

Andreas Balthasar (2009): Evaluationen in der Schweiz: Verbreitung und Verwendung, in: Thomas Widmer; Wolfgang Beywl; Carlo Fabian (Hrsg.) Evaluation. Ein systematisches Handbuch, VS Verlag für Sozialwissenschaften, Wiesbaden, S. 486–497.

Andreas Balthasar

Evaluationen in der Schweiz: Verbreitung und Verwendung

„In the end, the worth of evaluations must be judged by their utility.“ (Rossi et al. 1999: 431)

Evaluationen werden durchgeführt, um zu verwendbaren Ergebnissen zu gelangen. Die Verwendungsorientierung gehört also mit zur Definition von Evaluation (Vedung 1999: 12). Die Frage nach dem Ausmass der Verwendung von Evaluationen durch Verwaltung und Politik ist somit ein Kernthema der Evaluationsforschung. Der vorliegende Beitrag geht am Beispiel der schweizerischen Bundespolitik auf diese Thematik ein; er verfolgt dazu drei Fragen vertieft:

Da die erste Bedingung für die Verwendung von Evaluationen deren Durchführung ist, wird zunächst die Verbreitung der Evaluationstätigkeit in der schweizerischen Bundesverwaltung untersucht. Danach wird die Frage nach der Verwendung von Evaluationen gestellt. Es wird geprüft, ob und in welchem Umfang die realisierten Evaluationen auch genutzt wurden, wobei kurz auf die Faktoren eingegangen wird, die die Verwendung beeinflussen. Schliesslich geht es um die Frage nach den Akteurgruppen, von denen die Evaluationen verwendet werden. Dabei interessieren insbesondere Politikerinnen und Politiker – diese signalisierten in der 1999 erneuerten schweizerischen Bundesverfassung ein besonderes Interesse an Evaluationen: „Die Bundesversammlung sorgt dafür, dass die Massnahmen des Bundes auf ihre Wirksamkeit überprüft werden“, wie Artikel 170 der Bundesverfassung festhält.

Angesichts der grossen Bedeutung der Verwendung von Evaluationen erstaunt es nicht, dass sich die evaluationswissenschaftliche Verwendungsforschung in den USA schon seit Jahren als eigenständiger Forschungszweig etabliert hat (vgl. z. B. Cousins/Leithwood 1986; Lee 2000). Auch in Europa sind verschiedene Arbeiten entstanden, die sich mit der Verwendung von Evaluationen auseinandersetzen (z. B. Valovirta 2002; Van der Knaap 1995). In der Schweiz gehen die ersten diesbezüglichen Arbeiten auf die 1980er Jahre zurück (Klöti/Schneider 1989; Zimmermann et al. 1988). In neuerer Zeit haben sich in der Schweiz insbesondere Thomas Widmer und Peter Neuenschwander mit der Institutionalisierung und der Verwendung von Evaluationen befasst (Widmer/Neuenschwander 2004; Widmer et al. 2001).

Grundlage des vorliegenden Beitrags bildet eine breit angelegte Untersuchung über Verbreitung und Verwendung von Evaluationen in der Schweizer Bundesverwaltung (Balthasar 2007). „Evaluation“ wurde dabei definiert als eine wissenschaftliche und empirisch gestützte retrospektive Beurteilung der Konzeption, des Vollzugs und/oder der Wirksamkeit staatlicher Aktivitäten, seien dies Massnahmen, Programme oder Projekte. Evaluationen bewerten staatliche Aktivitäten nach transparenten Kriterien und stellen Kausalzusammen-

hänge zwischen Aktivitäten und Wirkungen fest. Die Untersuchung beinhaltet neben zehn Fallstudien auch eine Erhebung und Befragung zu allen Evaluationen, welche die Bundesverwaltung in den Jahren 1999 bis 2002 abgeschlossen hat. Kontaktiert wurden die Verantwortlichen aller rund 300 in diesen Jahren abgeschlossenen Evaluationen. Bei diesen Personen handelt es sich nicht um *Evaluationsdurchführende*, sondern um *Evaluationsbetreuende*. Anders als in den bisherigen Arbeiten zur Verwendungsforschung wurde nicht eine allgemeine Einschätzung der Nutzung von Evaluationen vorgenommen, sondern es wurde nach deren konkreten Verwendung gefragt. Es wurden zudem nicht nur Evaluationsverantwortliche, sondern auch weitere mögliche Nutzerinnen und Nutzer der Evaluation befragt, wobei sich herausstellte, dass sich die Ergebnisse der beiden Gruppen kaum unterschieden.

1 Verbreitung von Evaluationen

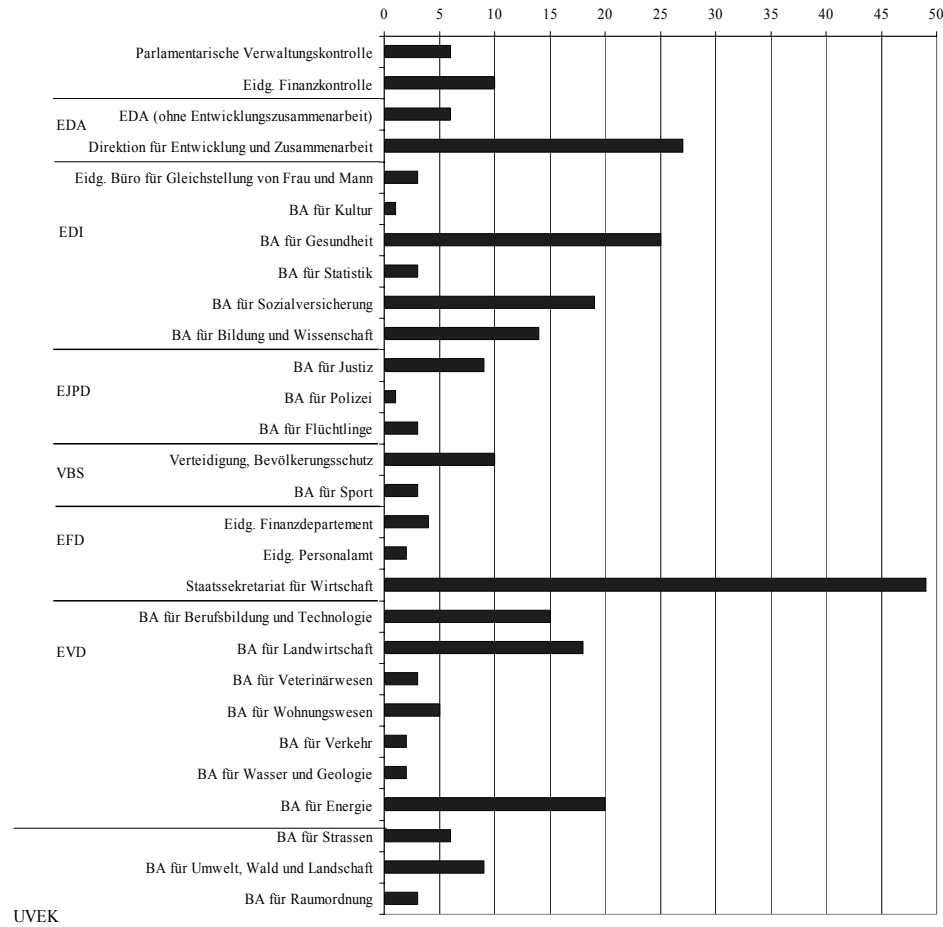
Der Ursprung der Evaluationstätigkeit in der Schweiz liegt zum einen im Nationalen Forschungsprogramm 6 „Entscheidungsvorgänge in der schweizerischen Demokratie“ (Linder 1987) und zum andern in der Arbeitsgruppe Gesetzesevaluation des Eidgenössischen Justiz- und Polizeidepartements (Arbeitsgruppe Gesetzesevaluation 1991). Zimmermann zählte für die Zeit von 1979 bis 1988 42 Evaluationen, die von der Bundesverwaltung in Auftrag gegeben worden waren (Zimmermann 1990: 80). Die diesem Beitrag zugrunde liegende Befragung geht von 308 Evaluationen in den Jahren 1999 bis 2002 aus. Der Vergleich der Zahlen weist auf ein starkes Wachstum der Evaluationstätigkeit in den letzten zwanzig Jahren hin. Tatsächlich hat die Institutionalisierung der Evaluation in der Schweiz auf Bundesebene einen hohen Grad erreicht. Zusammen mit Australien wird der Eidgenossenschaft diesbezüglich sogar ein Spitzenplatz zugeschrieben (Mader 2007). Treibende Kraft dieser Entwicklung waren nicht zuletzt die Bemühungen um eine qualitativ gute Gesetzgebung sowie um eine wirkungsorientierte Verwaltungsführung und -kontrolle.

Die Intensität der Evaluationstätigkeit unterscheidet sich jedoch innerhalb der Bundesverwaltung beträchtlich. In der Befragung wurden 45 Institutionen der Bundesverwaltung berücksichtigt. In rund einem Drittel dieser Institutionen wurden im Untersuchungszeitraum keine Evaluationen durchgeführt. Dazu gehören unter anderen die Bundeskanzlei, das Bundesamt für Polizei und das Bundesamt für Kommunikation. In den übrigen 29 Ämtern und Dienststellen wurden zwischen einer und 49 Evaluationen gezählt. Dies macht die nachfolgende Abbildung 1 deutlich. Sie listet für alle befragten Dienststellen der Bundesverwaltung die Zahl der Evaluationen auf, zu denen der Fragebogen beantwortet wurde. Weil die von uns durchgeführte Befragung einen Rücklauf von über neunzig Prozent erzielte, können wir die Zahl der Evaluationen, die wir in unsere Untersuchung einbeziehen konnten, mit der Zahl der durchgeführten Evaluationen gleichsetzen.

Wie Abbildung 1 zeigt, sind am meisten Evaluationen im Staatssekretariat für Wirtschaft abgewickelt worden. Rund ein Sechstel aller Evaluationen stammt aus dieser Einheit. Sie befassen sich mit wirtschaftspolitischen Themen im Inland und mit der Entwicklungszusammenarbeit. Zahlreiche Evaluationen wurden auch bei der Direktion für Entwicklung und Zusammenarbeit, beim Bundesamt für Gesundheit, beim Bundesamt für Sozialversicherung, beim Bundesamt für Landwirtschaft und beim Bundesamt für Energie durchge-

führt. 22 Ämter oder Dienststellen haben zwischen einer und zehn, deren sieben zwischen elf und 49 Evaluationen realisiert.

Abbildung 1: Übersicht über die Anzahl der Evaluationen nach Bundesämtern und Dienststellen (1999 bis 2002)



Legende: Grundgesamtheit ist die Anzahl der Antworten in der von Balthasar (2007) durchgeführten Befragung von Evaluationsverantwortlichen.

Einfache Erklärungen für die Unterschiede zwischen den Dienststellen gibt es nicht. Kissling-Näf und Knoepfel gehen davon aus, dass Ämter, die mit dem Management mehrerer Politikbereiche betraut sind, für Lernprozesse offener sind als solche, die nur einen einzigen „verwalten“. Für die beiden Autoren spielen aber auch Flexibilität und Aufnahmefähigkeit, das Alter der Institution und des Politikbereichs sowie der Politiktyp eine Rolle (Kissling-Näf/Knoepfel 1994: 123). Zum Teil lässt sich die unterschiedliche Verbreitung von Evaluationen auf den unterschiedlichen Legitimationsbedarf der Politikbereiche zurückführen. Da-

bei ist anzunehmen, dass der Legitimationsbedarf neuer Bereiche, wie sie die Umwelt-, die Drogen- oder die Energiepolitik darstellen, grösser ist als bei den traditionellen, etwa der Verteidigung, der Landwirtschaft oder der Alterssicherung. Allerdings sind auch die Kultur- oder die Migrationspolitik Bereiche mit hohem Legitimationsbedarf, obwohl sie nur sehr wenige Evaluationen aufweisen. Weiter werden programmatisch konzipierte Politiken wohl eher evaluiert als der klassische Gesetzesvollzug. Dies könnte die hohe Evaluationsintensität der Direktion für Entwicklung und Zusammenarbeit und des Bundesamts für Gesundheit erklären. Allfällige Einflüsse dieser Art werden aber überlagert von situationspezifischen Gegebenheiten, wie sie etwa bei Evaluationsprogrammen auftreten können, die in der Schweiz im Bereich der Berufsbildung oder der Krankenversicherung durchgeführt wurden (Bussmann 1997: 32).

Die Verbreitung von Evaluationen spiegelt sich nicht nur in der Zahl der Evaluationen pro Dienststelle wider, sondern auch im Anteil des Budgets, das für Evaluationen aufgewendet wird. Die Untersuchung zeigte, dass 28.4 Prozent der Evaluationen interne und externe Kosten von unter 50'000 Franken auslösten. In 26.6 Prozent der Fälle lagen die Kosten zwischen 50'000 und 99'999 Franken. 21.2 Prozent der Evaluationen kosteten zwischen 100'000 und 250'000 Franken, während die Kosten bei 15.5 Prozent der Evaluationen höher veranschlagt wurden.

Mit bestimmten Annahmen konnte errechnet werden, dass für die in die Untersuchung einbezogenen Evaluationen insgesamt rund 32 Millionen Franken aufgewendet wurden.¹ Für die Zeit von 1999 bis 2002 können die Evaluationsaufwendungen der Bundesverwaltung pro Jahr somit auf rund 8 Millionen Franken geschätzt werden. Dieser Betrag entspricht rund 0.02 Prozent der Ausgaben des Bundes in diesen Jahren. Aus der nachfolgenden Abbildung 2 wird ersichtlich, dass die prozentualen Ausgaben der Bundesämter und Dienststellen für Evaluationen überall unter einem Prozent lagen.²

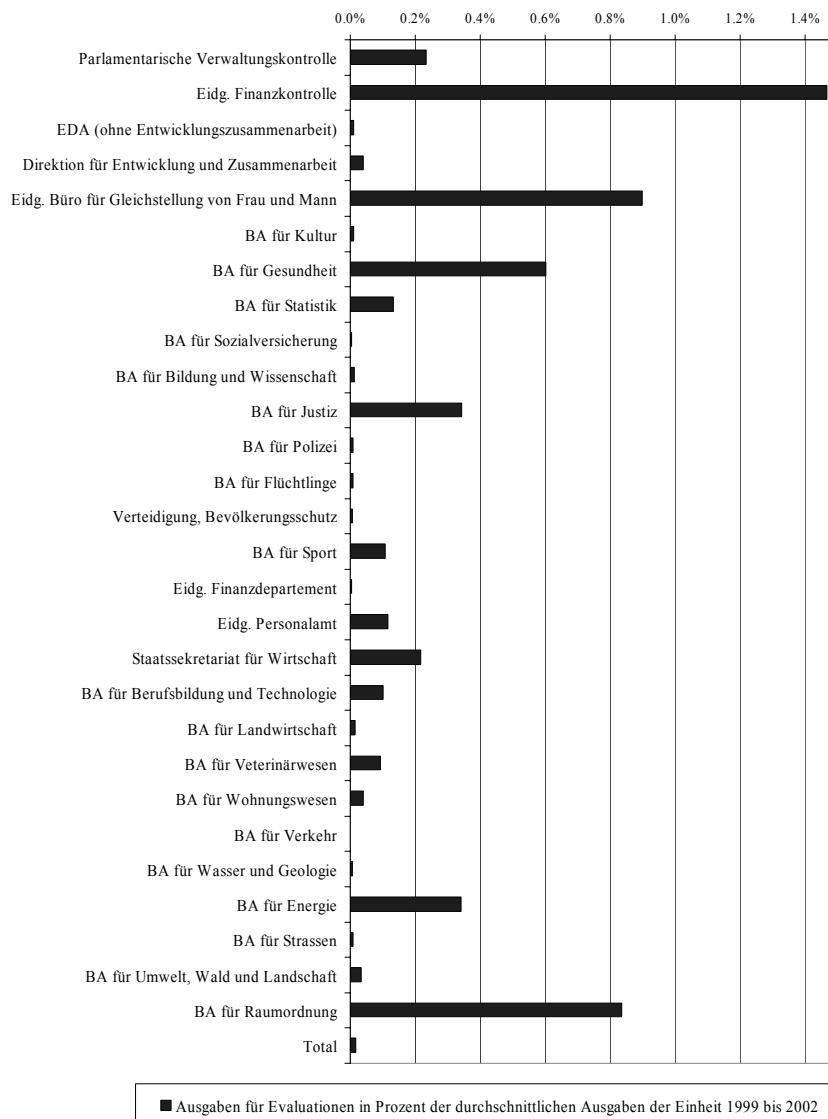
Die Abbildung 2 macht deutlich, dass sich auch die finanziellen Aufwendungen der Ämter und Dienststellen für Evaluationen stark unterscheiden. Im Energiebereich etwa wird im Verhältnis zum Budget viel evaluiert. Dies liegt daran, dass das im Untersuchungszeitraum laufende Programm „Energie 2000“ systematisch beurteilt wurde. Im Bundesamt für Justiz dagegen hatte die Evaluation im Massnahmenvollzug deshalb grosse Bedeutung, weil dort unterschiedliche Modelle erprobt wurden. Das Bundesamt für Gesundheit wiederum legt seit Jahren grossen Wert auf die Evaluation und verfügt über den verwaltungsintern am besten ausgebauten Evaluationsdienst. Das grosse relative Gewicht, das der Evaluation in der Raumordnungs- und der Gleichstellungspolitik zukommt, erklärt sich nicht etwa durch grosse Evaluationsbudgets, sondern gerade durch die bescheidenen Mittel, welche diesen Politikbereichen generell zur Verfügung stehen. Die Tatsache, dass das Eidgenössische Büro für die Gleichstellung von Frau und Mann mit 0.9 Prozent der gesamten Aufwendungen den grössten Anteil der verfügbaren Mittel für Evaluationen verwendet,

¹ Es wurden folgende Annahmen getroffen: Die Durchschnittskosten einer Evaluation von unter 50'000 Franken sind 40'000 Franken, jene einer Evaluation zwischen 50'000 und 99'999 Franken sind 80'000 Franken, jene einer Evaluation zwischen 100'000 und 250'000 Franken sind 120'000 Franken, jene einer Evaluation von über 250'000 Franken sind 350'000 Franken. Für die Evaluationen ohne Kostenangabe wurde angenommen, dass diese 40'000 Franken gekostet haben.

² Die Eidgenössische Finanzkontrolle (EFK) bildet hier als Kontrollorgan eine Ausnahme.

muss vor diesem Hintergrund also relativiert werden. Die relativ grossen Aufwendungen der Eidgenössischen Finanzkontrolle für die Evaluation erklären sich durch deren Aufgabe als Kontrollorgan des Bundes.

Abbildung 2: Geschätzte prozentuale Ausgaben für Evaluationen nach Bundesämtern und Dienststellen



Legende: Geschätzte Ausgaben für Evaluationen in Prozent der Ausgaben der Ämter und Dienststellen der Jahre 1999 bis 2002 gemäss Schweizerische Eidgenossenschaft (2003) (vgl. Balthasar 2007)

Immer wieder wird auch die Frage aufgeworfen, wie gross die Aufwendungen für Evaluationen etwa sein müssen, um eine ausreichende Beurteilung der Wirkungen zu ermöglichen. Eine der ersten systematischen Auseinandersetzungen mit dieser Thematik stammt von Marvin Alkin und Brian Stecher (Alkin/Stecher 1983). Sie setzt sich vornehmlich mit der Zusammensetzung der Aufwendungen auseinander, nennt aber auch die prozentualen Aufwendungen verschiedener öffentlicher Institutionen für Evaluation. Diese liegen bei zwischen einem und zehn Prozent des Budgets (Alkin/Stecher 1983: 120). Das schweizerische Bundesamt für Gesundheit sieht zehn bis 15 Prozent der Programmkosten für die Evaluation vor (BAG 1997: 24). Das Aktionsprogramm Energie 2000 wendete etwa 1.3 Prozent des gesamten Budgets für Evaluationen auf (Balthasar 2000: 210).

Gemäss Auskunft des Leiters der zentralen Evaluationsstelle der für den Haushalt der Europäischen Gemeinschaft zuständigen Direktion gab die EU in den letzten Jahren zwischen hundert und 150 Millionen Euro für Evaluationen aus, dies bei einem Budget von 109 Milliarden Euro. Evaluationen machen somit zwischen 0.09 und 0.14 Prozent des EU-Haushalts aus.³ Die ermittelten Ausgaben des Bundes für die 1999 bis 2002 durchgeführten Evaluationen liegen bei durchschnittlich 0.02 Prozent und damit unter sämtlichen genannten Richtgrössen. Im Vergleich zur EU wird in der Schweiz somit relativ wenig Geld für Evaluationen eingesetzt. Ob die Mittel generell ausreichend, zu hoch oder zu tief sind, lässt sich jedoch so nicht beurteilen, weil dies wesentlich von der Fragestellung und von der Zielsetzung der jeweiligen Evaluation abhängt: Pilotprojekte, die im Erfolgsfall flächendeckend eingeführt werden sollen, müssen intensiver analysiert werden als Projekte, die sich bereits seit Jahren bewährt haben; politisch umstrittene Massnahmen sind auf umfangreichere Evaluationen angewiesen als Massnahmen, die über eine breite Akzeptanz verfügen; Programme, die seit langem laufen, aber noch nie evaluiert wurden, müssen eher evaluiert werden als Programme, die vor kurzem bereits beurteilt wurden.

2 Verwendung von Evaluationen

Im Zentrum der Befragung der Evaluationsverantwortlichen der schweizerischen Bundesverwaltung standen Fragen im Zusammenhang mit der Verwendung der durchgeführten Evaluationen und deren Ergebnissen. Untersucht wurden insbesondere das Ausmass der Verwendung sowie die Faktoren, welche die Verwendung beeinflussen.

2.1 Ausmass der Verwendung

65.8 Prozent der befragten Evaluationsverantwortlichen gingen von einem hohen oder eher hohen Grad der Verwendung der Evaluation aus, die sie begleitet haben. Dieses Resultat ist überraschend hoch, wird aber durch die Ansichten der befragten weiteren Personen, die als

³ Information des Leiters der zentralen Evaluationsstelle der Generaldirektion Haushalt der Europäischen Gemeinschaft anlässlich der Tagung „Evaluation in der Europäischen Gemeinschaft – Aktuelle Praxis und zukünftige Trends“ der Schweizerischen Evaluationsgesellschaft vom 24. November 2004 in Bern.

mögliche Nutzende der Evaluation in Frage kamen, bestätigt. Von dieser Gruppe sprachen mehr als sechzig Prozent von einer (eher) intensiven Verwendung. Nur zwischen 25 und dreissig Prozent aller Evaluationen waren wenig oder gar nicht verwendet worden. Zwar gibt es in der internationalen Literatur keine breit abgestützten Zahlen, die als Vergleichsgrössen herangezogen werden könnten. Die verfügbaren Hinweise lassen aber den Schluss zu, dass die erhobene Verwendungsintensität auch im internationalen Vergleich als ausgesprochen hoch beurteilt werden kann (Weiss 1998b). Dieses Ergebnis lässt sich einerseits darauf zurückführen, dass es der methodische Ansatz erlaubt hat, sehr gezielt nach den Effekten von einzelnen Evaluationen zu fragen. So wurden nicht bloss generelle Einschätzungen, sondern ganz spezifische Erfahrungen ermittelt. Auf der anderen Seite kann es aber auch sein, dass der gewählte Zugang über die für die Evaluation verantwortlichen Personen zu einer zu positiven Darstellung der Verhältnisse geführt hat, da diese oft selbst in die Durchführung der Evaluation oder der evaluierten Massnahme involviert waren. Allerdings wird die Glaubwürdigkeit der Einschätzungen durch die Tatsache unterstrichen, dass sich die Einschätzungen der ebenfalls befragten weiteren potenziellen Nutzenden nicht wesentlich von denjenigen der Evaluationsverantwortlichen unterschieden.

Die Untersuchung beschränkte sich nicht auf eine generelle Einschätzung der Verwendung der untersuchten Evaluationen, vielmehr wurden auch unterschiedliche Arten der Verwendung ermittelt. Dabei wurde unterschieden zwischen instrumenteller, konzeptioneller, prozessbezogener und symbolischer Verwendung. Während es bei der instrumentellen Verwendung um die direkte praktische Nutzung wissenschaftlicher Erkenntnisse geht, spricht die konzeptionelle Verwendung die indirekten Wirkungen an, und zwar durch „gradual sedimentation of insights, theories, concepts and ways of looking at the world“ (Weiss 1977: 535). Diese Form der Verwendung kommt dadurch zustande, dass die Informationen, die eine Evaluation bereitstellt, von den Entscheidungsverantwortlichen stets im Lichte ihrer eigenen Erfahrungen, ihres Wissens und ihrer Bedürfnisse betrachtet und interpretiert werden (Ledermann 2004: 13). Von prozessbezogener Verwendung wird dann gesprochen, wenn sich unter den am Evaluationsprozess Beteiligten ein gemeinsames Verständnis bezüglich relevanter Probleme im Bereich der untersuchten Massnahme entwickelt. Prozessbezogene Verwendung kommt dadurch zustande, dass Akteure allein durch die Teilnahme am Evaluationsprozess – und nicht erst aufgrund der Ergebnisse – Einblicke in das Programm gewinnen, die ihr Verhalten beeinflussen (Vedung 1999: 242). Von symbolischer Verwendung von Evaluationsergebnissen wird gesprochen, wenn Entscheidungstragende Evaluationen dazu nutzen, ihre Optik zu bestätigen und sich Legitimation zu verschaffen (Henry/Rog 1998: 90). Symbolische Verwendung kann eine positive Konnotation haben, wenn die bestehende positive Einschätzung bestätigt wird. Sie kann aber auch negativ besetzt sein: dann nämlich, wenn die Verantwortlichen aus einer Evaluation gar nichts lernen wollen. In der nachfolgenden Abbildung 3 sind die Resultate der Befragung zu den verschiedenen Arten der Nutzung von Evaluationen abgebildet.

Abbildung 3: Ausmass der Verwendung von Evaluationen

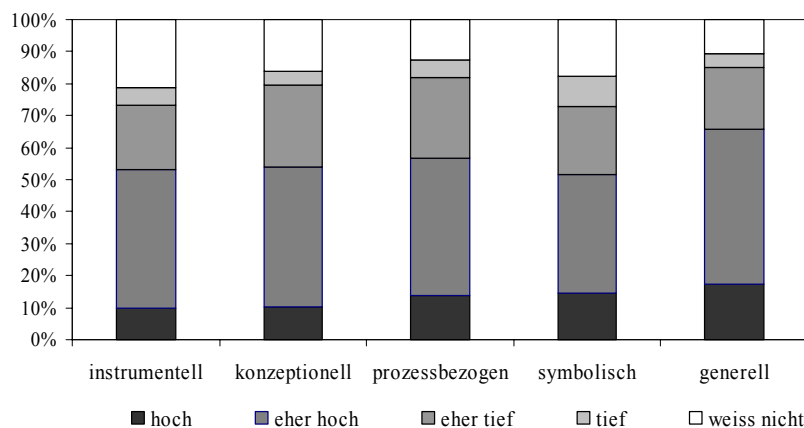


Abbildung 3 zeigt, dass das Ausmass der *prozessbezogenen* Verwendung als überdurchschnittlich hoch eingeschätzt wurde; dieses war gemäss den Angaben der Evaluationsverantwortlichen in 56.8 Prozent der Fälle „hoch“ oder „eher hoch“. Personen, die in die Evaluationsprojekte involviert waren, unterstreichen damit, dass sie im Prozess der Evaluation viel gelernt haben, und dass sie das Gelernte für ihre zukünftige Arbeit brauchen können. Auch die Verwendung der Empfehlungen, das heisst die *instrumentelle* Verwendung, wird als relativ hoch beurteilt. Die Evaluationsverantwortlichen gaben an, dass in 257 von 278 Evaluationen Empfehlungen formuliert worden sind. Bei 136 der 278 Evaluationen wurde von einer (eher) hohen instrumentellen Verwendung ausgegangen, da mehr als die Hälfte der Empfehlungen umgesetzt wurde (48.9 Prozent aller Evaluationen). In 66 Fällen (23.7 Prozent aller Evaluationen) wurde die instrumentelle Verwendung als „(eher) tief“ eingeschätzt, da weniger als die Hälfte der Empfehlungen umgesetzt wurde. Somit trat eine instrumentelle Verwendung sehr häufig auf. Damit bestätigt sich die Beobachtung von Stamm, dass „direkte Rezeption und Nutzung deutlich häufiger festgestellt werden kann als in der einschlägigen theoretischen Literatur behauptet“ (Stamm 2003: 326).

Eine *konzeptionelle* Verwendung trat mindestens ebenso häufig auf wie die instrumentelle. Von den Evaluationsbeauftragten gaben rund 54 Prozent an, dass die konzeptionelle Verwendung „hoch“ oder „eher hoch“ war. Carol Weiss und andere Autoren haben die Ansicht vertreten, dass die Verwendung von Evaluationen in erster Linie über das „enlightenment“ erfolge, also eine konzeptionelle sei (Weiss 1977). Unsere Arbeit zeigt dagegen, dass die konzeptionelle Verwendung nicht häufiger auftrat als die instrumentelle.

Der bedeutendste Unterschied zwischen den Evaluationsverantwortlichen und den weiteren potenziell Nutzenden zeigte sich bei der Einschätzung der *symbolischen* Verwendung. Über fünfzig Prozent der Evaluationsverantwortlichen schätzten die symbolische Verwendung als „(eher) hoch“ ein; bei den weiteren Nutzenden lag dieser Anteil bei fast 65 Prozent. Ob die konstatierte symbolische Verwendung positiv oder negativ konnotiert war, liess sich aus der Befragung nicht eruieren.

2.2 Einflussfaktoren auf die Verwendung

Die Verwendung der Ergebnisse von Evaluationen wird von ganz unterschiedlichen Faktoren bestimmt. Dies machten zahlreiche wissenschaftliche Untersuchungen deutlich (Balthasar 2007). Die diesem Beitrag zugrunde liegende Untersuchung ging von einem Wirkungsmodell aus, das die Verwendung von Evaluationen als Ergebnis des Zusammenspiels von institutionellen Faktoren versteht, nämlich: die Lancierung und Durchführung einer Evaluation, sodann Kontextfaktoren, die das Umfeld der Evaluation betreffen, und schliesslich prozessbezogene Faktoren. Das Modell knüpft an die Arbeiten von Cousins und Leithwood (1986; 1993) an, die kontext- und prozessbezogene Faktoren unterscheiden.

Als von herausragender Bedeutung für die Verwendung einer Evaluation hat sich das Interesse der Entscheidungstragenden an ihr erwiesen. War ein solches vorhanden, verbesserten sich die Chancen auf eine Verwendung massiv. Zudem müssen die Evaluationen so angelegt sein, dass ihnen ein Umsetzungspotenzial zugesprochen wird. Für die Befragten war also wichtig, dass die Evaluation die richtigen Fragen beantwortete, dass sie die Informationen zum richtigen Zeitpunkt lieferte, und dass der Sprachstil, in der die Evaluation abgefasst war, für die Nutzenden verständlich und hilfreich war. Schliesslich hat sich auch eine positive Einschätzung des Kosten-Nutzen-Verhältnisses der Evaluation günstig auf die Verwendung ausgewirkt.

Neben diesen prozessbezogenen Faktoren spielt auch die Institutionalisierungsform eine Rolle. Die Verwendung von Evaluationen scheint jedoch weniger von der Distanz zwischen Evaluierenden und Evaluierten an sich abzuhängen als vielmehr von der Ausprägung einzelner Faktoren, die diese Distanz bestimmen. So haben die Ergebnisse der empirischen Analyse gezeigt, dass die instrumentelle Verwendung, das heisst die Umsetzung der Empfehlungen einer Evaluation gefördert wird, wenn die verantwortlichen Institutionen die Kosten für die Evaluation selber tragen müssen. Demgegenüber ist eine ausgeprägt konzeptionelle Verwendung dann zu erwarten, wenn externe Evaluierende eingesetzt werden. Diese verfügen nämlich in der Regel über breite methodische Kenntnisse sowie Erfahrung in der Durchführung von Evaluationen, was für eine konzeptionelle Verwendung besonders wichtig ist. Die Untersuchung zeigt weiter, dass sich die prozessbezogene Verwendung signifikant verbesserte, wenn die Evaluation durch diejenigen Personen ausgelöst wurde, die sie später auch umsetzten, und/oder wenn sie amtsintern durchgeführt wurde.

Generell stützt unsere Arbeit die These der lernorientierten, interaktionistischen Evaluationstheorien, die besagen, dass eine intensive Verwendung der Resultate dann zu erwarten ist, wenn die Evaluierenden eng mit den Evaluierten zusammenarbeiten. Michael Pattons Forderung, Evaluationen sollten sich ganz an den Bedürfnissen der Nutzenden orientieren, gilt besonders für die prozessbezogene Verwendung (Patton 1997). Nur in Ausnahmefällen ist die Verwendung von Evaluationen stärker, wenn die Distanz zwischen Evaluierenden und Evaluierten zunimmt. Am ehesten ist dies bei Evaluationen der Fall, die konzeptionelle oder symbolische Verwendung finden. Allerdings sollten auch in diesen Fällen die Evaluationen so angelegt sein, dass ihnen von den möglichen Nutzenden ein Umsetzungspotenzial zugesprochen wird. Die in der Tradition von Donald Campbell – zum Beispiel von Davies, Nutley und Smith – vertretene Argumentation, dass die Distanz zwischen Evaluierenden und Evaluierten Unabhängigkeit signalisiere und damit die Akzeptanz der

Ergebnisse erhöhe (Campbell/Stanley 1963; Davies et al. 2000), kann aufgrund unserer Untersuchung am ehesten für die symbolische Verwendung bestätigt werden. Entgegen den Vermutungen von Ray Rist zeigt die Untersuchung ausserdem, dass die Glaubwürdigkeit der Information, verglichen mit der Orientierung an den spezifischen Bedürfnissen der Nutzenden, für die Verwendung eine untergeordnete Bedeutung hat (Rist 1994: 200).

3 Nutzung durch die Politik

Eine wichtige Zielgruppe von Evaluationen sind Politikerinnen und Politiker. Bei der Befragung der Evaluationsverantwortlichen haben wir festgestellt, dass bei rund zwanzig Prozent der untersuchten Evaluationen vorgesehen war, dass auch das Parlament deren Ergebnisse verwendet. Ein Beispiel dafür ist der Bericht „Das Kartellrecht: Standortbestimmung“, der von der Parlamentarischen Verwaltungskontrolle zuhanden der Geschäftsprüfungskommission des Nationalrates erarbeitet wurde (Parlamentarische Verwaltungskontrolle 2000). Die Verwendung dieses Berichts wurde in einer Fallstudie vertieft untersucht (Balthasar 2007). Dabei zeigte sich, dass die Geschäftsprüfungskommission mit dem Bericht unter anderem die Absicht verfolgte, „einen gewissen politischen Druck auf den Bundesrat auszuüben. Die Kartellgesetzesrevision war lange geplant, aber die Botschaft brauchte ziemlich viel Zeit, bis sie überhaupt erarbeitet wurde.“ Darüber hinaus wollte sich die Kommission mithilfe des Berichts in die Lage versetzen, die Qualität eines Entwurfs des Bundesrats zur Revision des Kartellgesetzes beurteilen zu können. Als nun aber der Bericht zur Diskussion vorlag, standen zusätzlich weitere Informationsquellen zur Entscheidungsfindung zur Verfügung, namentlich ein Vernehmlassungsentwurf des zuständigen Departements sowie eine vom Bundesrat veranlasste ergänzende Untersuchung. Gegenüber diesen beiden Informationsquellen kam der Bericht der Parlamentarischen Verwaltungskontrolle etwas spät: „Es gab eine Art Konkurrenz. Es gab keine grüne Wiese. [...] Wenn parallel zu einer Evaluation plötzlich etwas in der Verwaltung gemacht wird, wird es schwierig. Das ist von Zeit zu Zeit auch eine Taktik der Verwaltung: Man macht nichts und wenn man sieht, dass sich die Oberaufsicht um ein Thema kümmert, dann beginnt man, Aktivismus zu betreiben. Und wenn das Produkt dann vorliegt, kann man sagen: Wir haben schon alles vorgeschlagen.“

Vor diesem Hintergrund wurde der Bericht nach Meinung der Befragten eher wenig verwendet. Er sei wohl wahrgenommen und in der Botschaft des Bundesrates auch zitiert worden – die Entscheidungsfindung jedoch habe der Bericht kaum beeinflusst. Dessen Nützlichkeit wurde insbesondere durch den Umstand beeinträchtigt, dass die Verwaltung fast parallel dazu eigene Arbeiten durchführte und ein revidiertes Gesetz zur Stellungnahme unterbreitete. Dadurch kam die Kommission zeitlich unter Druck; sie konnte sich nicht wie sonst üblich vertieft mit dem Bericht auseinandersetzen. Immerhin nahm der Gesetzesvorschlag des Bundesrates in weiten Teilen die georteten Revisionsanliegen auf.

Das Beispiel zeigt, dass es nicht einfach ist, Evaluationen so anzulegen, dass Politikerinnen und Politiker diese auch verwenden. Zum einen stehen parlamentarische Organe in der Regel unter einem hohen Zeitdruck, dem die für Evaluationen notwendigen empirischen Arbeiten kaum entsprechen können. Zum anderen können Politikerinnen und Politiker oft auch auf andere wichtige Informationsinstrumente zur Entscheidungsfindung zu-

rückgreifen. In einer funktionstüchtigen Demokratie liefern auch Parteien und Interessengruppen, administrative und parlamentarische Aufsichtsorgane, Gerichte und Medien permanent Rückmeldungen über den Vollzug und die Wirkungen staatlicher Massnahmen (Janett 2003: 99).

Die Befragung der Evaluationsverantwortlichen hat dennoch ergeben, dass das Parlament die an es gerichteten Evaluationen mehrheitlich auch nutzt. 58.5 Prozent der Evaluationen mit dieser Zielgruppe haben gemäss Einschätzung der Evaluationsverantwortlichen dort auch Verwendung gefunden. Dieselbe Frage haben wir auch den weiteren Nutzenden gestellt. Obwohl in ihrer Einschätzung etwas zurückhaltender, gaben auch sie an, dass das Parlament 45 Prozent der an sie gerichteten Evaluationen genutzt habe. Aus der aktuellen Politik gibt es denn auch etliche Beispiele für die Verwendung von Evaluationen im Parlament. Dazu gehören etwa die Nationalratsdebatte über die Gleichstellung von Frau und Mann, die unter anderem auf der Grundlage eines Berichts des Bundesrates über die Evaluation der Wirksamkeit des Gleichstellungsgesetzes erfolgt ist. Ein weiteres Beispiel ist die Totalrevision des Opferhilfegesetzes, mit der sich der Ständerat befasst hat – sie wurde nicht zuletzt durch eine Evaluation des Vollzugs und der Wirksamkeit dieses Gesetzes ausgelöst (Mader 2007).

Insgesamt darf der Stellenwert von Evaluationen für die Politik aber nicht überschätzt werden. Politikerinnen und Politiker reagieren auf Evaluationen in erster Linie dann, wenn Missstände aufgedeckt werden, wenn sie in ihren eigenen Überzeugungen bestätigt werden, oder wenn sie sich mit Evaluationen profilieren können. Weil diese aber oft zu spät kommen, eignen sich Evaluationen kaum zur Profilierung. So reagieren Politikerinnen und Politiker weniger auf Evaluationen als vielmehr auf Krisen, auf Medienberichterstattungen und auf Interessengruppen, die an sie herantreten.

Obwohl sich Politikerinnen und Politiker also nur selten für die Ergebnisse einzelner Evaluationen interessieren, kann angenommen werden, dass sie den Wert von Vollzugs- und Wirkungsanalysen für die Verwaltung durchaus anerkennen. Politikerinnen und Politiker scheinen Evaluationen grundsätzlich als nützlich zu betrachten. Bestätigt wird dies nicht zuletzt durch die Tatsache, dass der Artikel 170 der Bundesverfassung im Vorfeld der Abstimmung kaum zu Diskussionen Anlass gab.

4 Schluss

Der vorliegende Beitrag setzt sich mit der Verbreitung und Nutzung von Evaluationen in der schweizerischen Bundesverwaltung auseinander. Aufgrund einer umfangreichen Erhebung und Analyse aller zwischen 1999 und 2002 abgeschlossenen Evaluationen konnte festgestellt werden, dass in der Schweiz auf nationaler Ebene eine rege Evaluationstätigkeit herrscht. Von einer flächendeckenden Evaluationskultur kann dennoch nicht gesprochen werden. Vielmehr variiert der Grad der Evaluationstätigkeit innerhalb der Bundesverwaltung sehr stark. Rund ein Drittel der Bundesämter und Dienststellen hat im Untersuchungszeitraum keine Evaluation durchgeführt. Bei einem weiteren Drittel kann höchstens von einer sporadischen Evaluationstätigkeit ausgegangen werden. Ein Drittel der Einheiten der Bundesverwaltung hingegen führt regelmässig Evaluationen durch – diese haben sich so in

den letzten Jahren wertvolle Erfahrung in Planung, Begleitung und Umsetzung von Wirkungsuntersuchungen aneignen können. Ein Vergleich der Ausgaben des Bundes für Evaluationen mit jenen der Europäischen Kommission hat allerdings ergeben, dass in der Schweiz relativ wenig Geld für Evaluationen ausgegeben wird.

Wie unsere Untersuchung zeigt, bescheinigen sowohl die befragten Evaluationsverantwortlichen als auch die weiteren möglichen Nutzenden den durchgeführten Evaluationen mehrheitlich eine gute Verwendung. Knapp zwei Drittel der befragten Evaluationsverantwortlichen gingen von einem hohen oder eher hohen Grad der Verwendung aus. Dabei waren es vor allem drei Faktoren, die sich günstig auf die potenzielle Verwendung einer Evaluation ausgewirkt haben: das Interesse der Entscheidungsträger an ihr, eine positive Beurteilung ihres Umsetzungspotenzials sowie eine positive oder ausgeglichene Kosten-Nutzen-Bewertung. Was die institutionellen Rahmenbedingungen betrifft, scheint die Verwendung von Evaluationen weniger von der Distanz zwischen Evaluierenden und Evaluierten an sich abzuhängen, als vielmehr von gewissen Faktoren, die diese Distanz bestimmen. So werden etwa die Empfehlungen einer Evaluation viel eher umgesetzt, wenn die für die Umsetzung verantwortlichen Ämter oder Dienststellen die Kosten der Evaluation selber tragen müssen.

Evaluationen richten sich in erster Linie an die Verwaltung. Von dieser Akteurgruppe werden sie denn auch häufig verwendet. Hingegen scheint es schwierig zu sein, Evaluationen so anzulegen, dass sie auch von Politikerinnen und Politikern verwendet werden. Evaluationen stehen nicht zuoberst auf deren Prioritätenliste, zumal, wenn es sich um retrospektive, das heisst die Vergangenheit betreffende Evaluationen handelt: „Politicians have to muddle through an information jungle – acting more like artists, using intuition and judging all sorts of available information“ (Lemne/Sohlman 2004: 10). Politische Entscheide fallen nach politischen Kriterien, sachliche Kriterien sind relevant, aber nicht dominant.