



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
Main Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 1996

---

**Evaluationen und ihre Qualität: Meta-Evaluation als Beitrag zur  
Qualitätssicherung in der Evaluationsforschung**

Widmer, Thomas

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-165428>

Published Research Report

Published Version

Originally published at:

Widmer, Thomas (1996). Evaluationen und ihre Qualität: Meta-Evaluation als Beitrag zur Qualitätssicherung in der Evaluationsforschung. Bern: Schweizerischer Nationalfonds.

1P2 NFP 27.27

Nationales Forschungsprogramm 27  
**Wirksamkeit staatlicher Massnahmen**

Reihe «Kurzberichte»

Thomas Widmer

# Evaluationen und ihre Qualität

**Meta-Evaluation als Beitrag  
zur Qualitätssicherung  
in der Evaluationsforschung**

Bibliothek  
**Politische Wissenschaft**

**Expertengruppe des NFP 27**

Ernst Buschor (Präsident), Heinz Hausheer, Ulrich Klöti, Wolf Linder, Luzius Mader,  
Pierre Moor, Kurt Nuspliger, Gerhard Schmid, Danielle Yersin

**Programmleitung**

Werner Bussmann, c/o Postamt, 3003 Bern 3 Bundeshaus, Tel. 031 322 47 98

Bern, 1996

431622

Die Studie, die im Auftrag des Nationalen Forschungsprogramms «Wirksamkeit staatlicher Massnahmen» (NFP 27) am Institut für Politikwissenschaft der Universität Zürich erarbeitet wurde und über die hier in Kurzform berichtet wird, befasst sich mit der Qualität von sozialwissenschaftlichen Evaluationen. Um dieses Thema anzugehen, musste in einem ersten Schritt bestimmt werden, was eine gute Evaluation auszeichnet und wie die Qualität einer spezifischen Evaluationsstudie erfasst werden kann. Zweitens behandelte die Studie die Frage nach der aktuellen Qualität der schweizerischen Evaluationspraxis. Weiter interessierte, wie in der gegenwärtigen Evaluationspraxis die Qualität verbessert werden kann.

Diese Fragen bearbeiteten wir anhand einer Meta-Evaluation, also einer Evaluation von Evaluationen. Konkret haben wir zehn sehr unterschiedliche Evaluationsstudien einer Evaluation unterzogen. Dies erfolgte in einem ersten Schritt in der Form von Fallstudien, in denen eine eingehende Qualitätsbewertung der einzelnen Evaluationen vorgenommen wird. Zweitens wurden auf dieser Grundlage die zehn Evaluationsstudien und deren Bewertungen miteinander verglichen. Daraus wurden - drittens - Folgerungen bezüglich der oben dargelegten Fragen gezogen.

## Die evaluierten Evaluationsstudien

Folgende zehn Evaluationsstudien wurden in einer eingehenden, empirischen Analyse einer Qualitätsbewertung unterzogen:

1. «Umweltpolitik und technische Entwicklung» von Andreas Balthasar und Carlo Knöpfel (1994)
2. «Unternehmerische Innovationsprozesse» von Willy Bierter und Hans-Martin Binder (1993)
3. «Wohneigentumsförderung durch den Bund» von Hans-Rudolf Schulz, Christoph Muggli und Jörg Hübschle (1993)
4. «Bedingt wirksam: Das Lohngleichheitspostulat des Heimarbeitsgesetzes» von Frohmüt W. Gerheuser und André Schmid (1993)
5. «Evaluation der Zürcher Gemeinschaftszentren (ZGZ)» von Andreas Näf et al. (1993)
6. «Schneeräumung von Urner Passstrassen» von Heini Sommer und Stefan Suter (1993)
7. «Evaluation Kehrichtsackgebühr» von Barbara Haering et al. (1990)
8. «Verbrauchsabhängige Heizkostenabrechnung – Förderung erneuerbarer Energien» von Wolf Linder et al. (1990)
9. «Evaluation des BUWAL» von Peter Knoepfel und Willi Zimmermann (1991)
10. «Evaluierung der bundeseigenen EWR-Informationenkampagne» von Claude Longchamp (1993)

Bei den ersten vier Evaluationen handelt es sich um Forschungsprojekte, die im Rahmen des NFP 27 entstanden sind. Die Studien 5 und 6 sind im Auftrag der zu-

ständigen Stellen erarbeitet worden, wobei sich das NFP 27 an der Finanzierung dieser Studien unter dem Titel «Kurzevaluationen» mitbeteiligt hat. Die Evaluationen 7 und 8 sind im Auftrag der Arbeitsgruppe Gesetzesevaluation (AGEVAL) des EJPD realisiert worden. Die letzten zwei Studien haben die Geschäftsprüfungskommission des Nationalrates bzw. das Integrationsbüro (EDA/EVD) veranlasst.

Bei der Auswahl wurde darauf geachtet, eine möglichst grosse Vielfalt an Evaluationsstudien zu erfassen. Wie aus der Liste der Titel der Schlussberichte (siehe Kasten) hervorgeht, stammen die evaluierten Studien aus diversen Politikfeldern und bewegen sich auf nationaler, kantonaler oder kommunaler Ebene. Sie verwenden sehr verschiedenartige Vorgehensweisen und unterscheiden sich auch hinsichtlich ihres zeitlichen und finanziellen Umfangs deutlich. Die evaluierenden Forscher sind in einem privaten oder in einem universitären Umfeld tätig und stammen zudem aus verschiedenen Disziplinen. Auch die den Studien zugrundeliegenden Erkenntnisinteressen und die zu bearbeitenden Fragestellungen variieren stark. Die Evaluationen befassen sich mit ganzen politischen Programmen, ausgewählten Einzelmassnahmen oder auch mit Institutionen. Sie sind teilweise summativ, teilweise aber auch formativ angelegt. Neben Studien, die eine rein quantitative oder qualitative Methodik einsetzen, enthält unsere Auswahl auch Evaluationen mit einem gemischten Ansatz (Methodenmix, Triangulation).

## Die Bewertungskriterien

Um eine Bewertung der einzelnen Evaluationen vornehmen zu können, musste ein Instrument eingesetzt werden, durch das die Qualität einer Evaluation genauer definiert und empirisch erfasst werden kann. Als Grundlage verwendeten wir einen Katalog von sogenannten Standards, der ursprünglich zu Beginn der 80er Jahre in den USA entwickelt wurde (vgl. Joint Committee on Standards for Educational Evaluation: The Program Evaluation Standards. 2nd Edition. Thousand Oaks: Sage 1994). Zur empirischen Bewertung der zehn Evaluationsstudien wurde ein auf dieser Vorgabe basierender Kriterienkatalog eingesetzt. Dieser besteht aus vier Kriteriengruppen, die sich mit der Nützlichkeit, der Anwendbarkeit, der Korrektheit sowie der Genauigkeit von Evaluationen befassen. Alle vier Gruppen zusammen umfassen insgesamt dreissig Einzelkriterien. Um den Kriterienkatalog auch auf Evaluationsstudien anwenden zu können, die einem alternativen erkenntnistheoretischen Ansatz folgen und damit einige der traditionellen wissenschaftlichen Qualitätspostulate (Validität, Reliabilität, Objektivität) negieren, wurde der Kriterienraster in einigen Punkten um Alternativkriterien ergänzt. Dabei stützten wir uns auf Vorarbeiten von Egon G. Guba und Yvonna S. Lincoln (Fourth Generation Evaluation. Newbury Park: Sage 1989: 236-243). So waren wir in der Lage, angepasst auf die erkenntnistheoretische Ausrichtung der Evaluationsstudie, ein alternativ-kriterium zu ersetzen. Bei den Kriterien G.5. und G.9. (siehe die nachfolgende Übersicht über die Kriterien) kommen je nach erkenntnistheoretischer Position der

Evaluationen entweder die Kriterien (a) (für traditionelle Positionen) oder (b), (c) sowie (d) für alternative Ausrichtungen in Betracht.

Die Anwendung des Kriterienrasters auf die zehn Evaluationsstudien geschah nicht zuletzt auch mit der Absicht, die Tauglichkeit der Kriterien zur Qualitätserfassung zu überprüfen. Zudem sollte der Einsatz des Rasters auch zeigen, ob und wenn ja welche Änderungen am Kriterienraster erforderlich sind, um die Adäquatheit des Rasters zu verbessern. Die empirische Umsetzung zeigte, dass sich der Kriterienraster grundsätzlich bewährt, dass aber trotzdem noch einige kleinere Anpassungen nötig sind. Im Rahmen des Projektes haben wir, basierend auf den in den Fallstudien gewonnenen Erfahrungen einen revidierten Kriterienkatalog ausgearbeitet, der folgendermassen ausgestaltet ist:

### Kriterienraster zur Bewertung von Evaluationen

#### **Nützlichkeit (N.)**

*Identifikation der beteiligten Gruppen (N.1.):* Die Beteiligten oder Betroffenen einer Evaluation sollten identifiziert werden, damit deren Interessen berücksichtigt werden können.

*Glaubwürdigkeit des Evaluators (N.2.):* Die Personen, welche die Evaluation durchführen, sollten sowohl vertrauenswürdig wie auch kompetent sein, damit die Ergebnisse der Evaluation ein Höchstmass an Glaubwürdigkeit und Akzeptanz erreichen.

*Informationsrahmen und -auswahl (N.3.):* Die beschafften Informationen sollten von einem Umfang und einer Auswahl sein, welche die Behandlung sachdienlicher Fragen zum Evaluandum ermöglichen und gleichzeitig die Interessen der Beteiligten und Betroffenen berücksichtigen.

*Klarheit des Berichts (N.4.):* Der Evaluationsbericht sollte das Evaluandum und seinen Kontext ebenso beschreiben wie die Absichten, die Verfahren und die Ergebnisse der Evaluation, damit die wesentlichen Informationen zur Verfügung stehen und leicht verstanden werden können.

*Rechtzeitigkeit und Verbreitung des Berichts (N.5.):* Wichtige Zwischenergebnisse und Evaluationsberichte sollten den Beteiligten und Betroffenen in einer Weise zukommen, so dass sie rechtzeitig genutzt werden können.

*Wirkung der Evaluation (N.6.):* Evaluationen sollten so geplant, durchgeführt und beschrieben werden, dass die Beteiligten und Betroffenen dazu ermuntert werden, der Evaluation zu folgen, damit sich die Wahrscheinlichkeit erhöht, dass die Evaluation genutzt wird.

#### **Anwendbarkeit (A.)**

*Praktische Verfahren (A.1.):* Die Evaluationsverfahren sollten praktisch sein, so dass Störungen minimal gehalten und die benötigten Informationen beschafft werden können.

*Politische Überlebensfähigkeit (A.2.):* Evaluationen sollten mit Voraussicht auf die unterschiedlichen Positionen der verschiedenen Interessengruppen geplant und durchgeführt werden, so dass deren Kooperation erreicht werden kann und so dass mögliche Versuche irgendeiner dieser Gruppen, die Evaluationsoperationen zu beschneiden oder die Resultate zu verzerren oder zu missbrauchen, vermieden werden können.

*Kostenwirksamkeit (A.3.):* Die Evaluation sollte Informationen mit einem Wert produzieren, der die gemachten Aufwendungen rechtfertigt.

#### **Korrektheit (K.)**

*Formale Vereinbarung (K.1.):* Die formalen Parteien einer Evaluation sollten den Vereinbarungen in schriftlicher Form zustimmen, damit sie zur Einhaltung aller Bedingungen verpflichtet sind oder damit sie diese aufs Neue verhandeln müssen.

*Deklaration von Interessenkonflikten (K.2.):* Interessenkonflikte sind offen und aufrichtig zu behandeln, damit sie den Prozess und die Resultate der Evaluation nicht kompromittieren.

*Vollständige und aufrichtige Offenlegung (K.3.):* Mündliche und schriftliche Evaluationsberichte sollten in der Darlegung der sachdienlichen Ergebnisse und der Grenzen der Evaluation offen, direkt und ehrlich sein.

*Herstellung von Öffentlichkeit (K.4.):* Die formalen Parteien einer Evaluation sollten das Recht der Öffentlichkeit auf Information respektieren und sicherstellen, wobei die Begrenzungen durch andere verwandte Prinzipien und Regelungen, wie die öffentliche Sicherheit und das Recht auf Privatsphäre, berücksichtigt werden müssen.

*Schutz individueller Menschenrechte (K.5.):* Evaluationen sollten so geplant und durchgeführt werden, dass die Rechte und das Wohlergehen der Menschen respektiert und geschützt sind.

*Menschliche Interaktionen (K.6.):* Evaluatoren sollten in den Interaktionen mit anderen Personen im Zusammenhang mit der Evaluation die Würde und den Wert der Menschen respektieren, damit die Beteiligten nicht gefährdet oder geschädigt werden.

*Ausgewogene Einschätzung (K.7.):* Evaluationen sollten in der Überprüfung und in der Präsentation der Stärken und Schwächen des Evaluandums vollständig und fair sein, so dass die Stärken weiter ausgebaut und die Problemfelder benannt werden können.

*Finanzielle Verantwortlichkeit (K.8.):* Die Zuweisung und Ausgabe von Ressourcen durch den Evaluator sollte unter Berücksichtigung fundierter und verantwortlicher Verfahren und auch anderweitig klug und ethisch verantwortlich erfolgen.

#### **Genauigkeit (G.)**

*Dokumentation des Programms (G.1.):* Das zu evaluierende Programm sollte klar und genau

beschrieben und dokumentiert werden, so dass es deutlich identifiziert werden kann.

*Kontextanalyse (G.2.):* Der Kontext des Evaluandums sollte ausreichend detailliert untersucht werden, so dass mögliche Beeinflussungen des Evaluandums identifiziert werden können.

*Beschreibung von Zielen und Vorgehen (G.3.):* Die Ziele und das Vorgehen der Evaluation sollten ausreichend genau beobachtet und beschrieben werden, so dass sie identifiziert und eingeschätzt werden können.

*Verlässliche Informationsquellen (G.4.):* Die Informationsquellen sollten ausreichend detailliert beschrieben werden, so dass die Adäquatheit der Informationen eingeschätzt werden kann.

*Valide und reliable Information (a) bzw. Glaubwürdigkeit (b), Übertragbarkeit (c) und Verlässlichkeit (d) (G.5.):*

(a) Die Verfahren zur Erhebung von Informationen sollten so gewählt oder entwickelt und dann umgesetzt werden, dass die Gültigkeit und die Zuverlässigkeit der erhaltenen Informationen für den gegebenen Zweck sichergestellt sind.

(b) Die Rekonstruktionen durch die Evaluation sollten eine korrekte Abbildung der Konstruktionen der Akteure darstellen.

(c) Eine präzise, detaillierte und vollständige Beschreibung (thick description) des untersuchten Kontextes sollte es erlauben, zu entscheiden, ob die Aussagen der Evaluation auf andere Kontexte übertragbar sein könnten.

(d) Der Evaluationsprozess sollte von aussen erkennbar sein, damit auch die getroffenen Entscheidungen und Interpretationen in der Evaluation untersucht und beurteilt werden können.

*Systematische Datenüberprüfung (G.6.):* Die in einer Evaluation gesammelten, bearbeiteten und angegebenen Daten sollten überprüft und korrigiert werden, so dass die Resultate der Evaluation nicht fehlerhaft sind.

*Analyse qualitativer und quantitativer Information (G.7.):* Qualitative und quantitative Informationen einer Evaluation sollten angemessen und systematisch analysiert werden, so dass die Fragen der Evaluation effektiv beantwortet werden.

*Begründete Folgerungen (G.8.):* Die in einer Evaluation erfolgten Folgerungen sollten explizit begründet sein, so dass die Beteiligten und Betroffenen diese einschätzen können.

*Unvoreingenommene Berichterstattung (a) bzw. Bestätigbarkeit (b) (G.9.):*

(a) Die Evaluationsverfahren sollten über Vorkehrungen verfügen, um die Ergebnisse und Berichte einer Evaluation gegen Verzerrungen durch persönliche Gefühle und Vorlieben irgendeiner Partei in der Evaluation zu schützen.

(b) Die Konstruktionen der Evaluation sollten auf ihre Ursprünge und Wurzeln im empirischen Material auch von aussen rückführbar sein.

*Meta-Evaluation (G.10.):* Die Evaluation selbst sollte formativ und summativ mit den vorliegenden oder anderen Kriterien evaluiert werden, so dass die Durchführung entsprechend angeleitet werden kann und damit die Beteiligten und Betroffenen bei Abschluss der Evaluation ihre Stärken und Schwächen gründlich überprüfen können.

Der durch uns revidierte Kriterienraster umfasst nunmehr 27 Einzelkriterien, wovon sechs Kriterien auf die Gruppe der «Nützlichkeit», drei auf die Gruppe der «Anwendbarkeit», acht auf jene zur «Korrektheit» und zehn auf die Gruppe der «Genauigkeit» entfallen. Die Kriterien und

die Kriteriengruppen sind im Hinblick auf ihre Bedeutung für die Gesamtqualität einer Evaluation bewusst nicht gewichtet worden. Dass die Gruppe mit den Kriterien zur «Anwendbarkeit» nur gerade drei Kriterien umfasst, bedeutet somit keinesfalls, dass diesem Aspekt die geringste Bedeutung zukommen soll.

Das hier vorgestellte Kriterienraster eignet sich nicht nur für die (nachträgliche) Bewertung von Evaluationsstudien in einer Meta-Evaluation, sondern kann auch als Orientierungslinie für die praktische Durchführung von Evaluationen oder als Hilfsmittel bei der Aus- und Weiterbildung in Evaluationsforschung eingesetzt werden.

## Die Folgerungen für die Evaluationspraxis

Der Fallstudienvergleich hat ergeben, dass die Qualität der untersuchten Evaluationsstudien durchaus ein beachtliches Niveau erreicht. Dies gilt vor allem für die Bereiche der «Nützlichkeit», der «Anwendbarkeit» und der «Korrektheit», in denen mehrheitlich gute bis sehr gute Bewertungen resultierten. Hingegen hat die Untersuchung auch gezeigt, dass im Bereich der «Genauigkeit» noch einige Schwachpunkte festzustellen sind. Deshalb sehen wir auch primär im zuletzt genannten Bereich Ansatzpunkte zur Qualitätsverbesserung.

Aufgrund dieser Resultate ist dem Bereich der Genauigkeit bei der Durchführung von Evaluationen grösseres Gewicht beizumessen. Dies gilt sowohl für die Konzipierung einer Evaluation wie auch für die Umsetzung und die Berichterstattung. In all diesen Phasen haben sich Evaluationen

neben den praxisorientierten auch an den wissenschaftlichen Anforderungen zu orientieren. Beide Aspekte sollten angemessen berücksichtigt werden. Wird eine Seite allzu stark vernachlässigt, verliert die Evaluationsforschung als praxisorientierte Forschung ihre Daseinsberechtigung.

In mehreren der untersuchten Evaluationen bestand aufgrund unserer Untersuchung ein Missverhältnis zwischen dem Umfang und der Komplexität der zu bearbeitenden Fragestellung und den für die Evaluation zur Verfügung stehenden Ressourcen. Den Evaluatoren gelang es trotz dieser ungünstigen Voraussetzungen mehrheitlich, den Ansprüchen der Praxis gerecht zu werden. Hingegen mussten im Bereich der Wissenschaftlichkeit zum Teil Kompromisse eingegangen werden, die zu einem Qualitätsverlust führten. Mit einem gewissen Recht haben die betroffenen Forscher geltend gemacht, gerügte Mängel ihrer Arbeiten seien darauf zurückzuführen, dass zu wenig Zeit und Geld für eine den Kriterien gerecht werdende Durchführung der Evaluation zur Verfügung gestanden hätten.

Aufgrund dieser Beobachtungen empfehlen wir potentiellen Auftraggebern, Evaluationen mit angemessenen Ressourcen auszustatten, damit nicht qualitativ unbefriedigende Studie entstehen. Die Bemessung der bereitzustellenden Mittel hat sich primär an der Zielsetzung des Projekts zu orientieren. Die von der Evaluation zu beantwortenden Fragen müssen im Bewusstsein formuliert werden, dass eine Ausdehnung des Gegenstandsbereichs – ohne entsprechende Anpassung der Ressourcen – für die Untersuchung negative Konsequenzen hat. Sind die Ressourcen

cen zu knapp bemessen, wird der Handlungsspielraum der evaluierenden Forscher zu stark eingeschränkt. Die Evaluatoren sind dazu gezwungen, bei der Umsetzung der Evaluation Kompromisse einzugehen, die der Qualität der Evaluation schaden oder gar dazu führen, dass die Evaluation gänzlich scheitert. Evaluationen, die ihre Zielsetzungen verfehlen, können Kosten verursachen, die das zugesprochene Evaluationsbudget bei weitem überschreiten (vgl. dazu auch Widmer, Thomas / Rothmayr, Christine / Serdült, Uwe (1996): Kurz und gut? Qualität und Effizienz von Kurzevaluationen. Zürich: Rüegger).

### **Die Möglichkeiten zur Qualitätsverbesserung**

Um die wissenschaftliche Qualität von Evaluationen sicherzustellen und zu verbessern, sind verschiedene Massnahmen zu ergreifen:

- Es können Evaluationsexperten als Mitglieder von Begleitgruppen in den Evaluationsprozess integriert werden. So erhält die Wissenschaftlichkeit in diesen oft einflussreichen Begleitgremien ein angemessenes Gewicht. Der zumeist am Evaluationsgegenstand orientierte Sachverstand in den Begleitgruppen wird durch evaluationsspezifisches Know-how ergänzt.
- Die öffentliche Hand soll Vorkehrungen treffen, um den Aufwand für Evaluationsprojekte massgeblich zu reduzieren und deren Datengrundlage erheblich zu verbessern. Im Vordergrund steht dabei die (weitere) Verbesserung der Datenlage durch die amtliche Statistik, aber auch die Erhebung und Archivierung von Prozessdaten im Vollzug sowie die Verankerung

des Öffentlichkeitsprinzips bei Bund, Kantonen und Gemeinden.

- Während oft ausreichende Einsichten in die technischen Wirkungsmechanismen und die ökonomischen Zusammenhänge zur Verfügung stehen, fehlt es indessen in verschiedenen Bereichen an Grundlagenwissen über soziale und politische Strukturen und Prozesse. Deshalb ist die Grundlagenforschung in diesen Bereichen verstärkt zu fördern. Dadurch lässt sich nicht nur der Aufwand für Evaluationsstudien reduzieren, sondern es wird damit gleichzeitig auch die Gefahr von Überinterpretationen und Fehleinschätzungen in Evaluationsstudien reduziert.

- Die im Evaluationsprozess anfallenden Informationen sollten laufend systematisch erfasst und aufgearbeitet werden. Dies erleichtert die methodische Selbstreflexion der Evaluierenden genauso wie die Erarbeitung einer fundierten Berichterstattung.

- Zur Qualitätssicherung in der Evaluationsforschung sind vermehrt Meta-Evaluationen durchzuführen. Besonders Institutionen, die sehr häufig Evaluationen durchführen (lassen), sollen regelmässig Meta-Evaluationen veranlassen.

- Evaluationsverträge sind routinemässig mit einer Meta-Evaluationsklausel zu versehen, die es den beteiligten Parteien gestattet, eine Meta-Evaluation in Auftrag zu geben, und die den Vertragspartner zur Mitwirkung an der Meta-Evaluation verpflichtet.

- Die Berichterstattung über die Evaluation soll verstärkt darauf ausgerichtet werden, dass die vorgenommenen Untersuchungsschritte in transparenter und

nachvollziehbarer Form dargestellt werden.

In der Schweiz existiert bisher nur eine relativ kleine Zahl von Evaluatoren, die in der Lage sind, qualitativ hochstehende Evaluationsstudien durchzuführen. Dies hängt sicher mit der Kleinheit der Schweiz zusammen. Diese Tatsache birgt zwei ernstzunehmende Gefahren: Erstens fehlt es in einzelnen Politikbereichen an Konkurrenz zwischen den Forschungsteams. Zweitens – und das fällt unseres Erachtens stärker ins Gewicht – kann dies dazu führen, dass zwischen den involvierten Verwaltungsstellen und bestimmten spezialisierten Evaluationsequipen Interessenbindungen entstehen, die sich negativ auf die Qualität und die Glaubwürdigkeit von Evaluationsstudien auswirken können. Weiter hat sich gezeigt, dass die an Evaluatoren zu stellenden Anforderungen sehr anspruchsvoll sind und nicht immer vollständig erfüllt werden können. Ein professioneller Evaluator muss sich in kurzer Zeit in sehr unterschiedlichen Untersuchungsfeldern zurechtfinden können. Er sollte eine breite Palette von Evaluationsansätzen und sozialwissenschaftlichen Methoden beherrschen. Auch die kommunikativen Tätigkeiten und das Projektmanagement stellen hohe Anforderungen.

Deshalb ist der Kreis der Evaluatoren, die fähig sind, qualitativ hochstehende Evaluationen durchzuführen, zu erweitern, und die Kenntnisse und Fertigkeiten der aktiven Evaluatoren sind weiter zu vertiefen. Dazu dienen die folgenden Massnahmen:

- Die Einrichtung interdisziplinär ausgerichteter, evaluationsspezifischer Lehrgänge in der Aus- und Weiterbildung kann

einen wichtigen Beitrag zur Verbreiterung und Vertiefung der Wissensbasis leisten.

- Die vor kurzem gegründete Schweizerische Evaluationsgesellschaft (SEVAL) kann einen wichtigen Beitrag zur Weiterbildung leisten und ein bedeutendes Forum des Gedankenaustausches für die an Evaluationen Interessierten darstellen.

Im Rahmen der Meta-Evaluation stiessen wir auf zwei weitere Aspekte, die zu beachten sind. Erstens kamen wir aufgrund der empirischen Analyse zum Schluss, dass zwei Studien nicht als Evaluationen bezeichnet werden sollten. Der Begriff «Evaluation» sollte nicht inflationär für verschiedenste Typen von Auftragsverhältnissen (Beratung, Organisationsentwicklung, Begleitforschung etc.) verwendet werden. Nicht alles, was ein Evaluator tut, ist Evaluation. Auch wenn der Begriff der Evaluation heute populär ist, sollten nur Studien mit einer evaluativen Zielsetzung und einem entsprechenden Instrumentarium auch als Evaluationen bezeichnet werden.

Zweitens fiel auf, dass die vertraglichen Abmachungen zwischen Auftraggeber und Evaluator bei den untersuchten Evaluationen oft nicht ausreichend waren. Der Evaluationsvertrag sollte eine unumstrittene, eindeutige und ausreichend präzise Basis für die Evaluation bilden. Es ist zu vermeiden, dass Evaluationsverträge zu pauschal ausfallen, dass sie Widersprüche aufweisen oder dass sie Regelungen enthalten, die nicht von allen Beteiligten getragen werden. Insbesondere ist davor zu warnen, mehrere Offerten oder ganze Briefwechsel als bindend zu erklären und als integralen Bestandteil des Evaluationsvertrages zu bezeichnen.

Ebenfalls zu meiden sind Vereinbarungen, die nur rudimentäre Angaben über die zu leistenden Arbeiten enthalten. Eine präzise Festlegung der Inhalte und Methoden einer Evaluation dient Auftraggebern und Auftragnehmern gleichermaßen. Da die korrekte und präzise Erarbeitung von Offerten und Vertragsentwürfen zum Teil schon einen recht hohen Aufwand verursacht, sollte bei Bedarf eine separate Entschädigung für diese Arbeiten geprüft werden. Dies würde es potentiellen Auftragnehmern erlauben, mehr Ressourcen für die Auftragsvorbereitung einzusetzen.

Insgesamt betrachtet hat die Meta-Evaluation aber klar gezeigt, dass in der schweizerischen Evaluationsforschung in den letzten Jahren gute Arbeit geleistet wurde. Die durch die AGEVAL und das NFP 27 ausgelösten Aktivitäten haben Früchte getragen und zu einer hohen Qualität der schweizerischen Evaluationsforschung beigetragen. Die vorgängig genannten Massnahmen verfolgen die Absicht, diese Entwicklung zu unterstützen. Die Sicherstellung einer hohen Qualität soll dazu beitragen, dass Evaluationen ihrer bedeutenden Aufgabe an der Nahtstelle zwischen Wissenschaft und Praxis auch in Zukunft gerecht werden können.

#### Adresse des Autors

Institut für Politikwissenschaft  
Universität Zürich, Karl Schmid-Strasse 4  
8006 Zürich  
Telefon 01 257 38 41, Fax 01 252 54 11  
E-Mail thow@pwi.unizh.ch

#### Der Schlussbericht wurde wie folgt veröffentlicht

##### Hauptbericht

Thomas Widmer: Meta-Evaluation: Kriterien zur Bewertung von Evaluationen. Bern: Paul Haupt 1996.

##### Fallstudien

Thomas Widmer: Fallstudien zur Meta-Evaluation, Teil I: «Umweltpolitik und technische Entwicklung» (Studie Nr. 1), Reihe Schlussberichte NFP 27, Bern 1996.

Thomas Widmer: Fallstudien zur Meta-Evaluation, Teil II: «Unternehmerische Innovationsprozesse» (Studie Nr. 2), Reihe Schlussberichte NFP 27, Bern 1996.

Thomas Widmer: Fallstudien zur Meta-Evaluation, Teil III: «Wohneigentumsförderung durch den Bund» (Studie Nr. 4), Reihe Schlussberichte NFP 27, Bern 1996.

Thomas Widmer: Fallstudien zur Meta-Evaluation, Teil IV: «Evaluation der Zürcher Gemeinschaftszentren» (Studie Nr. 5), Reihe Schlussberichte NFP 27, Bern 1996.

Thomas Widmer: Fallstudien zur Meta-Evaluation, Teil V: «Evaluation Kehrtrichtersackgebühre» (Studie Nr. 7), Reihe Schlussberichte NFP 27, Bern 1996.

Thomas Widmer: Fallstudien zur Meta-Evaluation, Teil VI: «Heizkostenabrechnung und erneuerbare Energien» (Studie Nr. 8), Reihe Schlussberichte NFP 27, Bern 1996.

Thomas Widmer: Fallstudien zur Meta-Evaluation, Teil VII: «Evaluierung der EWR-Informationskampagne» (Studie Nr. 10), Reihe Schlussberichte NFP 27, Bern 1996.

Der Hauptbericht ist im Buchhandel erhältlich. Die Fallstudien können zu einem Preis von je 5.- Franken bezogen werden über: Schweiz. Nationalfonds, Abt. IV/NFP 27, Postfach 2338, 3001 Bern, Tel. 031 308 23 49.

#### Veröffentlichungen aus dem Nationalen Forschungsprogramm «Wirksamkeit staatlicher Massnahmen»

**Beim Schweizerischen Nationalfonds (Schweizerischer Nationalfonds, Abt. IV/NFP27, Postfach 2338, 3001 Bern, Tel. 031 308 23 49) erhältliche Veröffentlichungen:**

##### Schlussberichte

Rey Michel, et al., La politique valaisanne de promotion économique: une évaluation privilégiant le point de vue des acteurs. Berne 1993, 210 p., 10 francs

Gerheuser Frohmüt W., Schmid André, Bedingt wirksam: Das Lohngleichheitspostulat des Heimarbeitsgesetzes. Bern 1993, 154 Seiten, Fr. 10.–

Evaluationen in der Agrarpolitik I. Bern 1994, 192 Seiten, Fr. 20.–  
Stefan Felder, Peter Rieder, Analyse des Angebotsverhaltens und der Produktionsstruktur im Agrarsektor mit Hilfe eines LP Strukturmodells

Evaluationen in der Agrarpolitik II. Bern 1994, 154 Seiten, Fr. 10.–  
Priska Baur, Leo Meyer, Peter Rieder; Politökonomische Analyse der Getreidepolitik  
Stefan Felder; Leo Meyer, Peter Rieder; Vorzeitige Stilllegung von Höchsttierbeständen in der Fleisch- und Eierproduktion 1980-1984

Gion Cavegn, Leo Meyer, Peter Rieder; Analyse des Angebotsverhaltens der Getreideproduzenten

Stalder Peter, Frick Andres, Gaillard Serge, Salzgeber Renate, Die Auswirkungen der Ausländerpolitik auf den Strukturwandel der Schweiz. Bern 1994, 97 Seiten, Fr. 20.–

Lüscher Claude, Wyss Christian, Auswirkungen des UVP-Verfahrens in einem Schweizer Kanton. Bern 1994, 85 Seiten, Fr. 20.–

Blanc Jean-Marc, Boyer François, Piotet Daniel. Evaluation comparative du processus d'apprentissage de la gestion de crise au sein de l'administration suite à la catastrophe de Tchernobyl. Berne 1995, 119 pages, 20 francs

Urio Paolo, Bari Jean-Paul et Glauser Christoph. L'évaluation comme instrument d'apprentissage et de changement organisationnels: l'exemple des programmes cantonaux de promotion de santé. Berne 1996, 203 pages, 20 francs

Widmer Thomas, Fallstudien zur Metaevaluation. 7 Teile. Bern 1996, 52 bis 78 Seiten, je Fr. 5.–

##### Kurzberichte

Rey Michel, Vodoz Luc, La promotion économique valaisanne. Berne 1993

Schulz Hans-Rudolf, Muggli Christoph, Hübschle Jörg, Die Wirksamkeit der staatlichen Wohneigentumsförderung. Bern 1993

Gerheuser Frohmüt W., Schmid André, Bedingt wirksam: Das Lohngleichheitspostulat des Heimarbeitsgesetzes. Bern 1993

Bierter Willy, Binder Hans-Martin, Unternehmerische Innovationsprozesse, Die Wirksamkeit staatlicher Förderungsmassnahmen-eine Evaluation. Bern 1994

Balthasar Andreas, Knöpfel Carlo, Umweltvorschriften und technische Entwicklung. Bern 1994

Meyer Leo, Rieder Peter,  
Evaluationen in der Agrarpolitik - Fallbeispiele im Getreidebau und der Tierhaltung. Bern 1994

Hanser Christian, Kuster Jürg, Cavelti Guido,  
Auswirkungen der staatlichen Hotellerieförderung. Bern 1994

Stalder Peter, Frick Andres, Gaillard Serge, Salzgeber Renate,  
Die Auswirkungen der Ausländerpolitik auf den Strukturwandel der Schweiz. Bern 1994

Lüscher Claude, Wyss Christian, Geringere Umweltbelastung dank  
Umweltverträglichkeitsprüfung?. Bern 1994

Delley Jean-Daniel. Impact et limites des mesures de protection de l'air:  
Le cas de l'industrie et de l'artisanat. Berne 1994

Boyer François, Piotet Daniel. Leçons tirées de Tchernobyl par les administrations suisse  
et allemande. Berne 1995

«Avenir du tourisme de ski», Communauté de travail (éd.). Tourisme de ski: de l'analyse  
à la stratégie. Berne 1995

Widmer Thomas, Evaluationen und ihre Qualität. Bern 1996

Widmer Thomas, Kurzevaluationen: Eine Alternative zu umfangreicheren Evaluationsstudien?.  
Bern 1996

Bussmann Werner, Mehr Transparenz über die Wirkungen staatlichen Handelns: Ergebnisse  
aus dem Nationalen Forschungsprogramm «Wirksamkeit staatlicher Massnahmen» (NFP 27).  
Bern 1996

Bussmann Werner, Plus de transparence sur les effets de l'action étatique: Résultats du  
programme national de recherche «Les effets des mesures étatiques» (PNR 27). Berne 1996

### **Diskussionsbeiträge**

Evaluationen: Wozu und für wen? Tagung vom 14. August 1991 in Bern, mit Beiträgen  
von Eleanor Chelimsky, Patrick Viveret, Dieter Freiburghaus und Werner Bussmann, 52 Seiten,  
gratis, lieferbar solange Vorrat

### **Im Buchhandel erhältliche Veröffentlichungen**

#### **Synthese**

Bussmann Werner, Evaluationen staatlicher Massnahmen erfolgreich begleiten und nutzen:  
ein Leitfaden. Rüegger, Chur/Zürich 1996 (2. Auflage)

Bussmann Werner, Accompagner et mettre à profit avec succès les évaluations des  
mesures étatiques, guide de réflexion. Edition Georg, Genève 1995

Bussmann Werner, Klöti Ulrich und Knoepfel Peter (Hrsg.). Einführung in die Evaluations-  
forschung (erscheint ab 1997, à partir de 1997)

Verwaltung & Management, Mai/Juni 1996: S. 129 -141 (Beiträge von  
Bundesrat A. Koller, E. Buschor, W. Bussmann, U. Klöti, Thomas Widmer und L. Mader)

Bussmann Werner (Hrsg.) Lernen in staatlichen Verwaltungen und Policy-Netzwerken.  
Rüegger, Chur/Zürich 1994

Morand Charles-Albert (ed.), Evaluation législative et lois expérimentales. Presses universitaires  
d'Aix-Marseille, Aix-en-Provence 1993 (commandes par fax au numéro 0033 42 17 2903)

### **Projekte**

Sangra Emmanuel, Evaluation des actions d'information de la protection civile sur le compor-  
tement de la population au cas d'alarme, Editions J-M Blanc SA, Le Mont-sur-Lausanne 1992

Schulz Hans-Rudolf, Muggli Christoph, Hübschle Jörg, Die Wirksamkeit der staatlichen Wohn-  
eigentumsförderung, Schriftenreihe Wohnungswesen (Hrsg. Bundesamt für Wohnungswesen),  
Band 55. Bern 1993, Bezug bei der EDMZ

Bierter Willy, Binder Hans-Martin, Unternehmerische Innovationsprozesse, Die Wirksamkeit  
staatlicher Förderungsmassnahmen-eine Evaluation, Rüegger, Chur/Zürich 1994

Balthasar Andreas, Knöpfel Carlo, Umweltschutz und technische Entwicklung,  
Helbing & Lichtenhahn. Basel 1994

Hanser Christian, Kuster Jürg, Cavelti Guido, Hotellerieförderung durch Bund und Kantone,  
Evaluation der Auswirkungen in der Hotellerie, Schriftenreihe BIGA, Beiträge zur Tourismus-  
politik Nr. 3. Bern 1994, Bezug bei der EDMZ

Delley Jean-Daniel, Impact et limites des mesures de protection de l'air:  
Le cas des installations industrielles et artisanales.  
CETEL (Faculté de droit, Uni Mail, CH-1211 Genève 4), Genève 1994

Brandner, Balthasar et al., Skitourismus: Von der Vergangenheit zum Potential der Zukunft.  
Rüegger, Chur/Zürich 1995

Widmer Thomas 1996. Meta-Evaluation. Kriterien zur Bewertung von Evaluationen.  
Haupt, Bern 1996

Widmer Thomas, Rothmayr Christine, Serdült Uwe, Kurz und gut?  
Qualität und Effizienz von Kurzevaluationen. Zürcher Beiträge zur Politikwissenschaft, Band 19.  
Rüegger, Zürich 1996 (im Erscheinen).

Sauvain, Paul et al., Zukunft des Skitourismus / Avenir du tourisme de ski. Schriftenreihe  
BIGA/OFIAMT: Beiträge zur Tourismuspolitik Nr. 7 (Bezugsquelle EDMZ), Bern 1995

Hohermuth Susanne, Meier Hans-Peter, Lokale Chancen für Nachhaltigkeit: Wintertourismus  
aus der Sicht der Bevölkerung. Edition cultur prospektiv (Bezugsquelle: cultur prospektiv,  
Gasometerstrasse 28, 8005 Zürich), Zürich 1996

Kissling-Näf Ingrid, Lernprozesse und Umweltverträglichkeitsprüfung - Staatliche  
Steuerung über Verfahren und Netzwerkbildung in der Abfallpolitik. Dissertation Universität  
St. Gallen Nr. 1744, St. Gallen 1996

### **Kurzevaluationen im Rahmen des NFP 27**

ECOPLAN, Gestaltungsmöglichkeiten und Auswirkungen verschiedener Submissionsordnungen,  
Evaluation im Auftrag der Bau-, Verkehrs- und Energiedirektion des Kantons Bern, 1993,  
200 Seiten, Fr. 30.- (erhältlich bei der BVE, Reiterstrasse 11, 3011 Bern, Tel. 031 633 31 63)

ECOPLAN, Schneeräumung von Urner Passstrassen, Evaluation im Auftrag der Baudirektion  
des Kantons Uri, 1993 (erhältlich beim Bauamt des Kantons Uri, 6460 Altdorf, Tel. 044 4 26 11)

Forster Sophie/Jeanrenaud Claude, La promotion économique neuchâteloise, bilan et  
perspectives, Neuchâtel (peut être obtenu auprès de l'Institut de recherches économiques  
et régionales, tél. 038 21 10 85)

Institut für Politikwissenschaft



UM866427



Ziel des NFP 27 ist die Erprobung und Verbesserung von Methoden zur Erfassung und Beurteilung der Wirksamkeit staatlicher Massnahmen. Es werden Evaluationen, Kurzevaluationen und Studien über Lernprozesse in Verwaltungen durchgeführt, um

- ein Know How über Evaluationsmethoden aufzubauen
- Evaluationsmethoden zu erproben und daraus Erfahrungen zu gewinnen
- Anhaltspunkte für die Zeitgerechtigkeit, Zuverlässigkeit und Relevanz von Evaluationen zu gewinnen
- zu überprüfen, inwieweit durch Evaluationen Lernprozesse in Gang gesetzt werden können.

Thomas Widmer

# Evaluationen und ihre Qualität

Meta-Evaluation als Beitrag  
zur Qualitätssicherung  
in der Evaluationsforschung

IPZ

NFP

27:27

E f f e t s  
des mesures étatiques  
Wirksamkeit  
staatlicher Massnahmen  
E f f e t t i  
delle misure statali

E f f e t s  
des mesures étatiques  
Wirksamkeit  
staatlicher Massnahmen  
E f f e t t i  
delle misure statali

E f f e t s  
des mesures étatiques  
Wirksamkeit  
staatlicher Massnahmen  
E f f e t t i  
delle misure statali