

# Evaluationssynthesen: Bedingungen, Möglichkeiten und Grenzen

Andreas Balthasar | *Das Bedürfnis, staatliche Massnahmen auf ihre Wirksamkeit zu überprüfen, ist gross. Evaluationen müssen aber nicht nur wissenschaftlichen Ansprüchen genügen, sie sollen auch praktischen Nutzen haben. Evaluationssynthesen sind ein möglicher Weg.*

Seit dem Start des Nationalen Forschungsprogramms 27 «Wirksamkeit staatlicher Massnahmen» sind gerade zehn Jahre vergangen. Seither sind zu verschiedenen Politikbereichen zahlreiche Evaluationen erarbeitet worden. Mit der Zunahme der Zahl der Evaluationen von einzelnen Massnahmen ist ein Bedürfnis nach einer Synthese der Evaluationsergebnisse entstanden. Dabei zeigte sich relativ bald, dass gehaltvolle Synthesen nur möglich sind, wenn die einzelnen Massnahmenevaluationen in ein Gesamtkonzept eingebunden sind. Auf welchen Grundpfeilern muss nun ein Evaluationskonzept aufgebaut sein und welche Art von Synthesen sind möglich?

Mit diesen Fragen beschäftigt sich der vorliegende Artikel.<sup>1</sup> Im ersten Abschnitt werden drei Grundpfeiler der Gestaltung umfassender Evaluationssysteme vorgestellt. Danach werden im zweiten Abschnitt fünf Konsequenzen diskutiert, die sich daraus ergeben. Zum Abschluss werde ich mich der Thematik der Evaluationssynthesen zuwenden. Ich werde die Behauptung zu begründen versuchen, dass nützliche Evaluationssynthesen zwar möglich, aber nicht wissenschaftlich ableitbar sind.

## 1 Prinzipien der Gestaltung umfassender Evaluationssysteme

Die Frage, ob das systematische Ermitteln der Wirksamkeit staatlicher Massnahmen notwendig ist, wurde nicht erst mit Artikel 170 der neuen Bundesverfassung beantwortet. Dieser hält fest, dass die Bundesversammlung dafür sorgt, dass die Massnahmen des Bundes auf ihre Wirksamkeit überprüft werden. Schon vorher verfolgten immer mehr Bundesstellen eine eigentliche Evaluationspolitik. Die Direktion für Entwicklung und Zusammenarbeit, die Parlamentarische Verwaltungskontrolstelle, die Verwaltungskontrolstelle des Bundesrates, die Eidgenössische Finanzkontrolle, das Bundesamt für Energie, das Bundesamt für Justiz (Opferhilfe), das Bundesamt für Umwelt, Wald und Landschaft, das Bundesamt für Landwirt-

schaft, das Bundesamt für Sozialversicherung, das Bundesamt für Bildung und Wissenschaft, das Bundesamt für Berufsbildung und Technologie, das Bundesamt für Gesundheit, das Bundesamt für Wirtschaft und Arbeit sowie der Schweizerische Wissenschaftsrat und auch andere Dienststellen des Bundes lassen regelmässig evaluieren. Auch auf kantonaler Ebene erweiterte sich der Einsatz von Evaluationen in den letzten Jahren stark (Busmann 1996).

Manche Bundesämter, wie beispielsweise das Bundesamt für Energie, Sozialversicherung, für Wirtschaft und Arbeit oder für Gesundheit, haben Jahresprogramme für die Evaluation entwickelt, die Budgets und spezifische Zielsetzungen enthalten. An diesen orientieren sich die Ausschreibungen einzelner Evaluationsprojekte. Das Anliegen besteht darin, nicht einfach eine mehr oder weniger zufällige Abfolge einzelner Massnahmen zu evaluieren. Vielmehr sollen – ausgehend von einer ganzheitlichen Perspektive – Evaluationen einzelner Massnahmen in ein Konzept eingebettet sein, das die Inhalte der Evaluationstätigkeiten im zeitlichen Ablauf festlegt, Schwerpunkte setzt, institutionelle Vorgaben bestimmt sowie Koordinations- und Synergiepotenziale aufzeigt. Auf dieser Basis sollen die Evaluationen effizient, lernorientiert und unter Vermeidung von Doppelspurigkeiten durchgeführt werden. Das Bundesamt für Energie verfügt seit 1992 über ein solches Konzept, das Bundesamt für Sozialversicherung seit 1996, das Bundesamt für Wirtschaft und Arbeit seit 1998 (BFE 1992, BSV 1998; Balthasar 1998).

Die Grundlage dieser Evaluationssysteme bilden die Erfahrungen aus dem In- und Ausland. Diese beziehen sich allerdings meist auf die qualifizierte Durchführung einzelner Evaluationen. Die Anforderungen an Evaluationssysteme wurden bisher nicht systematisch thematisiert.

Aus den verfügbaren Quellen lassen sich in meinen Augen aber drei Grundsätze herauskristallisieren, die beim Aufbau umfassender Evaluationssysteme zu berücksichtigen sind.<sup>2</sup> Es handelt sich erstens um die Notwendigkeit, an komplexe Materien projektorientiert heranzugehen. Zweitens müssen Evaluationen prozessorientiert angelegt sein. Und drittens müssen sie auf einem partizipativen Ansatz aufbauen. Den drei Grundsätzen kann nur nachgelebt werden, wenn die Evaluationsthematik kontinuierlich verfolgt wird. Dazu ist eine stabile institutionelle Verankerung der Evaluation im betreffenden Politikbereich notwendig. Im Folgenden gehen wir auf die drei genannten Aspekte näher ein.

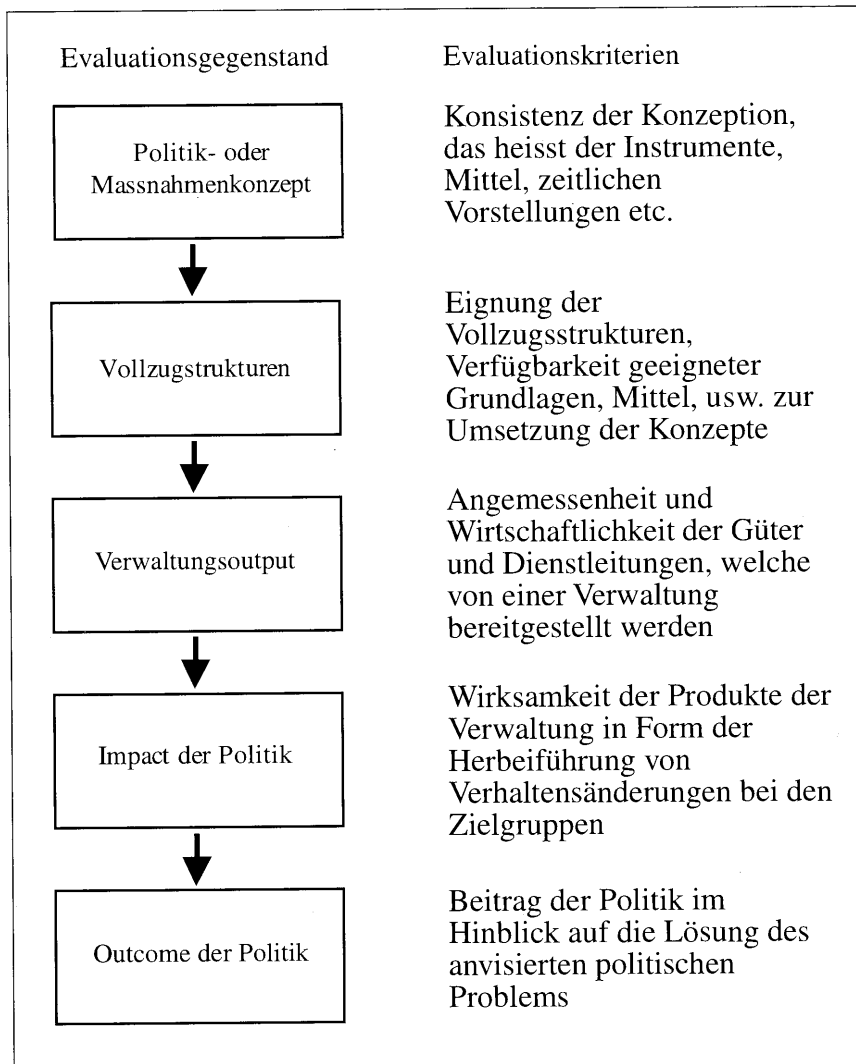
## 1.1 Projektorientierte Herangehensweise

Gute Evaluationen müssen nicht nur wissenschaftlich korrekt gemacht, sondern auch praktisch nützlich sein (Widmer 1996). Diese anerkannte Qualitätsanforderung lässt sich nur einlösen, wenn sich Wirkungsuntersuchungen auf präzise, sorgfältig abgegrenzte und auf Grund der vorhandenen oder zu erhebenden Daten gut bearbeitbare Fragestellungen konzentrieren. Bei der Planung der Evaluation von vielschichtigen Politikbereichen sind demnach verschiedene kleinere Projekte zu formulieren. Der direkte Weg über eine einzige «Gesamtevaluation» hat sich als nicht erfolgversprechend herausgestellt. Dies haben zum Beispiel die Auseinandersetzung mit der Evaluation in der Regionalpolitik oder auch die Vorbereitung des Aktionsprogramms Energie 2000 deutlich gemacht (Balthasar 1998; Linder/Rieder/Varone 1992). Beim Start von Energie 2000 stand die systematische Verfolgung einer immensen Zahl von Indikatoren zur Diskussion, welche alle Bereiche von Energie 2000 hätten abdecken sollen.

In ihrem Bericht zur Evaluierbarkeit des Programms haben Wolf Linder und seine Mitarbeiter Stefan Rieder und Frédéric Varone aufgezeigt, dass der Weg über Projektevaluationen eingeschlagen werden muss. Sie argumentieren, dass der Aufbau von Energie 2000 – wie die meisten Politiken – keiner programmeigenen, organisatorischen oder technischen Logik folgt. Es gibt nur die politische Logik, welche die vorgesehenen staatlichen Aktivitäten bestimmte. Auf Grund dieser Ausgangslage war nicht zu erwarten, dass ein Muster oder ein allgemein anwendbarer Ansatz für eine Programmevaluation auf der Makroebene zu finden war: «Da die Programmelemente aus politischen Argumenten angehäuft wurden, ohne Rücksicht auf ihre Beziehungen und ihre Wechselwirkungen (z. B. Zielkonflikte, Substitutionseffekte), ist keine ganzheitliche Evaluation des Aktionsprogramms durchführbar.» (Linder/Rieder/Varone 1992, 15) Als praktikabel wurden demgegenüber Evaluationen der Wirkungen von einzelnen Massnahmen erachtet. Diese bilden die Mosaiksteine, welche am Schluss zu einem ganzen Bild zusammengefügt werden müssen. Ob und wie dies gelingen kann, haben die Autoren allerdings offen gelassen. Es wird darum darauf zurückzukommen sein.

## 1.2 Prozessorientierter Ansatz

In der in kürzester Zeit zum Standardwerk gewordenen «Einführung in die Politikevaluation» unterscheiden Peter Knoepfel und Werner Bussmann verschiedene Evaluationsgegenstände entlang dem Policy-Cycle (Bussmann/Knoepfel 1997). Die darin angelegte dynamische und prozessorien-



*Darstellung: Evaluationsgegenstände und Evaluationskriterien, vereinfacht nach Bussmann/Knoepfel, 1997, S. 70*

tierte Betrachtungsweise ist erfahrungsgemäss ganz wesentlich für die Anlage komplexer Evaluationssysteme. Vereinfacht lassen sich insbesondere fünf Evaluationsgegenstände unterscheiden, die zeitlich nacheinander stehen. Sie können als Hürden aufgefasst werden, die übersprungen werden müssen, wollen Politiken wirksam sein. Jedem dieser Hindernisse muss sich die Evaluation zuwenden.

Die erste Hürde bildet die Qualität der Konzeption einer Politik. Evaluationen müssen prüfen, ob eine Politik so angelegt ist, dass sie überhaupt zielwirksam sein kann. Im Kontext des Krankenversicherungsgesetzes stellte sich beispielsweise die Frage, ob der Risikoausgleich zwischen den Krankenkassen einen Beitrag zum Ausgleich der Prämienunterschiede einzelner Kassen leisten kann oder ob andere Faktoren, wie beispielsweise unterschiedliche Alters- und Geschlechtsstrukturen, zu berücksichtigen sind (vgl. BASS 1999).

Die zweite Gruppe von Evaluationsgegenständen thematisiert die Vollzugsebene. Dort wird unter anderem danach gefragt, ob und wie eine konzipierte Massnahme durch die Verwaltung implementiert wird. Insbesondere steht die Wirksamkeit des Vollzugs zur Debatte. Die Auswertung der Datenbank für die Evaluationen in der Energiepolitik zeigte, dass die Frage des Vollzugs in mindestens der Hälfte aller Untersuchungen gestellt wurde (Balthasar 2000).

Die dritte Gruppe von Untersuchungsgegenständen kreist um die Leistungen der Verwaltung, das heisst um den so genannten Output. Auch wenn Leistungen nachweisbar sind, können die Wirkungen noch auf sich warten lassen. Zuerst muss das Zielpublikum die Massnahme kennen und akzeptieren. Die Bekanntheit von Unterstützungsangeboten ist gerade in der Regionalpolitik eine immer wieder aufgeworfene Frage (Hanser/Kuster/Cavelti 1994).

Erst wenn die Leistungen dem Zielpublikum im erwarteten Ausmass zukommen, sind zielrelevante Wirkungen oder so genannte Impacts zu erwarten. Damit beschäftigt sich die vierte Gruppe von Evaluationsgegenständen. Impact-Studien fragen nach dem Verhalten der Zielgruppen, deren Handeln durch das Gesetz direkt beeinflusst wird. Zentral ist in diesem Zusammenhang die Frage nach dem Anteil des Mitnahmeeffektes, also jener Fördermittel, die ausgeschüttet werden, ohne dass spezielle Leistungen dafür erbracht werden (Dreher/Balthasar 1997).

Die Politik will Verhaltensänderungen bei beteiligten Akteuren herbeiführen, um zentrale politische Ziele zu erreichen. Die letzte Gruppe von Evaluationsgegenständen kreist um die gewünschten gesellschaftlichen

Veränderungen, das heisst um die so genannten Outcomes. Ein Beispiel dafür stellt die Evaluation des Investitionsprogrammes dar, welche auch nach den energetischen und wirtschaftlichen Effekten des Programms fragte (Rieder/Iten 1999).

Der Erfolg einer Politik hängt davon ab, dass alle Stufen des Policy-Prozesses erfolgreich durchlaufen werden. Die Evaluation muss dabei Fragen auf allen Stufen stellen. Zudem ist zu beachten, dass je nach Zeitpunkt unterschiedliche Themen evaluierbar sind. Dies machte eine ganzheitliche und langfristig orientierte Sichtweise notwendig.

### **1.3 Partizipativer Ansatz**

Evaluationen verlangen eine gewisse Distanz zwischen Evaluierenden und Evaluierten. Nur dadurch ist eine unabhängige Beurteilung der Wirksamkeit staatlicher Aktivitäten möglich. Neuere Untersuchungen weisen auch darauf hin, dass Wirkungen von schriftlichen Berichten nicht unterschätzt werden dürfen. Aus diesem Grund ist unbedingt auf die formale Qualität von Berichten zu achten. Diese können zu umfangreich, zu deskriptiv oder zu theoretisch abgefasst sein und dadurch die Umsetzung der Ergebnisse behindern. Obwohl gute Berichte weiterhin wichtig sein werden, ist klar, dass Evaluationen vielfach nicht nur auf Grund ihrer Schlussberichte Wirkungen haben. Oft sind die Prozesse, die im Zuge der Untersuchungen eingeleitet werden, mindestens ebenso bedeutsam. Vielfach führt erst eine Evaluation dazu, dass schon vor ihrem Abschluss relevante Vollzugs- und Wirkungsindikatoren systematisch erhoben werden. Auch lässt sich beobachten, dass Ziele neu und zum Teil präziser formuliert werden.

Damit solche Lerneffekte möglich sind, müssen Evaluationen partizipativ angelegt sein. Konkret bedeutet dies, dass die beteiligten und die betroffenen Akteurguppen (z. B. andere Bundesstellen, Kantone, Private) in die Planung und Durchführung einer Evaluation einbezogen werden müssen. Damit wird ein optimaler Informationsaustausch aufgebaut und ein Interesse für die Resultate der Evaluation geschaffen. Dies stellt einerseits sicher, dass alle für die Durchführung der Evaluation relevanten Informationen zur Verfügung gestellt werden. Andererseits ist dieses Interesse die Voraussetzung dafür, dass Ergebnisse später auch in der einen oder anderen Form umgesetzt werden. Nur wenn man weiss, wie etwas zu Stande gekommen ist, glaubt man es auch. Bei der Analyse der Wirkung der Prämienverbilligung wurde im Herbst 1997 beispielsweise ein Workshop zusammen mit den Kantonen durchgeführt. Dort zeigte sich, dass die Vorbehalte, welche das Bundesamt gegenüber der zu diesem Zeitpunkt vorliegenden Studie

hatte, von den Kantonsvertretungen in weiten Teilen geteilt wurden. Die Kantone machten bei dieser Gelegenheit einen Vorgehensvorschlag, den das BSV umsetzen liess. Die Ergebnisse wurden im Januar 1999 veröffentlicht. Sie zeigen ein Bild, das die Kantone nicht mehr in Zweifel zogen (Balthasar 1999). Einzelne Kantone haben sogar reagiert und die Kriterien zur Berechnung der Prämienverbilligung angepasst. Grundsätzlich sind bei Evaluationen Überraschungen zu vermeiden. Die Verantwortlichen sollten auf Grund kontinuierlicher und guter Information über den Prozess der Erfolgskontrolle am Schluss das Gefühl haben, dass sie die Probleme schon gekannt haben. Dann ist es für sie leichter, diese anzugehen. Kurze, regelmässige Feedbacks zwischen Erfolgskontrolle und Tagesgeschäft sind unabdingbar, damit Verbesserungen sofort an die Hand genommen werden können.

Die genannten Prinzipien zur Gestaltung komplexer Evaluationssysteme sind im Grundsatz unbestritten. Dies hängt auch damit zusammen, dass man sich ihrer Konsequenzen im Einzelnen zu wenig bewusst ist.

## **2 Konsequenzen aus den Prinzipien**

Die genannten Prinzipien umfassender Evaluationssysteme haben verschiedene Konsequenzen mit durchaus bitterem Nachgeschmack.

### **2.1 Hohe finanzielle Aufwendungen**

Die projektorientierte Herangehensweise mit ihrer Konzentration der Evaluation auf einzelne, klar abgegrenzte Fragestellungen macht zahlreiche einzelne Untersuchungen notwendig. Dies hat auch finanzielle Folgen. Die bisherigen Erfahrungen zeigen, dass komplexe Materien nur mit erheblichen Mitteln zuverlässig evaluiert werden können. In der Regel sind rund zwei Prozent der Budgetmittel für die Evaluation zu reservieren.

Aus Sicht der Evaluation scheint dies zunächst unproblematisch, ja sogar erfreulich. Das Personal der öffentlichen Verwaltungen sehen jedoch sofort, dass die notwendigen Mittel vielfach nicht verfügbar sind. Die Folge davon ist, dass entweder auf Evaluationen ganz verzichtet wird, dass die Budgets für Einzelprojekte unrealistisch tief angesetzt werden oder – und das scheint mir der einzig richtige Weg – dass Priorisierungen vorgenommen werden. Dies ist jedoch eine politisch brisante Übung, weil damit bereits Bewertungen vorweggenommen werden. So argumentieren Verwaltungsangehörige oft, den Vollzug müsse man nicht untersuchen, er sei in Ordnung oder seine Schwächen seien bekannt und Verbesserungen im Gange. Aussenstehende sehen dagegen gerade im Vollzug die grössten Probleme bei vielen Politiken.

## 2.2 Ganzheitliches Bild

Das zweite Problem, das im Zusammenhang mit der projektorientierten Herangehensweise besteht, betrifft die Gefahr, dass durch die Konzentration auf einzelne Fragestellungen das Ganze aus dem Blickfeld gerät. Dem ersten Konzept zur Evaluation des Krankenversicherungsgesetzes wurde nicht zu Unrecht vorgeworfen, dass ein ganzheitlicher Ansatz nicht erkennbar sei. Die Auswahl der Einzelprojekte würde einen eher zufälligen Eindruck machen. Mit einem Ausführungsplan, der die einzelnen Projekte auf der Basis der Wirkungszusammenhänge im Bereich der Krankenversicherung definierte, wurde dem Anliegen nach einem ganzheitlichen Ansatz Rechnung getragen (BSV 1998). Er macht deutlich, dass Einzelthemen sorgfältig ausgewählt und aufeinander abgestimmt wurden. Im Allgemeinen ist es allerdings schwierig, eine solche Abstimmung zu erzielen. In der Regionalpolitik wird versucht, mit Leitfragen zu operieren. Darin liegt aber bereits erheblicher Zündstoff, weil es unterschiedliche Sichten darüber gibt, wie das Ganze zusammenhängt und wo Prioritäten zu setzen sind.

## 2.3 Wirkungsaussagen sind oft erst nach Jahren möglich

Aus dem prozessorientierten Ansatz ergibt sich, dass transparent gemacht wird, dass die – aus politischer Sicht – wirklich interessanten Wirkungsfragen meist erst einige Jahre nach der Einführung einer Massnahme gestellt werden können. Damit die Ergebnisse in den politischen Entscheidungsprozess einfließen können, ist es dafür oft sogar zu spät. Weil die Umsetzung einer Politik einfach Zeit braucht, müssen Evaluationen sich in den ersten Jahren auf Vollzugsfragen konzentrieren. Die Informationen, die dadurch generiert werden, sind für die Politik oft nicht spannend genug, dass dafür Geld ausgegeben werden könnte.

Im guten Fall, zum Beispiel in der Energiepolitik, wird die Evaluation weitergeführt, aber von der Verwaltungsaussenwelt wenig beachtet. Im schlechten Falle wird die Evaluationsübung abgebrochen. Eine in meinen Augen interessante, aber nicht ungefährliche Entwicklung geht dahin, dass der Ruf nach Arbeiten wächst, die rasch Antworten auf Wirkungsfragen geben. Weil die Politik Wirkungsaussagen verlangt, werden solche generiert, damit die Budgets für die Programme weiterlaufen. Die wissenschaftlich gestützte Evaluation kann derartige Aussagen nicht liefern. In der Energiepolitik sind die Angaben über energetische Wirkungen und Arbeitsplatzeffekte von Energie 2000 Schätzungen von Beteiligten. In der Gesundheitspolitik zeichnet sich etwas Ähnliches ab, weil die Informationsbedürfnisse der Politik zum Teil den Möglichkeiten seriöser Empirie nicht entsprechen.



## 2.4 Fehlende Distanz zwischen Evaluierenden und Verantwortlichen

Der partizipative Ansatz birgt die Gefahr, dass die für die Beurteilung staatlichen Handelns notwendige Distanz zwischen den Verantwortlichen und den Evaluierenden nicht mehr selbstverständlich ist. Dieses Problem stellt sich auch dann, wenn die Durchführung der Untersuchungen durch unabhängige Expertenteams erfolgt. Hilfestellung bieten in diesem Zusammenhang sicher die professionellen Standards, welche Leitplanken für die Evaluationspraxis setzen (Widmer 1996). Eine Grundfrage, die sich in der Energiepolitik immer wieder stellt, ist jene, ob Evaluatorinnen und Evaluatoren die Begleitung der Weiterentwicklung der Politik und damit die Umsetzung ihrer Vorschläge selber an die Hand nehmen sollen oder nicht. Es geht grundsätzlich auch hier um die Frage der Distanz der Evaluation zur Politik. In meinen Augen handelt es sich bei der Evaluation und bei der Strategieberatung um zwei sehr unterschiedliche Aufgaben, die unterschiedliche Kompetenzen verlangen. In diesem Sinne ist es kaum möglich, dass Evaluatorinnen und Evaluatoren systematisch für die Strategieberatung beigezogen werden. In Einzelfällen mag dies zweckmässig sein. Es ist aber nicht zu erwarten, dass eine qualifizierte Evaluation der bisherigen Arbeiten auch sagen kann, was genau in Zukunft zu tun ist. Sie kann Probleme aufzeigen, das Entwickeln von Lösungen ist eine andere Aufgabe.

## 2.5 Grundsätzliche Fragen

Schliesslich zeigen die Erfahrungen deutlich, dass mit den heutigen Systemen zur Generierung von Evaluationsfragestellungen grundsätzliche Themen kaum oder nur selten aufgeworfen werden. Evaluationen konzentrieren sich in der Praxis meist auf das unabdingbare «fine tuning», dies ist auch eine Folge des partizipativen Ansatzes. Das «fine tuning» zu thematisieren ist wichtig, aber daneben sollten auch Grundsatzfragen gestellt werden. Zum Teil wird behauptet, dass diese Themen von der Verwaltungskontrollstelle oder der Finanzkontrolle aufgegriffen werden. Es gibt durchaus Beispiele, wo dies der Fall ist. In meinen Augen gibt es aber noch zahlreiche schwarze Löcher, die weder von der einen noch von der anderen Stelle bearbeitet werden. Auch hier sind in den Evaluationskonzepten Wege vorzusehen, damit grundsätzliche Fragen thematisiert werden können.

Ideale Modelle zur Neutralisierung der genannten Probleme gibt es nicht. Es macht jedoch gerade den Reiz der Evaluation aus, dass der ideale Weg immer neu gesucht und ausgehandelt werden muss.

### 3 Welche Art von Synthesen sind möglich?

Von ganzheitlich angelegten Evaluationssystemen, welche die genannten Probleme mehr oder weniger gut umschiffen, wird erwartet, dass aus ihnen nützliche Ergebnisse für die Weiterentwicklung der Politik gewonnen werden können. Die Evaluationsforschung kennt verschiedene Instrumente zur Synthetisierung von Evaluationsergebnissen.

Zum einen gehört dazu die «kumulative Sekundäranalyse» oder «Substanzorientierte Metaevaluation» (Widmer 1996). Das Ziel dieses Ansatzes besteht darin, in einer forschungsökonomischen sparsameren Vorgehensweise bei gleichzeitiger Erhöhung der empirischen Evidenz Resultate verschiedener Studien auszuwerten. Entscheidend ist aber, dass es sich dabei um Studien zu denselben Fragestellungen handelt. Es wird also die gleiche Frage immer wieder in unterschiedlichen Untersuchungen behandelt. Zum anderen handelt es sich auch bei den Metaevaluationen, welche die Qualität von Evaluationen bewerten, um einen Ansatz, der im Zusammenhang mit Evaluationssynthesen oft erwähnt wird. Dort geht es darum, zu prüfen, ob die wissenschaftliche Qualität, die Praxisrelevanz, die Nutzung und die Wirkung einer oder mehrere Evaluationen einem vorgegebenen Standard entsprechen. Das Ziel von Metaevaluationen besteht darin, ein empirisch erprobtes Kriterienset für die Bewertung von Evaluationen heranzuziehen und an Einzelfällen zu prüfen, wie weit diese die Kriterien erfüllen.

Im Kern sind beide genannten Syntheseinstrumente aber rückwärts gewandt. Evaluationen analysieren Vergangenes, und diese Rahmenbedingung gilt auch für deren Synthesen. Was die Politik jedoch von Synthesen erwartet, sind Hinweise darauf, wie sie in Zukunft ihre Arbeit besser machen könnte. Diese Hinweise sollen auf den gesammelten Erfahrungen von empirisch gestützten Evaluationen aufbauen. Damit will die Politik nur das einlösen, was ihr die Evaluation immer wieder verspricht: die Möglichkeit, aus den Evaluationen zu lernen. Die Evaluation steht in einem instrumentellen Kontext, auch wenn man die Ansicht nicht teilt, dass sie als elitär-technokratisches Werkzeug dient, das sich vom ursprünglichen Lasswell-schen Postulat der Policy-Analyse als einer Demokratiewissenschaft weit entfernt hat (Fischer 1993).

Evaluationssynthesen, die für die Politikgestaltung und für die Weiterentwicklung der gesellschaftlichen Problemlösungsfähigkeit wertvoll sein sollen, müssen mehr sein als blosse Zusammenstellungen von Evaluationsergebnissen. Sie müssen anstehende Fragen vor dem Hintergrund der Ergebnisse von «Ex-post-Evaluationen» spiegeln. Solche Evaluationssynthesen können kein wissenschaftliches Produkt aus wissenschaftlichen Evaluatio-

nen sein. Sie sind vielmehr eine intelligente Mischung aus wissenschaftlichen Ergebnissen von Evaluationen und Kenntnissen bezüglich der anstehenden Probleme eines Politikbereichs. Zugespißt bedeutet dies, dass die Wissenschaftlichkeit für nützliche Evaluationssynthesen zwar eine notwendige, aber keine hinreichende Bedingung ist.

Dies lässt sich an einem Beispiel erläutern: Energieagenturen sind gegenwärtig in der Energiepolitik ein viel diskutiertes Thema. Dies nicht zuletzt darum, weil das anfangs 1999 in Kraft getretene schweizerische Energiegesetz die Schaffung von Energieagenturen vorsieht. Allerdings bestehen unterschiedliche Vorstellungen bezüglich der konkreten Ausgestaltung und der Aufgabenstellung von Energieagenturen. Es stellt sich nun die Frage, ob die Resultate von Evaluationen zu diesem Thema etwas Nützliches beitragen können.

Die eine Antwort lautet, sie kann es nicht, weil es bisher keine den geplanten schweizerischen Energieagenturen entsprechenden Einrichtungen gab, deren Erfahrungen man wissenschaftlich genau übertragen kann. Die andere Antwort lautet, dass dies möglich ist, allerdings nicht eins zu eins, sondern transformiert und angereichert mit vertieften Erkenntnissen des Politikbereichs. Die Resultate aus der Evaluation legen es etwa nahe zu behaupten, dass Energieagenturen dann erfolgreich sein können, wenn sie einen Marktauftritt garantieren, der die Vorteile von staatlichem Image und privatwirtschaftlicher Dynamik verbindet. Weiter weisen die Resultate aus Evaluationen darauf hin, dass Energieagenturen auf ihre Wirksamkeit überprüft werden müssen (Balthasar/Rieder 1997). Diese Erkenntnisse finden sich aber nicht in einzelnen Evaluationen. Es lassen sich nur wenige eindeutige Referenzen angeben.

Ich bin der Auffassung, dass sich die Evaluation den Herausforderungen der Politikberatung stellen muss. Evaluationen müssen sich auch dadurch rechtfertigen, dass aus ihnen gelernt werden kann. Die wissenschaftliche Arbeitsweise bleibt selbstverständlich eine notwendige Bedingung für die Durchführung von Untersuchungen. Diese Grundlage ist aber nicht hinreichend für das Formulieren von Evaluationssynthesen. Diese müssen auf den Erkenntnissen kontinuierlicher und unabhängiger Auseinandersetzung mit einem Politikbereich beruhen.

Evaluationssynthesen sind nur möglich, wenn ein Politikbereich kontinuierlich evaluiert wird. Dazu ist eine gewisse Institutionalisierung notwendig. Sie sichert den Bestand von Spezialistinnen und Spezialisten, die nicht nur etwas von Evaluation verstehen, sondern auch den Politikbereich kennen. Die Institutionalisierung stellt sicher, dass Evaluationen nicht

Stückwerk bleiben, sondern zu einem ganzheitlichen Bild verbunden werden und auf diese Weise einen Beitrag zur Steigerung der Problemlösungsfähigkeit des Staates leisten können.

#### Anmerkungen

- 1 Grundlage der Ausführung bildet ein Vortrag, der anlässlich der Jahresversammlung der Schweizerischen Evaluationsgesellschaft 1998 in Solothurn gehalten wurde.
- 2 Interessante Hinweise dazu liefern:  
Bussmann/Klöti/Knoepfel 1997; Bussmann 1996; Widmer 1996; Rossi/Freemann 1993; Linder/Rieder/Varone 1992; Maurer/Marti/Stapfer 1997.

#### Literatur

- Balthasar, A./Rieder, St., 1997, Energieagenturen als Element schweizerischer Energiepolitik. Eine Bewertung von Varianten, Diskussionpapier zuhanden des Bundesamtes für Energie.
- Balthasar, A., 1998, Strategie für die Evaluation des raumordnungspolitischen Instrumentariums zuhanden des Bundesamt für Wirtschaft und Arbeit.
- Balthasar, A., 1999, Die sozialpolitische Wirksamkeit der Prämienverbilligung in den Kantonen. Evaluationsergebnisse und Monitoringkonzept, Beiträge zur sozialen Sicherheit, BSV-Forschungsbericht Nr. 21/98.
- Balthasar, A., 2000, Evaluationen im Energiebereich: Eine Übersicht über Inhalt und Methode von 120 Evaluationen (in Vorbereitung).
- BASS, Büro für Arbeits- und Sozialpolitische Studien, 1999, Wirkungsanalyse des Risikoausgleichs in der Krankenversicherung, Beiträge zur sozialen Sicherheit, BSV-Forschungsbericht Nr. 1/99.
- BFE, Bundesamt für Energie, 1992, Konzept für die Evaluation von Energie 2000.
- BSV, Bundesamt für Sozialversicherung, 1998, Ausführungsplan zur Wirkungsanalyse KVG.
- Bussmann, W./Knoepfel P., 1997, Die öffentliche Politik als Evaluationsobjekt, in: W. Bussmann/U. Klöti/P. Knoepfel, *Einführung in die Politikevaluation*, S. 58–77.
- Bussmann W., 1996, Evaluationen – Mehr Transparenz über die Wirkungen staatlichen Handelns. Ergebnisse aus dem Nationalen Forschungsprogramm «Wirksamkeit staatlicher Massnahmen» NFP 27.
- Dreher C./Balthasar A., 1997, Evaluierung des Schweizer CIM-Aktionsprogramms 1990–1996. Endbericht an das Bundesamt für Konjunkturfragen.
- Fischer F., 1993, Bürger, Experten und Politik nach dem Nimby-Prinzip: ein Plädoyer für die partizipatorische Policy-Analyse, *Politische Vierteljahresschrift* 34, Sonderheft 24, S. 451–470.
- Hanser, C./Kuster, J./Cavelti, G., 1994, Hotellerieförderung durch Bund und Kantone. Evaluation der Auswirkungen auf die Hotellerie, Schriftenreihe des BIGA, Beiträge zur Tourismuspolitik Nr. 3.
- Linder, W./Rieder, St./Varone, F., 1992, Schlussbericht des Vorprojekts Vollzugevaluation des Aktionsprogrammes «Energie 2000». Literaturauswertung, Bericht über die Evaluierbarkeitsprüfung, Übersicht über Massnahmen, Daten und Statistiken.
- Maurer, R./Marti, F./Stapfer, A., 1997, Kontrollprogramm Natur und Landschaft Kanton Aargau – Konzeption und Organisation von Erfolgskontrolle und Dauerbeobachtung, Grundlagen und Berichte zum Naturschutz Nr. 13, hrsg. vom Baudepartement des Kantons Aargau.
- Rieder, St./Iten, R., 1999, Evaluation des Investitionsprogramms Energie 2000.
- Rossi, P. H./Freemann, H. E., 1993, Evaluation: A Systematic Approach.
- Widmer, Th., 1996, Meta-Evaluation. Kriterien zur Bewertung von Evaluationen.

## **Résumé**

*Le nombre des évaluations a considérablement augmenté en Suisse ces dernières années. L'expérience montre que pour parvenir à des résultats consistants, les projets d'évaluation doivent être intégrés dans une conception globale. La présente contribution décrit les éléments constitutifs d'un système d'évaluation complet et orienté sur les résultats. L'auteur défend la thèse selon laquelle, le contexte politique évoluant rapidement, les résultats d'une évaluation ne peuvent généralement pas être établis par des moyens scientifiques, ce qui ne les empêche pas de justifier les efforts consentis pour les obtenir.*