben folgende Erfahrungen mit der Evaluationsmethode gemacht:

- Die Evaluationen waren unbedingt erforderlich, weil die Stärken und Schwächen der NPM- und WOV-Reformen, der aktuelle Handlungsbedarf und die weitere Entwicklungsrichtung systematisch und vollständig aufgezeigt wurden. So wurden z.B. das neue Rollenverständnis von Regierung und Parlament sowie das notwendige Einüben der neuen Zusammenarbeitsformen erst in der Evaluation richtig ersichtlich.
- Die Schlussfolgerungen der Erfolgs- und Wirkungskontrollen hätten den politischen Entscheidungsträgern in Regierung und Parlament noch klarer formuliert und noch fordernder vermittelt werden sollen, damit die Umsetzung der Evaluationsempfehlungen befördert wird. So hätte z.B. der Nachholbedarf bei der Aus- und Weiterbildung der Politikerinnen und Politiker viel aufrüttelnder und schonungsloser dargestellt werden müssen.
- Für die drei Projektleiter aus den Kantonen Aargau, Luzern und Thurgau ist das Instrument der Fremdevaluation zu langsam, um im Reformprozess kontinuierlich eingesetzt werden zu können, z.B. beim Erreichen wichtiger Meilensteine.

Fortsetzung der Debatte: "Welche Evaluationskultur braucht die Eidgenossenschaft"

P. Walther, A. Kühne

Abduktion – Beitrag zur Evaluationskultur

Stichwort: Evaluationskultur

Betrachtungen zur Evaluationslandschaft in der Schweiz – allerdings aus einer sehr summarischen Warte – zeigen, dass es heute zwei Tendenzen gibt, die sich - pointiert - folgendermassen charakterisieren lassen:

(1) Die „WissenschaftlerInnen“: Evaluationen sollen Hypothesen im Rahmen von Wirkungsmodellen empirisch (vorwiegend mit quantitativen Methoden) überprüfen und damit die Wirkung von Projekten und Programmen aufzeigen. Resultate sind in erster Linie empirische Studien, deren Schlussfolgerungen durch die Programmleitungen umzusetzen sind.

(2) Die „Prozess-Orientierten“: Evaluation wird als Prozess zwischen beteiligten Akteuren verstanden. Es wird viel mit qualitativen Methoden (Tiefeninterviews, Beobachtungen im Gelände) gearbeitet. Empfehlungen und Schlussfolgerungen werden zusammen mit den zentralen Akteuren - z.B. mittels einer Expertengruppe, die sich in der Auseinandersetzung mit dem Thema ein Urteil gebildet hat - formuliert.

Bei Definitionsversuchen des Begriffs der Evaluation treten beide Dimensionen (Evaluation als Studie und als Prozess) oft unvermittelt nebeneinander auf. Insbesondere die Diskussion um den Ansatz der „utilization focused evaluation“ (Patton¹) hat die Diskussion vom wissenschaftlichen Modell (1) zum partizipativen Modell (2) hin verschoben.

Unserer Ansicht nach ist es sinnvoll, sich einmal mit den Wurzeln der Pragmatik (Philosophen wie Pierce, James) sowie, in deren Folge, mit den (pragmatischen) Erkenntnisgrundlagen der interpretativen Sozialwissenschaften (Soziologen wie Strauss, Soeffner, Reichertz) zu beschäftigen. Und dies führt zur überraschenden Erkenntnis, dass deren Arbeiten zur Logik der Abduktion unter Umständen für den Aufbau einer Evaluationskultur und zur Ueberbrückung der Differenzen zwischen den beiden Positionen (1) und (2) relevant sind.

Was ist Abduktion?

Der amerikanische Philosoph Peirce beschäftigte sich vor über 120 Jahren mit dem Problem, nach welcher Logik neue Ideen, eine Hypothese, oder ein Urteil über eine Situation entsteht². Er bezeichnete Abduktion als das

„Herzstück“ der Pragmatik. Seine Arbeiten sind insbesondere im Bereich des Designs von Expertensystemen (künstliche Intelligenz) wieder relevant geworden. Gemäss Peirce gibt es neben der Deduktion (Schlussfolgerungen mit theoretischer Sicherheit) und der Induktion (Schlussfolgerungen mit empirischer Wahrscheinlichkeit) eine dritte logische Operation, die zu Schlussfolgerungen mit praktischer Plausibilität führen. Er nannte die Operation Abduktion.

Der abduktive Schritt ist der folgende:

Regel: Alle Bohnen aus diesem Sack sind weiss.

Beobachtung: Diese Bohne ist weiss.

Schlussfolgerung: Diese Bohne ist (aus mehr oder weniger plausiblen Gründen) aus diesem Sack.

Abduktion ist die Logik der Plausibilität. Sie spielt zum Beispiel in der medizinischen Diagnose eine zentrale Rolle. Der Arzt hat ein bestimmtes Grundwissen (Theorie), er beobachtet Symptome an PatientInnen, und er fällt dann ein Urteil über den Einzelfall (die Diagnose). Dieses Urteil ist weder sicher (deduktiv) noch wahrscheinlich (induktiv). Es ist plausibel.

Peirce beharrte darauf, dass die Geburt einer neuen Idee nicht nur mit psychologischen Phänomenen oder Intuition erklärt werden kann³. Es gibt immer schlechte oder gute Gründe für eine Hypothese. Nach seiner Auffassung ist „reasoning from a hypothesis“ (Induktion) grundsätzlich verschieden von „reasoning towards a hypothesis“ (Abduktion).

Auch in den interpretativen Sozialwissenschaften wurde die Abduktion als leitende Erkenntnislogik eingeführt⁴. Damit ist ein Vorgang gemeint, der auf der Grundlage gewisser theoretischer Annahmen über einen Gegenstand zu Beobachtungen kommt, welche die Annahmen weder falsifizieren noch verifizieren, sondern weiterentwickeln und sich so den vermeintlichen, d.h. immer bereits gesellschaftlich vorinterpretierten Bedingungen, Strategien und Erfolgsaussichten von Handeln annähern. Abduktion ist hier ein kreativer, theoriebildender Prozess, der bis zu einem gewissen Grad von empirischen Beobachtungen begleitet ist. Die entsprechende Theorie zielt auf die handlungsrelevanten, oft impliziten Annahmen der handelnden Akteure. Abduktion ist damit sowohl ein alltagspraktischer Erkenntnisvorgang, mit welchem wir unsere Alltagsentscheidungen treffen, wie

ein sozialwissenschaftlicher Erkenntnisvorgang, mit welchem Alltagsentscheidungen handelnder Akteure rekonstruiert werden können.

Was hat Abduktion mit Evaluation zu tun?

Unserer Ansicht liegt diese Logik der Abduktion auch Evaluationen zu Grunde. Es geht darum, plausible Urteile über Projekte und Programme und deren Wirkungen zu fällen. Die Hypothesen und die impliziten Wirkungsmodelle von Projekten und Programmen müssen nicht nur überprüft (induktiv) sondern auch rekonstruiert (abduktiv) werden⁵.

Evaluationen sind erst dann interessant, wenn bestehendes Wissen erweitert oder überschritten⁶ und damit ein Beitrag an die Verbesserung eines Projektes oder Programmes geleistet wird. Dies ist das abduktive Fundament von Evaluationen. In Anlehnung an den Titel dieses Aufsatzes kann man daher zum Schluss kommen, dass Abduktion in Evaluationen noch zu oft „logica utens“ (basierend auf Instinkt) und zu wenig „logica docens“ (kritische Logik) ist.

Gemäss Peirce ist Abduktion die kreativste Form der Logik. Beispiele für abduktive Regeln sind nach Peirce: (1) es ist unzweckmässig, eine Hypothese vorzuschlagen, die empirisch nicht überprüfbar ist; (2) es ist nicht zweckmässig, Hypothesen vorzuschlagen, die auf einfachere Sätze reduziert werden können; (3) Hypothesen sollten an Extrem- und Gegenbeispielen überprüft werden; und (4) Hypothesen sollen nicht in einem zu grossen Widerspruch mit der Alltagserfahrung („common sense“) stehen.

Regel (3) kann natürlich als ein Votum für den Einsatz von qualitativen Methoden in Evaluationen aufgefasst werden. Regel (4) schlägt eine Brücke zur Forderung, dass die Ergebnisse einer Evaluation umsetzbar sein sollten.

Insbesondere mit der Weiterentwicklung des Evaluationsparadigmas hin zur verwendungs- bzw. nutzungsorientierten Evaluation (Patton) gewinnt die Abduktion ihren bisher wenig reflektierten Stellenwert. So stellt etwa die Rekonstruktion von Wirkungsmodellen und Kausalzusammenhängen in Programmen im Rahmen der Programmtheorie⁷ erkenntnistheoretisch eine aktuelle Anwendungsweise der Abduktion dar.

Relevant für diesen Ansatz ist die Beteiligung der Akteure an zentralen Stellen einer Evaluation (Entwicklung der Evaluationsfragen, Bestimmung des Nutzens

einer Evaluation, Interpretation von Evaluationsergebnissen). Das entsprechende Verhältnis zwischen Evaluator und Evaluiertem ist die Interaktion und der Dialog, in welchem weder deduktive, noch induktive Schlussfolgerungen, sondern das gemeinsame Entwickeln und Testen von Interpretationen auf der Grundlage ihrer Plausibilität und mit manchmal unvorhergesehenen neuen Sichtweisen (Innovationen) im Zentrum stehen.

Fazit: Bausteine zu einer Evaluationskultur

Unserer Ansicht nach könnte die Belebung der Diskussion dieser diagnostischen Komponente von Evaluationen und der zugrundeliegenden Logik der Plausibilität (Abduktion) einen interessanten Beitrag an den Aufbau einer Evaluationskultur in der Schweiz liefern. Die Diskussion zeigt einen interessanten Mittelweg zwischen den obgenannten beiden Positionen.

„Aus möglichst vielen Beobachtungen ist das Gebäude der Diagnose aufzubauen“⁸; „Therapie ohne Diagnose sollte im allgemeinen unter Strafsanktionen gestellt werden“; oder „Die Diagnose ist an vielem Uebel schuld“⁹. Diese Zitate aus der „logisch verwandten Disziplin“ der medizinischen Diagnostik können vielleicht zeigen, welchen Stellenwert Evaluationen in der Politik und insbesondere in der Formulierung neuer Politiken einnehmen könnten.

Ein Mittelweg zwischen den eingangs genannten Positionen, der die Erkenntnisse der „Pragmatiker“ mit einbezieht, könnte etwa folgendermassen aussehen:

1. Evaluationsstellen nehmen die Chance, dass Evaluationen nicht einfach nur Ueberprüfungen von Hypothesen sind, sondern vielmehr der Logik der Plausibilität unterworfen sind, zur Kenntnis. Sie messen den Wert einer Evaluation an Kriterien wie Plausibilität der Ergebnisse, Erkenntnisgewinn und/oder „Wert der Diagnose für die Problemlösung“.

2. Wissenschaftliche Studien (mit induktiven Methoden) sind sehr wertvoll und dienen Evaluationen als sachlich-systematische Grundlagen, doch sind sie letztlich nicht mehr als „Beobachtungen am Patienten“ (siehe obige Metaphor der medizinischen Diagnose). Die Evaluation ist oft erst der Schritt danach, die Bildung eines plausiblen Urteils über die Situation.

3. Die Mobilisierung von Theorie und auch von Wissen über die politische Machbarkeit (von „Therapien“) sowie die Formulierung von plausiblen Schlussfolgerungen

sollte nicht allein der induktiv arbeitenden Wissenschaft sondern einer Evaluationsgruppe, die Erfahrungen in der Praxis hat und eine plausible Diagnose fällen kann, übergeben werden.

4. Nicht Punkt 2 sondern Punkt 3 ist die eigentliche Evaluation¹⁰, die zu plausiblen Ergebnissen und umsetzbaren Empfehlungen führt. In der Evaluationsgruppe sollen 2-3 Experten und mindestens ein betroffener Akteur Einsitz nehmen. Sie soll auf der Grundlage eines Pflichtenheftes einen kurzen Evaluationsbericht redigieren, dessen Schlussfolgerungen (die Diagnose) plausibel, und dessen Empfehlungen (die Therapie) umsetzbar sind.

Wir sind der Ansicht, dass dieser Vorschlag die Brücke zwischen den eingangs erwähnten Positionen (1) und (2) schlägt und die Diskussion darüber bereichert, wie Evaluationen im Spannungsfeld von Wissenschaft, Praxis/Politik und Öffentlichkeit zu definieren sind.

Abduktion ist dabei nicht eine Relativierung sondern letztlich ein Bekenntnis zur Wissenschaftlichkeit von Evaluationen. Pierce versuchte einfach aufzuzeigen, dass sich die Regeln der Plausibilität (Abduktion) grundsätzlich von der Induktion und der Deduktion unterscheiden. Dem Aufbau einer Evaluationskultur wurde damit ein grosses Tor geöffnet.

Kontakt:

Dr. P. Walther, EPM Walther Consultants,
Burgernzielweg 12, 3006 Bern;

Tel 0041 31 352 01 27; eMail pwalther@dial.eunet.ch

Dr. A. Kühne, Eidgenössische Finanzkontrolle,
Monbijoustr. 51a, 3003 Bern; Tel: 0041 31 323 10 53;
eMail: armin.kuehne@efk.admin.ch

¹ Patton, M.Q., 1997: Utilization-focused Evaluation. London.

² Vgl. Fann, K.T., 1970: Peirce's logic of abduction; oder Peirce, Ch., 1901: The nature of Hypothesis; gedruckt in Thomas, V., 1957: Essays in the philosophy of science. Indianapolis: 235-255.

³ Vgl. eine interessante Analyse der Positionen von Popper und Schiller/Reichenbach in Almeder, R., 1980: The philosophy of Charles S. Peirce: a critical introduction. Oxford.

⁴ Die Chicago School of Sociology (30-er-Jahre) berief sich dabei direkt auf Peirce.

⁵ Siehe auch Pawson, R. und Tilley, N., 1997: Realistic Evaluation. London.

⁶ in Anlehnung an das berühmte Zitat von Leibniz: Wissenschaft ist erst interessant, wenn Wissen überschritten wird.

⁷ siehe auch Leeuw, F., 2000: Analyzing Program Logics. Working Paper für den Jahreskongress der EES 2000 in Lausanne.

⁸ Rudolph, G., 1978: Zeitgedanken der Diagnostik und Semeiotik in der französischen Medizin des 18. und 19. Jahrhunderts; in Habrich, C. et al, 1978: Medizinische Diagnostik. München: S. 274

⁹ beide: Schweninger, E., 1926: Der Arzt. Radeburg: S. 128).

¹⁰ In der internationalen Praxis hat sich bewährt, klar zwischen „Background Studies“ und „Evaluationen“ zu unterscheiden. Evaluationsmissionen operieren auf der Basis eines Pflichtenheftes und haben einen mehr oder weniger offiziellen Status. Studien werden als „Resource Material“ betrachtet, auf das sich das Evaluatorenteam stützt.