

## Was und wie wird evaluiert? Eine Übersicht über die Evaluationspraxis in der Energiepolitik

**Andreas Balthasar** | *Evaluationen sind zu einem festen Bestandteil von Politik und Verwaltung geworden. Mit der sich entwickelnden Evaluationspraxis stellen sich zahlreiche Fragen nach den Gegenständen, Methoden und disziplinären Hintergründen von Evaluationen. Für den Bereich der Energiepolitik können erste zuverlässige Aussagen gemacht werden.*

### 1 Einführung

Auch in der Schweiz sind Evaluationen zu einem festen Bestandteil von Politik und Verwaltung geworden. Wurde die Wirksamkeit staatlichen Handelns bis anfangs der 80er-Jahre noch mittels spontaner Evaluationen durch die freie wissenschaftliche Forschung und im Rahmen Nationaler Forschungsprogramme eruiert, so ermöglichen heute rechtliche Bestimmungen eine systematische Erfolgskontrolle.<sup>1</sup> Evaluation ist heute – zumindest auf Bundesebene – in vielen Politikbereichen Alltag geworden, und die Zahl der Untersuchungen nimmt stetig zu. Vor diesem Hintergrund stellen sich zahlreiche Fragen im Zusammenhang mit der sich entwickelnden Evaluationspraxis: Offen ist beispielsweise, welche Evaluationsgegenstände dominieren, welche Evaluationsdesigns im Zentrum stehen, welchen disziplinären Hintergrund Evaluationen haben und welche Methoden vorherrschen.

Zwar haben Zimmermann/Kolly 1988 eine bibliographische und methodische Exploration der relevanten Evaluationen von Bundespolitik erarbeitet, doch die damit eingeleiteten Bestrebungen zur Gewinnung einer Übersicht über das Evaluationsschaffen in der Schweiz wurden nicht systematisch weitergeführt. Empirisch gestützt können die aufgeworfenen Fragen somit für die ganze Schweiz nicht beantwortet werden. Für den Bereich der Energiepolitik ist es jedoch möglich, zuverlässige Aussagen zu machen. Grundlage dafür bildet eine Datenbank der energiepolitischen Evaluationen, welche das Institut für Politikstudien Interface im Auftrag des Bundesamtes für Energie erstellt hat.

In der Energiepolitik ist die Erfolgskontrolle durch Evaluationsklauseln seit 1990 rechtlich verankert. Der Energienutzungsbeschluss vom 14. Dezember 1990 (AS 1991 1018), Rückgrat der gesetzlichen Massnahmen des Bundes im Rahmen von Energie 2000, forderte in Artikel 24 ausdrücklich eine Wirk-

samkeitskontrolle (SR 730.0). Denselben Grundsatz enthält auch Artikel 20 des Energiegesetzes vom 26. Juni 1998. Diese gesetzlichen Verpflichtungen wurden im Rahmen der Vereinbarungen über das Aktionsprogramm Energie 2000 gar noch ausgeweitet (Rieder/Balthasar 1997). Die Evaluation sollte sich nicht auf die Bewertung des Vollzugs und der Zielerreichung im Bereich der gesetzlich vorgeschriebenen Massnahmen beschränken, sondern auch die Umsetzung und die Wirksamkeit der freiwilligen Aktionen, welche durch die Ressorts von Energie 2000 eingeleitet worden sind, in die Beurteilung einbeziehen. Das Bundesamt für Energie (BFE) stellte dafür finanzielle Mittel bereit und gab Evaluationen in Auftrag.

Die erwähnte Datenbank erfasst die Evaluationen der Jahre 1990 bis 2000 des Bundes, aber auch von Kantonen, Gemeinden und anderen Akteuren, welche sich mit energiepolitischen Fragestellungen beschäftigen. Wie die meisten anderen Evaluationsklauseln, hat auch Artikel 24 des Energienutzungsbeschlusses «*retrospektiven Charakter*» (Bussmann 1997, 27). Die Datenbank erfasst daher Evaluationen, welche zumindest in Teilen auch ex post angelegt sind. Studien der Jahre 1980 bis 1989, welche die Thematik des Aktionsprogrammes behandelten, wurden ebenfalls berücksichtigt und in die Datenbank aufgenommen, so dass bis Ende 2000 165 energiepolitische Evaluationen registriert wurden. Davon waren zum Zeitpunkt dieser Untersuchung (Ende 1999) 147 Studien abgeschlossen, 13 liefen noch und fünf waren abgebrochen worden.

Die Datenbank unterscheidet drei Etappen einer Evaluation: Ein erster Teil umfasst Angaben zur Auftragsvergabe (Auftraggeber, Offertenverfahren, Auftragnehmer, Budget und Ziele der Studie). Der zweite Teil dokumentiert die methodische Anlage der Evaluation und Angaben zu Design und Methodik. Schliesslich wird das Endergebnis mittels einer Zusammenfassung sowie Angaben zur Bezugsquelle präsentiert.

Diese Daten wurden nach Ende des Aktionsprogramms Energie 2000 systematisch analysiert. Die 147 abgeschlossenen Evaluationen wurden in die Auswertung einbezogen. 25 davon stammen aus der Periode 1980 bis 1989, fallen also vor den Start des Aktionsprogramms Energie 2000. Der vorliegende Beitrag fasst die Ergebnisse dieser Analysen zusammen. Im Zentrum stehen die Fragen nach den zentralen Evaluationsgegenständen, nach den Evaluationsdesigns und nach den Evaluationsmethoden. Der erste Teil enthält Angaben zu Auftraggebern sowie Kosten der durchgeführten Evaluationen und fokussiert auf die Evaluationsgegenstände. Der zweite Teil widmet sich vorwiegend den Designs der durchgeführten Untersu-

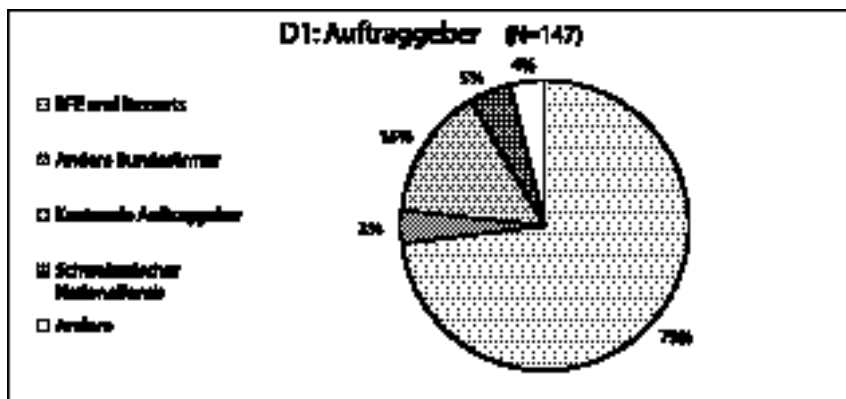
chungen. Schliesslich untersuchen wir im dritten Teil, mit welchen Techniken die nötigen Daten und Informationen erhoben und ausgewertet wurden. Dabei werden jeweils gängige Hypothesen aufgegriffen und auf ihre Gültigkeit im Energiebereich geprüft. In einer abschliessenden Betrachtung werden die Ergebnisse zusammengefasst und Schlüsse aus den Erkenntnissen gezogen.

## **2 Entstehungsbedingungen**

Um einen ersten Eindruck über die Art der Evaluationen im Energiebereich zu gewinnen, werden zunächst Hinweise über Auftraggeber und Kostenrahmen der Evaluationsstudien gemacht. Anschliessend wird untersucht, welche Fragestellungen am häufigsten im Zentrum standen.

### **2.1 Auftraggeber und Kosten der Evaluationen**

Mit dem Beginn des Aktionsprogramms verändert sich nicht nur (wie beabsichtigt) die Zahl der in Auftrag gegebenen Evaluationen, sondern auch das Gewicht der einzelnen Auftraggeber. Haben Aufträge vor 1990 etwa ein Drittel aller Studien ausgemacht, so sind das BFE und die Ressorts von Energie 2000 in den 90er-Jahren eindeutig zu den wichtigsten Initianten energiepolitischer Evaluationen geworden. Wie in Darstellung D1 ersichtlich wird, sind 73 Prozent der erfassten Evaluationen vom BFE oder den Ressorts von Energie 2000 in Auftrag gegeben worden. Obwohl der grosse Anteil des Bundes im Kontext des Aktionsprogramms Energie 2000 nicht erstaunen mag, gilt es zu bedenken, dass Zimmermann (1990) in seinem Bericht über Evaluationen von Bundespolitik darauf hingewiesen hat, dass Evaluationen meist in einem Anwendungskontext stehen und dass «nur wenige Evaluationen nicht von einer politischen Legitimation profitieren»<sup>2</sup> (Zimmermann 1990, 80). Im Energiebereich sind nur rund vier Prozent der Evaluationen nicht von Amtsstellen initiiert worden. Diese Evaluationen sind meist auf die Initiative universitärer Institute zurückzuführen und können zusammen mit den vom Schweizerischen Nationalfonds vergebenen Aufträgen in die Kategorie der «Forschung» eingeteilt werden, die damit rund zehn Prozent aller Evaluationen dieses Politikbereichs ausmachen.



Als wichtigster Auftraggeber der Evaluationsstudien stellt das BFE auch den grössten Teil der finanziellen Mittel für deren Erstellung bereit. Von den 122 während des Aktionsprogramms initiierten Evaluationen konnten deren 82 von Geldern des Bundesamtes profitieren. Dabei betrug der durchschnittliche Beitrag 63'000 Fr. Obwohl die finanzielle Unterstützung je nach Umfang des Projekts deutlich schwankte – die Beträge bewegen sich zwischen 2'000 Fr. und 150'000 Fr. –, war die Unterstützung in jedem Fall beträchtlich. In 78 Fällen wurden die Kosten ausschliesslich vom Bundesamt getragen. Wurde also eine Evaluationsstudie vom BFE finanziell unterstützt (und dies ist nur 26 Mal nicht der Fall), so trug das Bundesamt in der Regel die gesamten Kosten.<sup>3</sup>

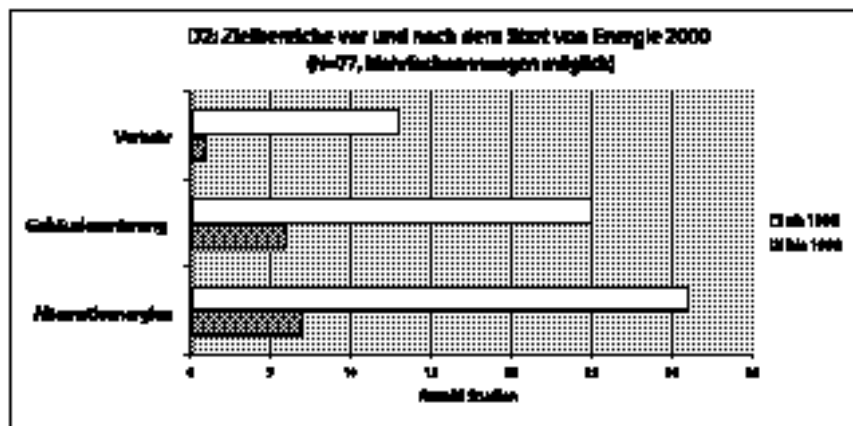
## 2.2 Thematische Schwerpunkte der Evaluationen

Die Förderung einer sparsamen und rationellen Energieverwendung sowie der erneuerbaren Energien bilden die Grundsätze des Energienutzungsbeschlusses vom 14. Dezember 1990 (Art. 2 Abs. 1). Für diese qualitativen Oberziele des Aktionsprogramms Energie 2000 wurden quantitative Zielsetzungen für die vier Energieträger Brennstoffe, Treibstoffe, Strom und erneuerbare Energien formuliert. Als Aktivitäten lassen sich alle im Rahmen von Energie 2000 zusammengefassten gesetzlichen Massnahmen und freiwilligen Aktionen zur Erreichung dieser Ziele bezeichnen. Dabei wird zwischen regulativen, finanziellen und persuasiven Aktivitäten unterschieden. Regulative Instrumente gründen auf Geboten und Verboten, finanzielle Instrumente beruhen auf monetären Anreizen, persuasive auf Überzeugung durch Information, Beratung und Ausbildung (Rieder/Balthasar 1997, 15ff.).

Im folgenden Abschnitt interessiert uns nun die Frage, zu welchen Aktivitäten des Aktionsprogramms am häufigsten Evaluationen erstellt werden. Dazu haben wir eine Liste mit relevanten Stichworten für das Aktionsprogramm Energie 2000 erarbeitet und in der Folge den Studien zugeordnet, wobei Mehrfachnennungen zugelassen wurden.<sup>4</sup> In einem ersten Schritt betrachten wir nun die Evaluationen zu den verschiedenen Zielsetzungen in den wichtigen Bereichen Alternativenergien, Gebäudesanierung und Verkehr. Danach wenden wir uns den vorwiegend evaluierten Instrumenten zu.

### 2.2.1 Welche Programmziele wurden am häufigsten evaluiert?

Die nachfolgende Darstellung D2 zeigt, wie sich die Evaluationen auf die verschiedenen Ziele von Energie 2000 verteilen.



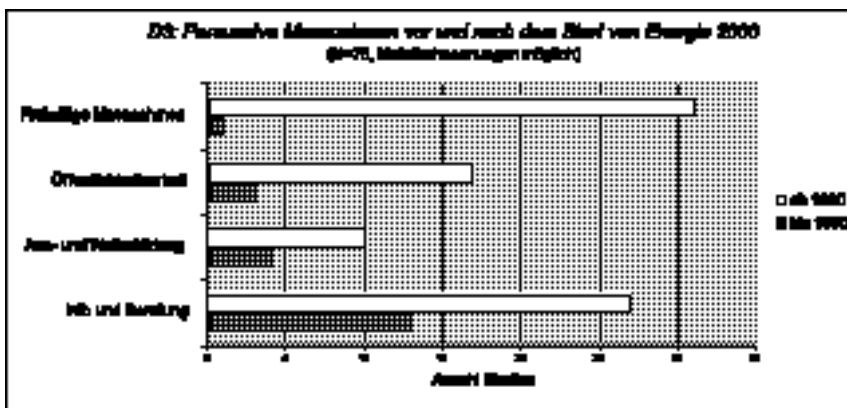
Betrachten wir die Darstellung D2 genauer, so lassen sich insbesondere zwei Feststellungen machen:

1. Es lässt sich feststellen, dass der Bereich Verkehr bis 1990 kaum evaluiert wurde. Zwar ist der Anteil an Verkehrsstudien auch nach dem Start des Aktionsprogramms noch relativ gering, doch ist die Zunahme von Evaluationen mit verkehrspolitischen Fragestellungen im Rahmen von Energie 2000 deutlich markanter als in den anderen Themenbereichen. Dieser Trend dürfte wohl auf die Einbindung der Verkehrsthematik in das Ressort Treibstoffe, und somit in die Energiepolitik, zurückzuführen sein.<sup>5</sup>

2. *Es manifestiert sich das grosse und zunehmende Gewicht von Evaluationen im Bereich der Alternativenergien.* Rund 50 Prozent der 77 Evaluationen kreisen um das Thema Alternativenergien. Dieser grosse Anteil erstaunt indessen wenig, ruft man sich in Erinnerung, dass mit dem «Energiefrieden», der dem Aktionsprogramm zugrunde lag, nebst der rationellen Energieverwendung der Einsatz erneuerbarer Energien als Konsensbereich identifiziert wurde. Obwohl uns eine Übersicht über die Zahl der während der gesamten Dauer des Aktionsprogramms eingeleiteten Aktionen im Bereich der erneuerbaren Energien fehlt, können wir doch sagen, dass in den Jahren 1992 bis 1995 im Vergleich zu jedem anderen Ressort mehr als doppelt so viele Aktionen vom Ressort regenerierbare Energien ausgehen (Rieder/Balthasar 1997, 62). Deshalb kann das Gewicht der Evaluationen in diesem Bereich kaum überraschen.

### 2.2.2 Welche Art von Aktivitäten wurden hauptsächlich evaluiert?

In einem zweiten Schritt wollen wir nun untersuchen, welche Art von Massnahmen und Aktivitäten am häufigsten evaluiert wurde. Diese Analyse beruht auf allen 147 in der Datenbank erfassten Studien, wobei Mehrfachnennungen zugelassen wurden. Die Betrachtung weist deutlich auf einen Schwerpunkt bei der Evaluation persuasiver Massnahmen hin. Die grosse Bedeutung der persuasiven Massnahmen mag vielleicht erstaunen, doch darf nicht vergessen werden, dass das Aktionsprogramm Energie 2000 derartigen Aktivitäten sehr grosses Gewicht beimass. Um uns im Folgenden einen besseren Überblick über die Art der persuasiven Massnahmen zu verschaffen, betrachten wir nachstehende Darstellung D3.



Die Darstellung macht deutlich, dass sich die Evaluation sowohl vor als auch nach dem Start des Aktionsprogramms Energie 2000 vor allem mit dem Instrument der Information und der Beratung beschäftigte. Etwa jede zweite Evaluation von persuasiven Massnahmen war dem Instrument der Information und der Beratung gewidmet. Diese Schwerpunktsetzung lässt sich insbesondere dadurch erklären, dass zwischen 1992 bis 1995 etwa die Hälfte der von den Ressorts und Aktionsgruppen lancierten Aktionen in den Bereich Information (33 Prozent) und Beratung (18 Prozent) fielen (Rieder/Balthasar 1997, 64).

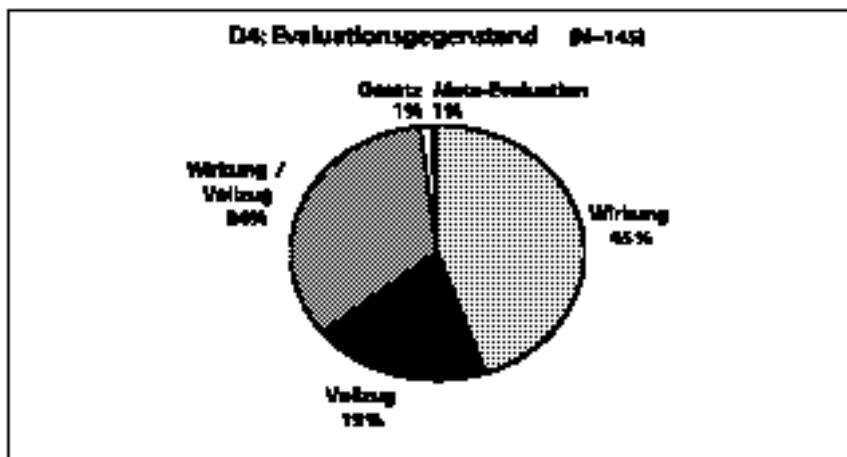
Ähnlich häufig wie Informations- und Beratungsaktivitäten werden freiwillige Massnahmen wie Betriebsoptimierungsprogramme oder der Kauf von gelabelten Geräten evaluiert. Rund 40 Prozent aller Evaluationen persuasiver Massnahmen gehören zu diesem Bereich. Anders als die Evaluationen von Information und Beratung gewinnt die Untersuchung freiwilliger Massnahmen jedoch erst mit dem Start des Aktionsprogrammes Energie 2000 an Bedeutung. Dadurch wird die strategische Ausrichtung des Programmes Energie 2000, das Freiwilligkeit in der Konzeption verankerte, deutlich.

### 2.3 Gegenstand der energiepolitischen Evaluationen

Evaluationen können verschiedene Zielsetzungen und Gegenstände haben. Politikwissenschaftlerinnen und Politikwissenschaftler orientieren sich meist an der Typologie des Policy-Cycle, um die Evaluationsgegenstände im Kontext eines Gesamtprozesses abzubilden und aufzugliedern. Dabei werden verschiedene Stufen der Politikgenerierung und –umsetzung unterschieden, welche allesamt Gegenstand einer Evaluation sein können (vgl. dazu Knoepfel/Bussmann 1997, 58–77, und Balthasar 2000). Werden die Ziele einer Politik nicht erreicht, kann dies bereits auf Defizite in der Konzeption der Politik zurückgeführt werden. Beispielsweise können Wirkungsdefizite auf Grund ungeeigneter Kausalhypothesen entstehen. In diesem Zusammenhang wird auch von «vorprogrammierten Vollzugsdefiziten» gesprochen (Knoepfel/Bussmann 1997, 92). Aus diesen Gründen werden Evaluationen der Gesetze durchgeführt. Denkbar ist aber auch, dass unerwünschte Wirkungen auf ungenügende Umsetzungsmechanismen zurückzuführen sind. In diesem Fall wird ein spezielles Augenmerk auf den Vollzug gerichtet sein. Weiter können die Wirkungen einer Politik evaluiert werden, wobei in der Regel verschiedene Wirkungsebenen unterschieden werden. Schliesslich haben wir noch die Kategorie der Meta-Evaluation berücksichtigt, das

heisst, Evaluationen, welche die Ergebnisse einer oder mehrerer Studien kritisch würdigen.

Die einzelnen Kategorien sind in der Praxis manchmal schwierig auseinander zu halten, und oft untersucht eine Studie gleichzeitig mehrere Gegenstände. In unserer Auswertung haben wir deshalb auch Kombinationen separat aufgenommen (vgl. D4).



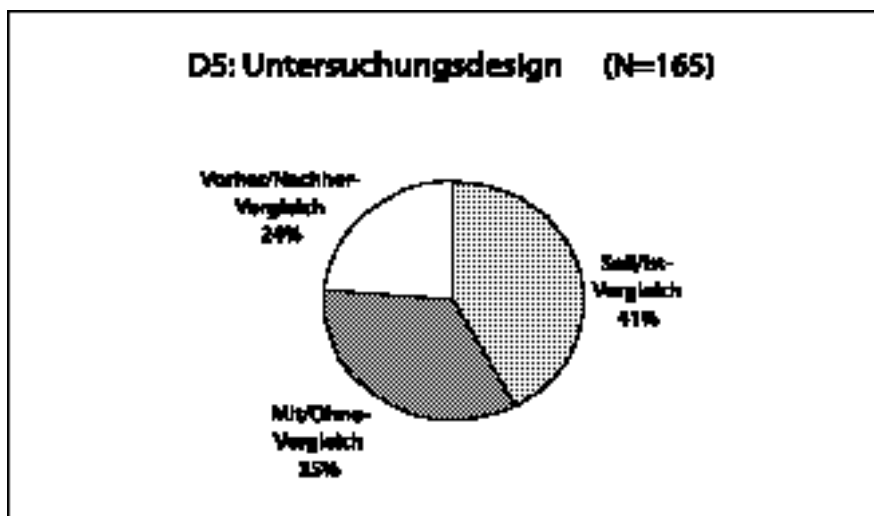
Auf Grund der Auswertung der uns verfügbaren Evaluationsstudien können wir feststellen, dass im Energiebereich Wirkungsanalysen vorherrschen. Die Kategorie «Vollzug» vereinigt deutlich weniger Studien auf sich; allerdings betrachtet ein Drittel aller Evaluationen Vollzug und Wirkungen gemeinsam. Dieses Bild steht in Kontrast zu den von Rieder/Varone (1991) gefundenen Resultaten, welche auf einer international angelegten Literaturliteraturanalyse beruhen. Die beiden Autoren kommen nämlich zum Schluss: «Bei den Evaluationsarten im Bereich Energiesparprogramme standen in der Vergangenheit eindeutig die Vollzugs- und Prozessevaluationen im Zentrum der Forschungstätigkeit» (Rieder/Varone 1991, 56). Diese Angaben beziehen sich auf die Zeit vor dem Aktionsprogramm Energie 2000. Die Datenbank, die wir ausgewertet haben, zeigt in dieser Periode auch für die Schweiz eine grosse Anzahl reiner Vollzugsanalysen. Rund ein Drittel der Evaluationen in den Jahren vor dem Aktionsprogramm Energie 2000 hatte den Vollzug zum Gegenstand. Allerdings dominierten auch in dieser Zeitperiode in der Schweiz die Wirkungsanalysen.



Die Auswertung der Datenbank zeigt aber auch, dass in den letzten Jahren die Bestrebungen intensiviert wurden, die Wirkungen von Massnahmen und Aktionen mit Evaluationen zu erfassen. Dabei wurde Wirkung nicht nur im Sinne von Outcomes, sondern auch von Impacts verstanden.

### 3 Evaluationsdesign

Das Untersuchungsdesign ist das konzeptionelle Herz einer Evaluation, geht es doch darum, festzulegen, wie die Fragestellung einer Evaluation in eine empirische Untersuchungsanlage umgesetzt wird. Die Wahl des Untersuchungsdesigns beinhaltet insbesondere die Entscheidung, welche Vergleichsebene der Untersuchung zugrunde gelegt werden soll und ob die Evaluation einen einzelnen Fall untersuchen oder einen Vergleich zwischen mehreren Fällen vornehmen will.



In Darstellung D6 ist zu erkennen, dass Soll/Ist-Vergleiche das am häufigsten verwendete Evaluationsdesign sind. Ähnlich weit verbreitet sind Vergleiche zwischen Situationen mit und ohne Massnahmen. Vorher/Nachher-Vergleiche sind demgegenüber etwas seltener.

Indem die Charakteristika der einzelnen Designs näher betrachtet werden, können die ersten Resultate noch ausdifferenziert werden. Knapp zwei Drittel der auf einem Soll/Ist-Vergleich basierenden Evaluationen sind Einzelfallstudien. Ein weiteres Fünftel sind vergleichende Fallstudien. Auch die Vorher/Nachher-Vergleiche sind je zu rund der Hälfte Fallstudien oder vergleichende Fallstudien. Anders ist die Verteilung bei den Mit/Ohne-Vergleichen. Diese sind naturgemäss oft als vergleichende Fallanalysen angelegt.

Dieser Befund bestätigt die Aussagen von Klöti/Widmer, welche betonen, dass «vergleichend angelegte Fallstudien in der Schweiz das am häufigsten verwendete Evaluationsdesign sind» (1997, 202). Fast 50 Prozent der Evaluationen der Energiepolitik sind als vergleichende Fallstudien angelegt. 40 Prozent konzentrieren sich auf einen einzigen Fall, und 10 Prozent sind Zeitreihenanalysen.

Die vergleichende Fallstudie bezieht verschiedene Fälle gleichwertig in die Untersuchung ein, um aufgrund festgestellter Differenzen auf Programmwirkungen zu schliessen. Dabei kommen typischerweise Querschnittanalysen zur Anwendung (Klöti/Widmer 1997, 202), und meist werden Situationen mit und ohne Massnahmen verglichen, wobei in vielen Fällen eher von Vergleichen verschiedener Massnahmen gesprochen werden kann. Auffällig selten – nämlich in weniger als fünf Prozent der Fälle – wurde ein internationaler Vergleich vorgenommen.

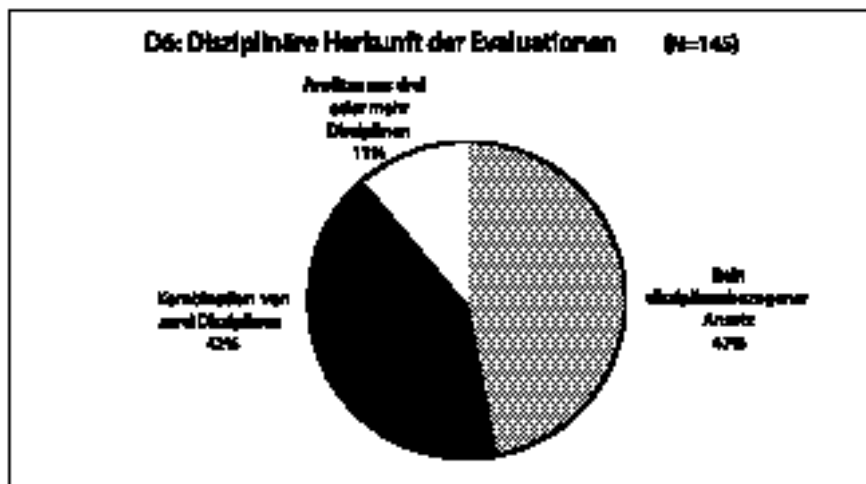
#### **4 Evaluationsmethoden**

Wenden wir uns nun den Erhebungs- und Analysetechniken zu, welche in den Evaluationen verwendet wurden. Da diese Methoden häufig durch den disziplinären Hintergrund der Untersuchungen beeinflusst werden, gehen wir zuerst darauf ein.

##### **4.1 Disziplinäre Herkunft der Evaluationen**

In ihrem Beitrag über Nutzungszusammenhänge von Evaluationen halten Bussmann und Knoepfel fest, dass «auch rund 20 Jahre Politikanalyse die immer noch stark disziplinenbezogene Ausrichtung von Politikstudien nicht ungeschehen machen können» (Bussmann/Knoepfel 1997, 129). Auf der anderen Seite wird immer häufiger betont, dass gerade im Umwelt- und Energiebereich das neue Forschungsprinzip der Interdisziplinarität oder Transdisziplinarität unabdingbar sei, um die Probleme der modernen Welt in all ihrer Vielfalt zu beleuchten (Mittelstrass 1995). Herrscht tatsächlich, wie es Bussmann/Knoepfel postulieren, die fachliche Parzellierung vor, oder zeichnet sich ein Trend zur Interdisziplinarität ab? Wenn ja, wann hat die Trendwende eingesetzt und welche Disziplinen scheinen sich für eine Zusammenarbeit zu eignen?

Um diese Fragen zu beantworten, haben wir die Evaluationsstudien nach ihrer disziplinären Zugehörigkeit sieben verschiedenen Kategorien zugeordnet.<sup>6</sup> Dabei sind beliebige Kombinationen möglich, um Studien mit eher interdisziplinärem Charakter Rechnung zu tragen.



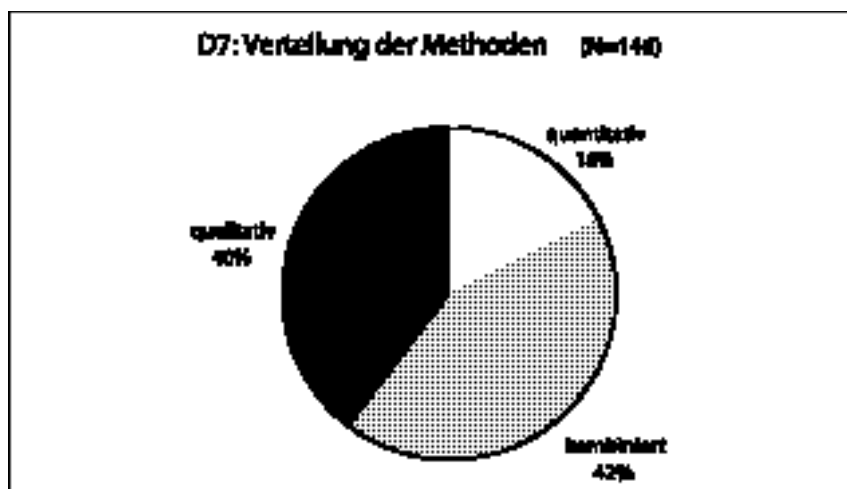
Wie der Darstellung D6 zu entnehmen ist, scheint die Aussage von Bussmann/Knoepfel auch im Energiebereich Gültigkeit zu haben, sind doch 47 Prozent aller Evaluationen einem einzigen disziplinären Ansatz verpflichtet. Innerhalb dieser Kategorie herrschen mit 40 Prozent die soziologischen Studien deutlich vor, gefolgt von Evaluationen mit politikwissenschaftlichem Ansatz (rund ein Drittel). Rein ökonomische Studien machen fast den ganzen Rest aus (20 Prozent). Somit steht auch fest, dass die juristischen, ökologischen, psychologischen und technischen Aspekte im Energiebereich jeweils nur Teil umfassenderer, «interdisziplinär»<sup>7</sup> angelegter Evaluationsstudien sind.

Trotz der nach wie vor weiten Verbreitung fachlicher Parzellierung zählen wir im Energiebereich fast ebenso viele Studien (nämlich 42 Prozent), welche die Perspektiven einer zweiten Disziplin ergänzend integrieren. Allerdings darf dieses Bild vor dem Hintergrund des zunehmenden Interesses an interdisziplinärem Arbeiten nicht allzu optimistisch stimmen. Der Anteil an Studien, welche Aspekte von zwei oder mehr Disziplinen in die Untersuchung einbeziehen, war in den 80er-Jahren höher als in der Zeit danach! Ob der Trend zu Interdisziplinarität nun gegenwärtig im Rahmen der Evaluation der energiepolitischen Massnahmen eher abnimmt, soll an dieser Stelle nicht beantwortet werden. Betrachten wir jedoch die disziplinen-übergreifenden Evaluationen ein wenig näher, so stellen wir fest, dass vor allem ökonomische, technische und politologische (oder aber soziologische) Gesichtspunkte kombiniert werden. Vielversprechend erscheint dabei vor allem der Brückenschlag zwischen Sozial- und technischen Wissenschaften, werden doch somit bis anhin als unvereinbar geltende Para-

digmen und Perspektiven miteinander verknüpft. Allerdings kann angenommen werden, dass es sich dabei um eine Eigenart der Evaluation energiepolitischer Massnahmen handelt.

#### 4.2 Forschungsmethoden

Jede Evaluation stützt sich auf Informationen und Daten, welche zunächst mit einem ausgewählten Forschungsverfahren erhoben und später ausgewertet werden. Die Auswahl des Forschungsverfahrens erfolgt meist in Anlehnung an die erkenntnistheoretische Position der Forschenden. Dabei werden in der Sozialwissenschaft zwei fundamentale Denkmodelle unterschieden. Das empirisch-analytische Paradigma, welches den Naturwissenschaften entspringt, wird vor allem mit quantitativen Forschungstechniken in Verbindung gebracht, das geisteswissenschaftliche interpretativ-hermeneutische Paradigma hingegen ist eng mit qualitativen Forschungsmethoden verknüpft (Widmer/Binder 1997, 214 ff.). Neben der Darstellung des Gewichts der einzelnen Ansätze in der energiepolitischen Evaluation ist es auch Ziel des folgenden Abschnittes, die von Zimmermann formulierte Hypothese, wonach «es keine Evaluation ohne qualitative Methoden gibt» (1990, 84),<sup>8</sup> für den Energiebereich zu überprüfen. Dazu wurden die Evaluationen nach Methode in die Kategorien *qualitativ*, *quantitativ* und *kombiniert* eingeteilt.<sup>9</sup>



Wie Darstellung D7 zeigt, nehmen die qualitativen Methoden auch im Energiebereich einen hohen Stellenwert ein, greifen doch 82 Prozent aller Evaluationen auf solche Techniken zurück. In diesem Sinne kann die Aussage Zimmermanns also bestätigt werden. Beachtlich ist ausserdem, wie häufig (42 Prozent der Fälle) qualitative und quantitative Methoden kombiniert zur Anwendung kommen. Dies deutet darauf hin, dass sich die Forschenden durchaus auf Methoden stützen, welche nicht unbedingt ihren Denkschulen entstammen.

Innerhalb der qualitativen Methoden werden am häufigsten, nämlich in rund 70 Prozent der Fälle, Experten-/Expertinnengespräche geführt. Dokumentenanalysen finden sich in etwas mehr als jeder zweiten rein qualitativen Evaluation, nicht standardisierte Befragungen in ungefähr jeder dritten Untersuchung. Studien, welche auf quantitativen Methoden basieren, enthalten fast immer standardisierte Befragungen. Ein Drittel der Studien greift auf Messungen zurück. Sekundärdatenanalysen, Modellrechnungen, Experimente und Simulationen bilden kaum je die einzige Grundlage einer Studie, sondern werden in der Regel mit qualitativen Methoden kombiniert.

Ein interessanter Zusammenhang lässt sich zwischen Forschungsmethode und Evaluationsgegenstand feststellen. Werden bloss elf Prozent der reinen Vollzugsanalysen mit quantitativen Methoden erstellt, so ist der Anteil der Wirkungsanalysen mit 53 Prozent in eben dieser Kategorie am höchsten. Gerade umgekehrt präsentiert sich das Bild, wenn man die mit qualitativen Methoden erstellten Evaluationen betrachtet. 25 Prozent dieser Evaluationen sind reine Vollzugsanalysen, was ziemlich genau der Hälfte aller Vollzugsanalysen entspricht. Dieses Bild deckt sich mit den Aussagen Bussmanns, welcher festhält: Die Vollzugsforschung «war und ist in der Regel qualitativ angelegt (Dokumentenanalysen, Interviews)» (1997, 24).

## 5 Schlussfolgerungen

Die Untersuchung hat, bezüglich der dominierenden Gegenstände in der Evaluation der Energiepolitik, einen Schwerpunkt bei den Wirkungsstudien zu Tage gefördert. Wirkungsfragen werden häufig allein oder gemeinsam mit Vollzugsfragen gestellt. Zusätzlich lässt sich erkennen, dass der Anteil der Vollzugsstudien im Verlauf der Zeit abgenommen hat.

Weiter lässt sich aus der Analyse der Datenbank erkennen, dass Evaluationen im Energiebereich meist als Soll/Ist-Vergleiche oder als Vergleiche mit und ohne Massnahmen angelegt sind. Fast 50 Prozent der Evaluationen der Energiepolitik sind vergleichende Fallstudien. 40 Prozent sind Einzelfallstudien. Diese Erkenntnis aus dem Energiebereich deckt sich mit den

von Klöti/Widmer (1997, 187) dokumentierten Erfahrungen. Andere Designs haben in der Evaluation energiepolitischer Massnahmen marginale Bedeutung.

Wenn wir schliesslich die eingesetzten Evaluationsmethoden betrachten, stellen wir eine klare Dominanz quantitativer Methoden fest. Vier von fünf Untersuchungen sind – auch – qualitativ angelegt. Dabei waren Experten-/Expertinnengespräche das am häufigsten angewandte Datenerhebungsverfahren. Dokumentenanalysen finden sich ebenfalls in mehr als der Hälfte aller qualitativ angelegten Evaluationen. Untersuchungen, die auf quantitativen Methoden basieren, greifen fast immer auf standardisierte Befragungen zurück. Dieses Ergebnis ist für die Ausbildung von Evaluatorinnen und Evaluatoren von Bedeutung. Es macht deutlich, dass fundierte Kenntnisse sowohl qualitativer wie auch quantitativer Verfahren zum Rüstzeug jener gehören müssen, welche Evaluationen durchführen.

Ob die Evaluation nun als Einzelfallstudie oder vergleichende Fallstudie angelegt ist – fast immer dienen qualitative Methoden wie Experten-/Expertinnengespräche und Dokumentenanalysen zur Erhebung und Auswertung der Daten. Dabei fällt allerdings auf, dass rund 40 Prozent der Studien ergänzend auf quantitative Methoden wie standardisierte Befragungen, Messungen oder Modellrechnungen zurückgreifen. Diese Kombination verschiedener Forschungstechniken kommt meist einem Brückenschlag von sozialwissenschaftlichen und technischen Disziplinen gleich. Auch wenn somit Perspektiven aus sehr unterschiedlichen Disziplinen zusammengeführt werden, kann nicht von eigentlicher Interdisziplinarität gesprochen werden. Allerdings beweist die beliebige Kombination von quantitativen und qualitativen Forschungstechniken, dass sich die Fronten zwischen den bis anhin als unvereinbar geltenden Disziplinen aufzuweichen scheinen. Die Forschenden sind im Gegenteil in der Lage, den Ansprüchen der Fragestellung, den Besonderheiten des Evaluationsgegenstandes und den Umständen der Evaluation Rechnung zu tragen und ihre Forschungsmethoden zweckgerichtet auszuwählen. Ob dieses Ergebnis als allgemeiner Trend einer Verbesserung der methodischen Qualität der schweizerischen Evaluationsforschung gedeutet werden kann, wird eine auch auf andere Politikbereiche abgestützte Metaevaluation zeigen müssen.

### Anmerkungen

Den Grundstein für diese Arbeit hat Frau Anke Tresch im Rahmen eines Praktikums bei Interface Politikstudien Ende 1999 gelegt. Ich möchte Ihr dafür bestens danken.

- 1 Eine aktuelle Übersicht betreffend die Evaluationsklauseln auf Bundesebene findet sich auf der Homepage des Bundesamtes für Justiz: (<http://www.bj.admin.ch/d/index.html>).
- 2 Eigene Übersetzung des französischen Originals.
- 3 Für 14 während des Aktionsprogramms realisierte Studien liegen uns keine Informationen über die Kosten und deren Träger vor.
- 4 Es sind dies folgende Stichworte: Alternativenergien, Aus- und Weiterbildung, Ausstellungen, Bekanntheit, Elektroheizungen, Energiesparmassnahmen, Evaluationsgrundlagen, Forschung und Entwicklung, freiwillige Massnahmen, Gebäudesanierung, Info und Beratung, Pilot- und Demonstrationsanlagen, Tarife, Veranstaltungen, Vereinbarungen, Verkehr, verbrauchsabhängige Heizkostenabrechnung.
- 5 Vgl. dazu auch Balthasar/Bächtiger 2000. Dieser Bericht zeigt unter anderem auf, dass die Verankerung der Evaluation in der schweizerischen Verkehrspolitik generell bescheiden ist.
- 6 Es sind dies die folgenden Kategorien: soziologisch, politologisch, ökonomisch, technisch, juristisch, psychologisch, ökologisch.
- 7 Wir sind uns bewusst, dass wir den Begriff der Interdisziplinarität allgemeiner und weiter fassen, als er generell verwendet wird. In unserem Zusammenhang bezeichnet er Studien, welche auf mehreren disziplinären Ansätzen basieren.
- 8 Unsere Übersetzung des Originaltextes.
- 9 Als qualitative Methoden gelten: Experten-/Expertinnengespräch, Dokumentenanalyse, nicht standardisierte Befragungen, teilnehmende Beobachtung und Gruppengespräche. Quantitative Methoden sind die folgenden: Simulationen, Experimente, Messungen, ökonomische Methoden, Modellrechnungen, Sekundärdatenanalysen und standardisierte, schriftliche und mündliche Befragungen. Die Kategorie «kombiniert» umfasst Studien, welche qualitative und quantitative Techniken verwendet haben (vgl. dazu Widmer/Binder 1997).

### Literatur

- Balthasar, Andreas, 2000a, Energie 2000: Programmwirkungen und Folgerungen aus der Evaluation, Chur/Zürich, Rüeegg.
- Balthasar, Andreas, 2000b, Evaluationssynthesen: Bedingungen, Möglichkeiten und Gren-

zen, Leges Gesetzgebung & Evaluation 2000/1, S. 13 – 25.

- Balthasar, Andreas/Bächtiger, Christine, 2000, Evaluationskonzept für die schweizerische Verkehrspolitik. Bestandesaufnahme bisheriger Ex-post-Evaluationen und Eckpfeiler für die Weiterentwicklung, Bericht im Rahmen des NFP 41 «Verkehr und Umwelt», Bern, EDMZ (Art. Nr. 801.690 d).
- Balthasar, Andreas/Rieder, Stefan, 2000, Learning from Evaluations. Effects of the Evaluation of the Swiss Energy 2000 Programme, Evaluation 6(3), p. 245–260.
- Bussmann, Werner, 1997, Evaluationen in der Schweiz, in: Bussmann, Werner/Klöti, Ulrich/Knoepfel, Peter, (Hgg.), Einführung in die Politikevaluation, Basel, S. 13–35, Helbing & Lichtenhahn.
- Bussmann, Werner/Knoepfel, Peter, 1997, Typische Nutzungszusammenhänge von Evaluationen, in: Bussmann, Werner/Klöti, Ulrich/Knoepfel, Peter, (Hgg.), Einführung in die Politikevaluation, Basel, S. 119–132, Helbing & Lichtenhahn.
- Klöti, Ulrich/Widmer, Thomas, 1997, Untersuchungsdesign, in: Bussmann, Werner/Klöti, Ulrich/Knoepfel, Peter, (Hgg.), Einführung in die Politikevaluation, Basel, S. 185–213, Helbing & Lichtenhahn.
- Knoepfel, Peter/Bussmann, Werner, 1997, Die öffentliche Politik als Evaluationsobjekt, in: Bussmann, Werner/Klöti, Ulrich/Knoepfel, Peter, (Hgg.), Einführung in die Politikevaluation, Basel, S. 58–77, Helbing & Lichtenhahn.
- Rieder, Stefan/Balthasar, Andreas, 1997, Die Entwicklung des Aktionsprogramms Energie 2000 zwischen 1990 und 1996. Dokumentation und Analyse, Bern.
- Rieder, Stefan/Varone, Frédéric, 1991, Fokus und Methoden sozialwissenschaftlicher Evaluationsforschung im Energiebereich. Eine Literaturauswertung, Forschungszentrum für schweizerische Politik, Bern.
- Widmer, Thomas/Binder, Hans-Martin, 1997, Forschungsmethoden, in: Bussmann, Werner/Klöti, Ulrich/Knoepfel, Peter, (Hgg.), Einführung in die Politikevaluation, Basel, S. 214–255, Helbing & Lichtenhahn.
- Zimmermann, Willy, 1990, Un inventaire des évaluations au niveau fédéral, in: Horber-Papazian, Katja, Evaluation des politiques publiques en Suisse. Pourquoi? Pour qui? Comment?, Lausanne.
- Zimmermann, Willy/Kolly, Doris, unter Mitarbeit von Mahon, Pascal, 1988, Die wissenschaftliche Evaluation von Bundespolitik. Eine bibliographische und methodische Exploration, Forschungsstelle Wissenschaft und Politik, Bern.

## Résumé

*Dans la plupart des domaines de la politique, les évaluations sont aujourd'hui monnaie courante et leur nombre continue d'augmenter. Une telle évolution amène les spécialistes à s'interroger sur la pratique de l'évaluation: quels sont les objets qui en font le plus souvent l'objet? Quel est le style d'évaluation qui prédomine? Sur quels fondements théoriques et interdisciplinaires l'évaluation repose-t-elle et quelles sont les méthodes les plus courantes? L'article d'Andreas Balthasar répond à ces questions sur la base de l'analyse de 147 évaluations effectuées entre 1990 et 2000 dans le domaine de la politique énergétique. Il en ressort, entre autre, que les techniques d'évaluation quantitatives et qualitatives se prêtent à de multiples combinaisons, démontrant indirectement la capacité des évaluateurs de tenir dûment compte des exigences de la recherche spécifique, des particularités de l'objet à évaluer et des circonstances de chaque évaluation et donc à choisir pour chaque cas particulier la méthode appropriée.*