

# Verwendung von Bewertungskriterien in den externen Evaluationen der Bundesverwaltung: Ergebnisse einer Forschungsstudie

**Christine Heuer** | *Evaluationen von Massnahmen des Bundes sollen transparent und von guter Qualität sein. Ein zentrales Kennzeichen von Evaluationen ist die Bewertung, die anhand eingangs festgelegter und offengelegter Kriterien erfolgen soll. Der vorliegende Artikel präsentiert die wichtigsten Ergebnisse einer Forschungsstudie, die der Frage nach der Offenlegung und der Verwendung der Bewertungskriterien in den extern in Auftrag gegebenen Evaluationen der Bundesverwaltung nachging. Untersucht wurden 29 Evaluationsberichte anhand einer Dokumentenanalyse. Die Autorin kommt zum Schluss, dass in den untersuchten Berichten den Bewertungskriterien sehr unterschiedlich und der Tendenz nach eher wenig Aufmerksamkeit geschenkt wird.*

## Inhaltsübersicht

- 1 Einleitung
- 2 Theoretische Einführung ins Thema
- 3 Forschungsfragen
- 4 Methodisches Vorgehen
- 5 Ergebnisse und Beantwortung der Forschungsfragen
- 6 Diskussion der Ergebnisse
- 7 Schlussfolgerungen

## 1 Einleitung

Im Rahmen der Totalrevision der Bundesverfassung der Schweizerischen Eidgenossenschaft im Jahr 1999 wurde Artikel 170 neu eingeführt. Dieser hält fest, dass die Bundesversammlung dafür zu sorgen hat, dass die Massnahmen des Bundes auf ihre Wirksamkeit überprüft werden. Als Folge der Einführung dieses Artikels setzte die Generalsekretärenkonferenz die «Interdepartementale Kontaktgruppe Wirkungsprüfungen» (IDEKOWI) ein und erteilte ihr den Auftrag, der Exekutive Empfehlungen zur Umsetzung von Artikel 170 zu unterbreiten (IDEKOWI 2004, 1). Der Bundesrat nahm den Schlussbericht der IDEKOWI am 3. November 2004 zur Kenntnis und beauftragte die Ämter, die Departemente und die Bundeskanzlei, die von der IDEKOWI vorgeschlagenen Massnahmen umzusetzen. Mit diesen Massnahmen sollte die Wirkungsorientierung in der Bundesverwaltung verstärkt, die Transparenz entsprechender Überprüfungen und deren Qualität verbessert sowie die Wirtschaftlichkeit stärker gewichtet werden (IDEKOWI 2004, 3).

Der Begriff «Wirksamkeit» ist in der Bundesverfassung weit gefasst. Die Überprüfung ist nicht auf die Wirksamkeit im engeren Sinn begrenzt (Zielerreichungsgrad einer Massnahme), sondern bezieht sich auch auf den Vollzug, die Wirkungen

sowie die Wirtschaftlichkeit. Daher empfahl die IDEKOWI, ihn auf Gesetzes- und Verordnungsebene mit den drei Begriffen «Wirksamkeit», «Zweckmässigkeit» und «Wirtschaftlichkeit» wiederzugeben (IDEKOWI 2004, 45). Damit sind der Bundesverwaltung drei zentrale Bewertungskriterien ihrer Massnahmen vorgegeben.

In professionell durchgeführten Evaluationen wird die Bewertung anhand explizit auf den Untersuchungsgegenstand ausgerichteter, präzise festgelegter und offengelegter Kriterien vorgenommen (Stockmann 2007, 27). Dies kommt auch im Standard C2 «Nachvollziehbare Bewertung und begründete Folgerungen» der Schweizerischen Evaluationsgesellschaft (SEVAL) zum Ausdruck: «Die Bewertung eines Evaluationsgegenstandes erfolgt auf systematische Weise. Die Kriterien und die empirischen Grundlagen der Bewertung und der Bewertungsvorgang werden nachvollziehbar dargelegt. Folgerungen werden aus den verfügbaren Daten hergeleitet und begründet.» (SEVAL 2016, 6)

Trotz seiner Relevanz ist der Bewertungsvorgang, zu dem die Bewertungskriterien gehören, gemäss Hense / Widmer (2013) bisher kaum Gegenstand der Evaluationsforschung geworden. Der Vorgang des (Be-)Wertens, das «valuing», habe bisher speziell in der empirischen Forschung kaum Aufmerksamkeit erhalten. Hense / Widmer vermuten, dass dies einerseits an der mangelnden Transparenz bei vielen Evaluationen und andererseits an den Schwierigkeiten der Sozialforschung im Umgang mit Bewertungsvorgängen liegt (2013, 270).

Die Autorin des vorliegenden Artikels ist Evaluationsmanagerin eines Bundesamtes. Eine ihrer wichtigsten Aufgaben im Rahmen des Evaluationsmanagements ist es, zuhanden der Evaluatorinnen und Evaluatoren eine systematische Bewertung des Entwurfs des Evaluationsberichts, eine sogenannte Meta-Evaluation, vorzunehmen. Bei dieser Arbeit fiel ihr immer wieder auf, dass mit dem Vorgang des Bewertens, dem zentralen Wesensmerkmal einer Evaluation, sehr unterschiedlich umgegangen wird. Da die Bundesverwaltung aufgrund ihres Verfassungsauftrags und ihrer gesetzlichen Grundlagen dem Anspruch der Transparenz und der Qualität besonders entsprechen sollte, entschied sie, die Verwendung von Bewertungskriterien in den extern in Auftrag gegebenen Evaluationen der Bundesverwaltung eingehend zu untersuchen. Der vorliegende Artikel präsentiert Teile der unveröffentlichten Abschlussarbeit zum Master of Advanced Studies (MAS) in Evaluation der Universität Bern «Bewertungskriterien in der Evaluationspraxis der Bundesverwaltung. Eine Forschungsstudie anhand von Fallbeispielen» aus dem Jahre 2014.» Für den vorliegenden Artikel ergänzte die Autorin die Teile der umfassenderen Arbeit mit weiterführenden Untersuchungen der ausgewählten Evaluationen.

## **2 Theoretische Einführung ins Thema**

### **2.1 Begriff «Bewertungskriterium»**

Der Aspekt des Bewertens ist ein allgemeines und zentrales Kennzeichen von Evaluationen (Stockmann 2010, 64; Wottawa/Thierau 2003, 14). Im Gegensatz zu Alltagsevaluationen sind bei professionell durchgeführten Evaluationen für die Bewertung Kriterien notwendig, die explizit auf den Evaluationsgegenstand ausgerichtet, präzise festgelegt und offengelegt werden (Stockmann 2007, 27). Die meisten Evaluatorinnen und Evaluatoren benutzen den Begriff Bewertungskriterium gemäss Stake (2004, 5, 7) als «an important descriptor or attribute». Scriven umfasst den Begriff als Cluster von Eigenschaften, die für einen Gegenstand charakteristisch und stets relevant sind (1959, 861–862). Er führt dazu aus: «... the term [...] include[s] indicators of success or merit, variables that are not part of success itself (or definitionally connected with it), but rather tied to it by empirical research» (1991, 111). Auf ihn Bezug nehmend definieren Beywl und Niestroj Bewertungskriterium als «Gesichtspunkt, auf welchen bei der Bewertung eines Evaluationsgegenstands explizit Bezug genommen wird» (2009, 66). Solche Gesichtspunkte können zum Beispiel auf sozialen Werten oder fachlichen Qualitätsansprüchen beruhen (Scriven 1959, 858).

Im vorliegenden Artikel wird der Begriff Bewertungskriterium gemäss dieser Definition von Beywl und Niestroj verwendet.

### **2.2 Vorgang des (Be-)Wertens in der Evaluationsliteratur**

Bewertungskriterien und ihre Verwendung sind nur ein Teil des Bewertungsprozesses, des sogenannten «valuing». Dieser wird in der Evaluationsliteratur kontrovers diskutiert. Ausgangspunkt ist die wissenschaftstheoretische Frage: Wie können Deskription (Ergebnisse empirischer Forschung) und Bewertung (Evaluation) in Einklang gebracht werden? Das heisst, wie können empirische Daten einen quasi-normativen Charakter erhalten (Scriven 2012; House/Howe 1999; Kromrey 2007)? Um dieses Problem zu lösen, müssen Strategien gefunden werden (Kromrey 2007, 114). Diese Strategien zeigen sich in den unterschiedlichen Ansätzen der Evaluationsliteratur. Zentrale Fragen, die dabei behandelt werden, sind gemäss Alkin et al. (2012) und Julnes (2012): Wer legt die Bewertungskriterien fest? Welche Rolle haben dabei die Evaluatorinnen und Evaluatoren? Wie werden die Bewertungskriterien festgelegt und welches ist ihre Quelle? Wie werden die empirischen Ergebnisse analysiert, synthetisiert und wie wird auf dieser Basis bewertet? Wer bewertet? Entgegen dieser kontroversen Diskussion ist man sich in der Evaluationsliteratur weitgehend einig, dass die Bewertungskriterien vor Beginn der Evaluation festgelegt werden sollen (Alkin et al. 2012, 36; Stake 2004, 65).

Fokussiert man die Fragen auf die Bewertungskriterien, so steht die Klärung, wer die Bewertungskriterien wie festlegt und wer sie verwendet, das heisst wer bewertet, im Zentrum.

### **3 Forschungsfragen**

Bei der Untersuchung der ausgewählten Evaluationsberichte der Bundesverwaltung zeigte sich rasch, dass nur wenige Berichte darüber Auskunft geben, wer die Bewertungskriterien festlegt, wie sie festgelegt werden und welches ihre Quelle ist, und dass diejenigen Berichte, die sich dazu äussern, diese Fragen nur fragmentarisch behandeln. Auch die Frage, wer bewertet, konnte nur auf der Basis von Vermutungen beantwortet werden. Um auf diese Fragen fundierte Antworten zu finden, wäre eine vertiefte Analyse weniger Evaluationen mittels Interviews notwendig gewesen.

Da die Autorin allfällig vorhandene Tendenzen erfassen wollte, entschied sie, statt einer vertieften Analyse der oben erwähnten Fragen eine grössere Anzahl Evaluationen zu untersuchen.

Der vorliegende Artikel präsentiert die Ergebnisse von Forschungsfragen, die sich einerseits aus der Praxis der Autorin als Evaluationsmanagerin stellen und andererseits anhand der Analyse von Evaluationsberichten der Bundesverwaltung Antworten erlauben:

#### *Offenlegung der Bewertungskriterien*

1. Werden die Bewertungskriterien in den Evaluationsberichten der Bundesverwaltung explizit als Bewertungskriterien offengelegt? Wenn nein: In welcher Art und Weise werden sie in den Evaluationsberichten offengelegt?
2. Welche Bewertungskriterien werden in den Evaluationsberichten offengelegt?

#### *Verwendung der Bewertungskriterien*

3. Wird der Evaluationsgegenstand anhand der offengelegten Bewertungskriterien bewertet?

Die erste und die dritte Forschungsfrage stammen aus der erwähnten MAS-Arbeit. Sie wurden beide für den vorliegenden Artikel mit zusätzlich erhobenen Daten und deren Auswertung ergänzt. Die zweite Forschungsfrage kam neu hinzu.

## **4 Methodisches Vorgehen**

### **4.1 Stichprobe und Untersuchungseinheit**

Um einen möglichst breiten Einblick in die externen Evaluationen der Bundesverwaltung zu erhalten, sollten möglichst viele verschiedene Bundesämter oder

-einheiten einbezogen werden. Eine Auswahl der wichtigsten Evaluationen des Bundes bietet der Bundesrat mit seinem Geschäftsbericht, mit dem er Rechenschaft ablegt über seine jährliche Zielerreichung. Im Anhang des Geschäftsberichts werden die wichtigsten Wirksamkeitsüberprüfungen der Bundesverwaltung aufgeführt (Evaluationen, Wirtschaftlichkeitsstudien, Controlling, Monitoring u. a. m.). Die beiden jüngsten Geschäftsberichte des Bundesrates, die zur Zeit der Planung der Forschungsstudie vorlagen, erfassen die Jahre 2012 und 2011. Diese beiden Berichte führen insgesamt 39 Wirksamkeitsüberprüfungen auf (2011: 23; 2012: 16). Bei 31 dieser Wirksamkeitsüberprüfungen handelt es sich um Evaluationen. Da in zwei Evaluationsberichten keine Fragestellungen aufgeführt sind und diese auch nicht aus den Zielen der Evaluation ableitbar sind, konnten lediglich 29 Evaluationsberichte analysiert werden. Diese stammen aus insgesamt 15 verschiedenen Ämtern und einer politischen Direktion respektive aus insgesamt 5 verschiedenen Departementen:

*Tabelle 1: Untersuchte Evaluationsberichte nach Departementen und Ämtern (n=29)*

Eidgenössisches Departement	Bundesamt	Anzahl untersuchter Berichte
Eidgenössisches Departement für äussere Angelegenheiten (EDA)	Direktion für Entwicklungszusammenarbeit (DEZA)	3
	Politische Direktion EDA	1
Eidgenössisches Departement des Innern (EDI)	Bundesamt für Gesundheit (BAG)	2*
	Bundesamt für Sozialversicherungen (BSV)	2
Eidgenössisches Justiz- und Polizeidepartement (EJPD)	Bundesamt für Justiz (BJ)	3
Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation (UVEK)	Bundesamt für Energie (BFE)	2
	Bundesamt für Strassen (ASTRA)	1
	Bundesamt für Verkehr (BAV)	1
	Bundesamt für Umwelt (BAFU)	1
	Bundesamt für Raumentwicklung (ARE)	1*
Eidgenössisches Volkswirtschaftsdepartement (EVD) (seit dem 1.1.2013)	Staatssekretariat für Wirtschaft (SECO)	5**
	Staatssekretariat für Bildung und Forschung (SBF)***	1
Eidgenössisches Departement für Wirtschaft, Bildung und Forschung [WBF]	Bundesamt für Berufsbildung und Technologie (BBT)***	1
	Bundesamt für Landwirtschaft (BLW)	1
	Bundesamt für Veterinärwesen (BVET)****	5
	Bundesamt für Wohnungswesen (BWW)	1
<b>Total</b>		<b>29</b>

\* Eine der aufgeführten Evaluationen wurde mit einem anderen Amt zusammen in Auftrag gegeben.

\*\* Zwei der aufgeführten Evaluationen wurden mit einem anderen Amt zusammen in Auftrag gegeben.

\*\*\* Das SBF und das BBT bilden seit dem 1.1.2013 zusammen das Staatssekretariat für Bildung, Forschung und Innovation (SBFI).

\*\*\*\*Das BVET ist seit dem 1.1.2014 mit der Abteilung Lebensmittelsicherheit des BAG vereint und heisst nun Bundesamt für Lebensmittelsicherheit und Veterinärwesen (BLV).

Untersucht wurden die von den Bundesämtern veröffentlichten Evaluationsberichte. Pflichtenhefte wurden dort in die Analyse einbezogen, wo sie veröffentlicht wurden.

Die Forschungsstudie erhebt mit dieser Stichprobe keinen Anspruch auf Repräsentativität für die von der Bundesverwaltung extern in Auftrag gegebenen Evaluationen. Die Ergebnisse erlauben jedoch, Tendenzen aufzuzeigen.

#### **4.2 Vorgehen bei der Datenerhebung und -auswertung**

Die 29 Evaluationsberichte wurden anhand einer Dokumentenanalyse untersucht. Zu diesem Zweck entwickelte die Autorin ein Raster, das im Folgenden kurz entlang der Forschungsfragen erläutert wird:

##### *Offenlegung der Bewertungskriterien*

Für die Beantwortung der ersten beiden Forschungsfragen 1) «Werden die Bewertungskriterien in den Evaluationsberichten der Bundesverwaltung explizit als Bewertungskriterien offengelegt? Wenn nein: In welcher Art und Weise werden sie in den Evaluationsberichten offengelegt?» und 2) «Welche Bewertungskriterien werden in den Evaluationsberichten offengelegt?» untersuchte die Autorin in einem ersten Schritt die einzelnen Berichte nach Bewertungskriterien, die explizit als Bewertungskriterien offengelegt werden, das heisst als solche bezeichnet wurden. Die gesichteten Bewertungskriterien nahm sie ins Raster auf. In einem zweiten Schritt durchsuchte sie die Berichte nach Bewertungskriterien, die zwar offengelegt werden, jedoch nicht als Bewertungskriterien benannt sind. Diese als Bewertungskriterien identifizierten «Gesichtspunkte», die die Evaluatorminnen und Evaluatoren zu Beginn des Berichts aufzuführen, nahm sie ins Raster auf. Im vorliegenden Artikel werden sie als implizit offengelegte Bewertungskriterien bezeichnet. Zudem notierte die Autorin, in welcher Art und Weise diese Bewertungskriterien offengelegt werden. Die verschiedenen Arten dieser impliziten Offenlegung kategorisierte sie.

##### *Verwendung der Bewertungskriterien*

Zur Beantwortung der Forschungsfrage 3 «Wird der Evaluationsgegenstand anhand der offengelegten Bewertungskriterien bewertet?» prüfte die Autorin, ob die explizit und die implizit offengelegten Bewertungskriterien bei der Bewertung des Evaluationsgegenstands verwendet werden oder nicht. Dabei konzentrierte sie sich auf alle evaluativen Fragestellungen. Die Autorin identifizierte in den 29 Berichten 229 solcher Fragestellungen. Beispiele dafür sind: «Ist die Konzeption geeignet, um die darin enthaltenen Zielsetzungen zu erreichen?» oder «Sind die Ziele aus der Sicht von Expertinnen und Experten richtig gesetzt?».

Da es evaluative Fragestellungen gibt, die das Bewertungskriterium nicht enthalten, wurden diese Fragestellungen in solche mit und solche ohne Nennung des Bewertungskriteriums eingeteilt. Beispiele für evaluative Fragestellungen

ohne Bewertungskriterium sind: «Wie beurteilen die urteilenden Personen die Rechte der [...] im Vergleich zu den Bestimmungen, die vor 2007 gegolten haben?» oder «Wie ist der Stand der Umsetzung der Programme zu bewerten?». Fragestellungen, die das Bewertungskriterium nicht beinhalten, wurden bei der Auswertung nicht berücksichtigt.

Zusätzlich prüfte die Autorin, ob die Gesamtbewertung des Evaluationsgegenstands anhand der explizit und implizit offengelegten Bewertungskriterien auf übergeordneter Ebene (wie Wirksamkeit, Zweckmässigkeit oder Wirtschaftlichkeit) vorgenommen wird.

Das für die Untersuchung verwendete Raster und die Kategorisierung der Daten sah somit wie folgt aus:

*Tabelle 2: Analyseraster der Forschungsstudie (für den vorliegenden Artikel angepasst)*

Bewertungskriterium	Art der Offenlegung	Verwendung des Bewertungskriteriums bei der Gesamtbewertung	Nennung des Bewertungskriteriums in der evaluativen Fragestellung	Verwendung des Bewertungskriteriums bei der Beantwortung der evaluativen Fragestellung
Protokollierung des Kriteriums	Als Bewertungskriterium Als Ziel / Zweck der Evaluation Als Konkretisierung der Fragestellung Als Teil der Methode Ausschliesslich in der Fragestellung	Ja Nein	Ja (Protokollierung des Kriteriums) Nein	Ja Nein

Das Raster wurde quantitativ ausgewertet. Die Ergebnisse bezüglich der Forschungsfrage nach der expliziten Offenlegung und der Verwendung der Bewertungskriterien werden im vorliegenden Artikel pro Bericht und im Total dargestellt. Die Ergebnisse bezüglich der Forschungsfragen nach der Art der impliziten Offenlegung der Bewertungskriterien und derjenigen, welche Bewertungskriterien in den Evaluationsberichten offengelegt werden, sind nur im Total aller Berichte aufgeführt.

## 5 Ergebnisse und Beantwortung der Forschungsfragen

### 5.1 Werden die Bewertungskriterien in den Evaluationsberichten der Bundesverwaltung explizit als Bewertungskriterien offengelegt?

In 11 der 29 untersuchten Evaluationsberichte werden Bewertungskriterien als Bewertungskriterien offengelegt. Dabei handelt es sich bei 10 Berichten ausschliesslich um Kriterien auf übergeordneter Ebene, das heisst um generische Be-

wertungskriterien, die anhand von Fragestellungen und gegebenenfalls von Detailfragestellungen konkretisiert und in den Fragen der Erhebungsinstrumente messbar gemacht werden müssen. Beispiele solcher Kriterien sind Wirksamkeit, Angemessenheit und Nachhaltigkeit. In einem Bericht werden die Bewertungskriterien erst auf der Ebene der detaillierten Fragestellungen als Bewertungskriterien offengelegt.

Anbei drei exemplarische Textauszüge aus den untersuchten Berichten, die zeigen, dass und wie Bewertungskriterien explizit als Bewertungskriterien offengelegt werden.

1: *«Ebenso sind [in der grafischen Darstellung, Anm. der Autorin] die zentralen Evaluationskriterien genannt, die für die Beurteilung entscheidend sein dürften»* (genannt sind: Angemessenheit und Effizienz im Vollzug, Effizienz des Verfahrens, Effektivität (Outcome) und Zielerreichung (Impact).

2: *«Das Kriterium der Wirksamkeit bezieht sich auf das Verhältnis zwischen den beabsichtigten und den tatsächlich eingetretenen Wirkungen.»*

3: *«Following the OECD / DAC evaluation criteria [...], this study examined the relevance, efficiency, effectiveness, impact and sustainability of [...]»*

In drei der untersuchten Berichte räumen die Evaluatorinnen und Evaluatoren den Bewertungskriterien ein eigenes Kapitel ein (1 davon im Anhang), drei führen sie bei den Fragestellungen oder beim Ziel beziehungsweise beim Zweck der Evaluation auf. In einem Bericht werden die Bewertungskriterien im Kapitel «area of enquiry» offengelegt (siehe Tabelle 3 unten).

## **5.2 Wenn die Bewertungskriterien nicht explizit als Bewertungskriterien offengelegt werden, in welcher Art und Weise werden sie implizit offengelegt?**

In 12 der 18 Evaluationsberichte, welche die Bewertungskriterien implizit offenlegen, werden diese als Ziel oder Zweck der Evaluation aufgeführt. Zwei Beispiele dafür sind:

1: *«Die Absicht der Evaluation ist es, einerseits Rechenschaft zur Relevanz, Zweckmässigkeit und Wirkung von [...] abzulegen,...»*

2: *«Ihr [Evaluation; Anm. der Autorin] Ziel besteht darin, erste Erkenntnisse zur Wirkung und die Grundlage für eine Fortsetzung der Vereinbarkeitsmassnahmen und Strukturen zu erhalten.»*

In drei Berichten werden die Bewertungskriterien als Teil der Methode offengelegt. Zum Beispiel: *«Zu prüfen sind folgende fünf Punkte [...] Auswirkungen auf [...], Zweckmässigkeit im [...]»*.

In zwei Berichten werden die Bewertungskriterien erst als Teil der evaluativen Fragestellungen offengelegt. Ein Beispiel für eine solche Fragestellung ist: *«Dabei gilt es auch abzuklären, wie es um die Transparenz der Gesuchsbeurteilung und der Darlehensverwaltung bestellt ist.»*



In einem Bericht werden die Bewertungskriterien in der übergeordneten Fragestellung offengelegt: «Offen ist insbesondere, ob der Personalverleih, wie er von [...] unterstützt wird, eine zweckmässige und wirkungsvolle Ergänzung zu den bestehenden und in den letzten Jahren stark ausgebauten [...] -Stellen ist.»

Eine Übersicht zu den Ergebnissen zeigt Tabelle 3.

Tabelle 3: Offenlegung der Bewertungskriterien im Total (n=29)

Wo werden die Bewertungskriterien explizit als Bewertungskriterien offengelegt? In welcher Art und Weise werden sie implizit offengelegt?			
Ort der expliziten Offenlegung	Anzahl Berichte	Art und Weise der impliziten Offenlegung	Anzahl Berichte
Beim Ziel / Zweck der Evaluation	3	Als Ziel / Zweck der Evaluation	12
		Als Teil der Methode	3
Bei den Fragestellungen	3	In den Fragestellungen	2
		Als Teil der übergeordneten Fragestellung	1
Bei den detaillierten Fragestellungen	1		
In einem separaten Kapitel	3		
Im Kapitel «area of enquiry»	1		
<b>Total</b>	<b>11</b>	<b>Total</b>	<b>18</b>

Es gibt Berichte, die in den Fragestellungen Bewertungskriterien enthalten, die nicht auf die offengelegten Bewertungskriterien auf übergeordneter Ebene zurückzuführen sind. Bei diesen Bewertungskriterien handelt es sich um zusätzliche. Solche hat die Autorin in 12 Berichten gefunden. Die Hälfte dieser Berichte legt die Bewertungskriterien explizit, die andere implizit offen. Der vorliegende Artikel beschränkt sich auf die inhaltliche Auswertung der explizit oder implizit offengelegten Bewertungskriterien auf übergeordneter Ebene (siehe Ziff. 5.3). 27 Berichte enthalten solche Bewertungskriterien.

### 5.3 Welche Bewertungskriterien werden in den untersuchten Evaluationen offengelegt?

Die 27 untersuchten Berichte, die Bewertungskriterien auf der übergeordneten Ebene explizit oder implizit offenlegen, führen zwischen 1 und 9 Bewertungskriterien auf. Insgesamt sind es 85.

In 22 Evaluationsberichten werden die Wirksamkeit, die (Aus-)Wirkungen beziehungsweise die Zielerreichung, die Effektivität oder der Impact als Bewertungskriterium offengelegt. In 7 Berichten werden je zwei dieser Aspekte genannt, zum Beispiel die Effektivität und die Wirksamkeit, «Effectiveness» und «Impact» oder Wirksamkeit und Zielerreichung. In 7 dieser 22 Berichte ist es das einzige aufgeführte Kriterium. In 13 Evaluationsberichten wird die Effizienz, die Wirtschaftlichkeit, die «Cost-Effectiveness» oder der «Net economic value» als Bewertungs-

kriterium offengelegt. Bei 3 Berichten handelt es sich um ökonomische Evaluationen, in 2 davon ist dies das einzige Kriterium. In 9 Berichten ist die Relevanz als Bewertungskriterium genannt, in 1 als einziges. Die Nachhaltigkeit und die Angemessenheit werden in je 5 Berichten offengelegt. In 4 Berichten wird die Kohärenz, in je 3 die Zweckmässigkeit oder die Vernetzung beziehungsweise die «Connectedness» aufgeführt; auch hier in Kombination mit anderen. In je 2 Berichten wird u. a. die Eignung oder die Umsetzungstreue offengelegt. Je 1 Bericht enthält die Sensibilisierung, das richtige Tun, die Qualität, die «Coverage», die «Dimension genre», die «Additionality», die Bewährung oder die «Justification» als Bewertungskriterium.

Die 6 Evaluationsberichte aus der Entwicklungszusammenarbeit legen ausschliesslich Kriterien des Development Assistance Committee (DAC) der OECD und teilweise als Ergänzung solche des Active Learning Network for Accountability and Performance (Alnap) offen.

Eine Übersicht über die verschiedenen Bewertungskriterien auf übergeordneter Ebene und die Anzahl der Berichte, in denen diese aufgeführt werden, zeigt Tabelle 4.

*Tabelle 4: Offengelegte Bewertungskriterien auf übergeordneter Ebene im Total (n=27)*

<b>Welche Bewertungskriterien werden explizit oder implizit offengelegt?</b>	<b>Anzahl Kriterien</b>	<b>Anzahl Berichte</b>
Wirksamkeit, (Aus-) Wirkungen, Zielerreichung, Effektivität, Impact	29	22
Effizienz, Wirtschaftlichkeit, «Cost-Effectiveness», «Net economic value»	15	13
Relevanz	9	9
Nachhaltigkeit	5	5
Angemessenheit	5	5
Kohärenz	4	4
Zweckmässigkeit	3	3
Vernetzung, «Connectedness»	3	3
Eignung	2	2
Umsetzungstreue	2	2
Sensibilisierung	1	1
Richtiges Tun	1	1
Qualität	1	1
«Coverage»	1	1
«Dimension genre»	1	1
«Additionality»	1	1
«Justification»	1	1
Bewährung	1	1
<b>Total</b>	<b>85</b>	

#### **5.4 Wird der Evaluationsgegenstand anhand der explizit oder implizit offengelegten Bewertungskriterien bewertet?**

Die Evaluatorinnen und Evaluatoren nehmen die Gesamtbewertung des Evaluationsgegenstands (in der Synthese, den Zwischenfazits, dem Summary oder dem Ergebnisteil) in insgesamt 13 der 27 Berichte anhand aller auf übergeordneter Ebene explizit oder implizit offengelegten Bewertungskriterien vor. In 10 Berichten werden diese Bewertungskriterien nur zum Teil bei der Bewertung verwendet. In vier Berichten werden keine dieser offengelegten Bewertungskriterien benutzt. In den letztgenannten beschreiben die Evaluatorinnen und Evaluatoren in einem Bericht die erreichten Ziele, statt die Zielerreichung zu bewerten. In drei Berichten wird der Evaluationsgegenstand im Fazit lediglich auf der Basis der Antworten auf die Fragestellungen oder gar auf der Basis der Antworten auf die Fragen in den Erhebungsinstrumenten bewertet.

Wertet man die Verwendung der Bewertungskriterien bei der Gesamtbewertung des Evaluationsgegenstands im Total aus, so kommt man zu folgendem Ergebnis: Die Evaluatorinnen und Evaluatoren gebrauchen von den 85 offengelegten Bewertungskriterien 53 (= 62%), und sie gebrauchen 32 (= 38%) nicht. In 18 Fällen benutzen sie andere Kriterien. In 14 Fällen fehlt die Bewertung ganz, das heisst, es existieren keine Ergebnisse zum Gegenstand der explizit oder implizit offengelegten Bewertungskriterien. Für die Fälle, in denen bei der Bewertung andere Bewertungskriterien verwendet werden, anbei zwei Beispiele:

1: Explizit oder implizit offengelegtes Bewertungskriterium und sein Gegenstand: Effizienz des Vollzugs; Bewertung: die Vollzugskosten sind hoch.

2: Explizit oder implizit offengelegtes Bewertungskriterium und sein Gegenstand: Zweckmässigkeit des Programms; Bewertung: konsistentes / stimmiges Konzept, Verzögerungen in der Umsetzung.

Betrachtet man die Bewertung des Evaluationsgegenstandes anhand der Bewertungskriterien in den Fragestellungen, so zeigt sich, dass die Evaluatorinnen und Evaluatoren lediglich in 4 Berichten alle in den Fragestellungen enthaltenen Bewertungskriterien bei der Beantwortung der evaluativen Fragestellungen verwenden. Diese Berichte weisen jedoch nur 1–5 evaluative Fragestellungen auf. In 2 Berichten werden die in den evaluativen Fragestellungen genannten Bewertungskriterien bei der Beantwortung nie benutzt. Beide Berichte führen je 7 evaluative Fragestellungen mit Bewertungskriterien auf.

Wertet man alle Antworten auf die evaluativen Fragestellungen im Total aus, so zeigt sich Folgendes: Von den 175 der insgesamt 229 als evaluativ identifizierten Fragestellungen, die ein Bewertungskriterium enthalten und von den Evalua-

torinnen und Evaluatoren als beantwortbar bezeichnet werden, verwenden die Evaluatordinnen und Evaluatoren das Bewertungskriterium in 97 Antworten (= 55 %); bei 78 Fragestellungen (= 45 %) ist dies nicht der Fall. Von diesen letztgenannten werden 24 nicht und 9 deskriptiv statt evaluativ beantwortet. In 45 Antworten benutzen die Evaluatordinnen und Evaluatoren andere Bewertungskriterien als die in den Fragestellungen enthaltenen. Zwei Beispiele für den Gebrauch von anderen Bewertungskriterien sind:

1: In Fragestellung enthaltenes Bewertungskriterium und sein Gegenstand: Eignung des Prozesses; Antwort: die Zusammenarbeit funktioniert gut, die Umsetzung ist unkompliziert.

2: In Fragestellung enthaltenes Bewertungskriterium und sein Gegenstand: richtige Prioritäten; Antwort: bewährt hat sich x, zu wenig Priorität hat y.

In 53 (= 23 %) der insgesamt 229 als evaluativ identifizierten Fragestellungen wird das Bewertungskriterium nicht genannt.

Tabelle 5 zeigt die Ergebnisse pro Bericht (siehe unten).

Richtet man den Blick auf die drei am meisten explizit oder implizit offengelegten Bewertungskriterien im Hinblick auf ihre Verwendung bei der Gesamtbeurteilung des Evaluationsgegenstands, so zeigt sich Folgendes: In 19 von 22 Berichten, welche die Wirksamkeit, die (Aus-) Wirkungen, die Zielerreichung, die Effektivität oder den Impact als Bewertungskriterium offenlegen, bewerten die Evaluatordinnen und Evaluatoren anhand dieser Kriterien. Bei der Relevanz ist dies in 8 von 9 Berichten der Fall. In den 13 Berichten, in denen die Effizienz, die Wirtschaftlichkeit, die «Cost-effectiveness» oder der «Net economic value» als Kriterien offengelegt werden, verwenden die Evaluatordinnen und Evaluatoren diese in 6 Berichten, in 1 zum Teil und in den weiteren 6 nicht.

Tabelle 5: Verwendung der offengelegten Bewertungskriterien pro Bericht

Werden die offengelegten Bewertungskriterien verwendet?						
Nr. Eval***	Bewertungskriterien auf übergeordneter Ebene			Bewertungskriterien in den Fragestellungen		
	Anzahl Verwendung	Anzahl nicht Verwendung	davon nicht beantwortet	Anzahl Verwendung	Anzahl nicht Verwendung	davon nicht beantwortet
1	0	8	8	2	10	3
2	4	5	0	8	7	4
3	1	0	0	8	4	0
4	1	0	0	0	0	0
5	1	0	0	4	<sup>d</sup> 2	1
6	1	1	1	7	<sup>ddd</sup> 8	0
7	1	0	0	2	0	0
8	0	1	0	0	0	0
9	2	1	1	3	1	0
10	5	0	0	4	<sup>d</sup> 4	0
11	1	0	0	3	1	0
12	6	0	0	4	2	2
13	0	<sup>d</sup> 1	0	1	0	0
14	2	1	1	4	1	1
15	4	0	0	5	0	0
16	1	0	0	0	4	3
17	1	0	0	0	0	0
18	2	1	0	5	3	1
19	1	0	0	2	1	1
20	4	1	1	9	2	1
21	1	1	0	2	2	0
22	---	---	---	2	3	1
23	1	0	0	1	0	0
24		<sup>d</sup> 2	1	0	7	0
25	---	---	---	4	3	3
26	4	2	1	6	1	0
27	2	1	0	6	2	0
28	3	<sup>dd</sup> 6	0	3	4	3
29	4	0	0	2	<sup>dddd</sup> 6	0
<b>Total Kriterien</b>	<b>53</b>	<b>32</b>	<b>14</b>	<b>97</b>	<b>78</b>	<b>24</b>
<b>**</b>		<b>85</b>			<b>*176</b>	
<b>Total Berichte</b>		<b>27</b>			<b>29</b>	

\* eine Fragestellung als nicht beantwortbar bezeichnet

\*\* Subtotale der übergeordneten Bewertungskriterien bzw. der evaluativen Fragestellungen

\*\*\* Die Reihenfolge der Evaluationsberichte entspricht nicht der Reihenfolge der aufgeführten Evaluationen in den Geschäftsberichten.

<sup>d</sup> Anzahl deskriptiv beantworteter Fragestellungen (<sup>d</sup>eine Fragestellung deskriptiv beantwortet, <sup>dd</sup> zwei Fragestellungen deskriptiv beantwortet etc.)

## **6 Diskussion der Ergebnisse**

### **6.1 Offenlegung der Bewertungskriterien**

Nur in gut einem Drittel (11 von 29) der untersuchten Evaluationsberichte werden die Bewertungskriterien explizit als Bewertungskriterien offengelegt: in 10 Berichten als Bewertungskriterien auf übergeordneter Ebene (zum Beispiel Wirksamkeit, Angemessenheit, Relevanz) und in 1 erst bei den detaillierten Fragestellungen der Evaluation. Von diesen 11 Berichten stammen 6 aus der Entwicklungszusammenarbeit, deren Evaluationen in der Regel anhand der DAC-Bewertungskriterien der OECD evaluiert werden. Dieses Ergebnis deutet darauf hin, dass sowohl die Bundesverwaltung als Auftraggeberin wie auch die Evaluatorinnen und Evaluatoren als Auftragnehmende der ausdrücklichen Offenlegung der Bewertungsgrundlage (ausser im Themenbereich der Entwicklungszusammenarbeit) tendenziell wenig Aufmerksamkeit schenken. Gründe dafür könnten sein, dass wenig Wissen zu den Bewertungskriterien vorhanden ist oder dass diesem Thema kaum methodische Wichtigkeit eingeräumt wird.

Wenn die Bewertungskriterien nicht als Bewertungskriterien offengelegt werden, sind sie in den untersuchten Berichten vor allem als Ziel oder Zweck der Evaluation formuliert (12 von 18). Ein Bezug zur Bewertung wird dabei nicht hergestellt, was die Transparenz der Bewertungsgrundlage der Evaluation und den Stellenwert ihrer Kriterien reduziert. Dass die Evaluatorinnen und Evaluatoren bei nicht expliziter Bezeichnung der übergeordneten Bewertungskriterien diese als Ziel oder Zweck der Evaluation aufführen, kann aus Sicht der Autorin daher kommen, dass sie die unpräzisen Formulierungen der Auftraggebenden vollständig übernehmen. Eine andere Erklärungsmöglichkeit sieht die Autorin darin, dass den Evaluatorinnen und Evaluatoren nicht bewusst ist, dass es sich um Bewertungskriterien handelt. Die Bewertungskriterien werden stattdessen als Fokus der Evaluation oder als Eingrenzung des zu untersuchenden Evaluationsgegenstandes verstanden. Da alle Berichte die meisten Bewertungskriterien spätestens in den Fragestellungen offenlegen, kann jedoch davon ausgegangen werden, dass dem zentralen Grundsatz professioneller Evaluation – der Offenlegung der Bewertungsgrundlage – mehrheitlich Genüge getan wird.

Nicht vertiefter untersucht wurde das Ergebnis, dass 12 der 29 untersuchten Berichte Bewertungskriterien in den Fragestellungen aufführen, die keinen (nachvollziehbaren) Bezug zu den offengelegten Bewertungskriterien auf übergeordneter Ebene haben. Solche zusätzlichen Bewertungskriterien finden sich unabhängig davon, ob die Bewertungskriterien in den Berichten explizit oder implizit als Bewertungskriterien bezeichnet werden. Wichtig ist dieser Befund, da er aufzeigt, dass der Ort, wo die umfassendste Bewertungsgrundlage offengelegt wird, die Fragestellungen sind.

Knapp ein Viertel der evaluativen Fragestellungen (53 von 229) enthalten kein Bewertungskriterium. Die Evaluatorinnen und Evaluatoren legen dieses erst in den Erhebungsinstrumenten offen oder gar nicht. Dies ist ein relativ hoher Anteil. Aus Sicht der Autorin hat dies verschiedene mögliche Gründe. Etwa, dass die Festlegung dieser Bewertungskriterien bei der Gesamtbewertung im Hintergrund steht oder nur ergänzend gedacht ist. Eine andere mögliche Erklärung wäre, dass weder die Auftraggeberin noch die Evaluatorinnen und Evaluatoren ein Bewertungskriterium für den Gegenstand benennen können oder wollen und die Bewertung ganz an andere Stakeholder delegieren. Wenn die Auftraggebenden die Fragestellung formuliert haben, ist ein möglicher Grund für die Nichtnennung des Bewertungskriteriums, dass sie die Festlegung dieser Kriterien den Evaluatorinnen und Evaluatoren überlassen, da diese die Evaluationsexpertinnen und -experten sind.

## **6.2 Offengelegte Bewertungskriterien**

Die in den untersuchten Berichten explizit oder implizit offengelegten Bewertungskriterien auf übergeordneter Ebene wie Wirksamkeit, Zweckmässigkeit oder Nachhaltigkeit legten vermutlich die Auftraggebenden in ihren Ausschreibungen fest. Dass das am meisten offengelegte Bewertungskriterium die Wirksamkeit beziehungsweise die Zielerreichung ist, überrascht daher nicht, entspricht diese doch dem zentralen Interesse der Bundesverwaltung. Dass im Rahmen einer einzigen Evaluation beispielsweise sowohl die Wirksamkeit wie auch die Effektivität untersucht werden, erstaunt auf den ersten Blick. Betrachtet man jedoch, was genau bewertet werden soll, also den Gegenstand der Bewertung, zeigt sich, dass es sich beispielsweise bei der Effektivität um die Zielerreichung auf Ebene Outcome handelt und bei der Wirksamkeit um die Zielerreichung auf Ebene Impact. Diese Begriffe (wie auch «Wirkung» und «Auswirkung») werden jeweils mit unterschiedlichen Inhalten gefüllt, was damit zu erklären ist, dass es in der Schweiz kein einheitliches Begriffsverständnis gibt.

Dass in 13 von 29, also fast der Hälfte der Evaluationen die Wirtschaftlichkeit oder entsprechende Bewertungskriterien verwendet werden, verblüfft zunächst. Bei genauerer Untersuchung wird ersichtlich, dass die 3 ökonomischen und 5 der 6 Evaluationen aus der Entwicklungszusammenarbeit, die hohe Anzahl erklären. Dennoch verwenden 5 weitere Evaluationen dieses Bewertungskriterium. Vermutlich wird der Wirtschaftlichkeit von politischen Interventionen mehr Beachtung geschenkt, da die Frage der Verteilung der Gelder immer wichtiger wird. Diese Begründung dürfte ebenfalls für die Verwendung des Bewertungskriteriums Relevanz in 9 Berichten zutreffen (wovon 6 aus der Entwicklungszusammenarbeit stammen). Weniger erstaunlich ist, dass oftmals die Fragestellungen

nach der Wirtschaftlichkeit nicht oder nur deskriptiv beantwortet werden. Kennzahlen zu den Kosten einer Massnahme oder deren quantitativ messbar gemachten Zielen liegen meist nicht vor. Dass die Zweckmässigkeit als übergeordnetes Bewertungskriterium eher selten Verwendung findet, kann aus Sicht der Autorin so interpretiert werden, dass dieser Begriff ausgesprochen unspezifisch ist und selten befriedigend messbar gemacht wird. Vermutlich wird die Zweckmässigkeit vor allem von Auftraggebern als Bewertungskriterium aufgeführt, die sich zu wenig darüber im Klaren sind, was genau bewertet werden soll. Sie lassen sich eventuell auch vorwiegend von den Kriterien Wirksamkeit, Zweckmässigkeit und Wirtschaftlichkeit des IDEKOWI-Berichts leiten. Dass aber auch spezifischere Bewertungskriterien wie Sensibilisierung, Vernetzung oder richtiges Tun genannt werden, zeigt, dass es Evaluationen gibt, in denen Bewertungskriterien formuliert werden, die spezifischer auf den Untersuchungsgegenstand und den Informationsbedarf der Auftraggeber ausgerichtet sind.

### **6.3 Verwendung der Bewertungskriterien**

Die Gesamtbewertung des Evaluationsgegenstands wird lediglich in knapp der Hälfte (13 von 27) der untersuchten Berichte vollkommen und in 10 zum Teil anhand der entsprechend offengelegten Bewertungskriterien vorgenommen. Auf der Ebene der evaluativen Fragestellungen wird nur bei einem guten Siebtel (4 von 29) der Berichte anhand aller in den Fragestellungen enthaltenen Kriterien bewertet. Hier muss jedoch festgehalten werden, dass die Anzahl der Bewertungskriterien auf übergeordneter Ebene und die Anzahl der evaluativen Fragestellungen pro Bericht stark variieren. Auffallend ist, dass die Evaluatorinnen und Evaluatoren trotz der niedrigen Anzahl verwendeter Bewertungskriterien in den evaluativen Fragestellungen (97 von 176) die Gesamtbewertung bei 53 von insgesamt 85 Bewertungen anhand der auf übergeordneter Ebene offengelegten Bewertungskriterien vornehmen. Eine mögliche Erklärung dafür könnte sein, dass sie für die Gesamtbewertung ergänzend Ergebnisse aus deskriptiven oder kausalen Fragestellungen verwendeten, da die entsprechenden Bewertungskriterien zusätzlich zu den evaluativen mit deskriptiven und kausalen Fragestellungen konkretisiert wurden.

Die Ergebnismerte für die Verwendung der explizit oder implizit offengelegten Bewertungskriterien bei der Gesamtbewertung, wie auch für die Verwendung der in den evaluativen Fragestellungen enthaltenen Bewertungskriterien sind tief in Anbetracht der Tatsache, dass das Bewerten ein zentrales Kennzeichen von Evaluation ist. Die Autorin geht davon aus, dass auch dieses Resultat ein Hinweis darauf sein kann, dass etliche Evaluatorinnen und Evaluatoren über zu wenig fachspezifisches Knowhow verfügen und diesem zentralen Aspekt methodisch



zu wenig Beachtung schenken. Eine mögliche Erklärung für die mangelnde methodische Beachtung wäre hier, dass die Gesamtbewertung beispielsweise anhand der Kriterien der evaluativen Fragestellungen vorgenommen wird und die evaluativen Fragestellungen anhand ihrer Indikatoren aus den Erhebungsinstrumenten beantwortet werden. Dies würde bedeuten, dass die Ebenen der Fragestellung (übergeordnete Fragestellung, Fragestellung und Frage im Erhebungsinstrument) und deren Beantwortung in Bezug auf das Bewertungskriterium nicht identisch sind. Da die Messbarmachung der Bewertungskriterien auf übergeordneter Ebene nicht transparent gemacht wird, ist die Bewertung nicht nachvollziehbar. Ein weiterer Grund für die Nichtverwendung der Bewertungskriterien kann sein, dass die Bewertungskriterien während der Durchführung der Evaluation angepasst wurden. Dieser Umstand sollte jedoch im Evaluationsbericht als solcher ausgewiesen werden, was in keinem der 29 untersuchten Berichte der Fall ist.

## **7 Schlussfolgerungen**

### **7.1 Für die Forschung**

Die Untersuchung der Offenlegung und der Verwendung der Bewertungskriterien in den extern in Auftrag gegebenen Evaluationen der Bundesverwaltung anhand ihrer Berichte hat sich als sehr aufwendig und teilweise schwierig erwiesen. Die untersuchten Evaluationsberichte sind alle anders strukturiert, und nicht alle weisen die elementarsten Informationen wie die (klare) Formulierung von Fragestellungen, die systematische Beantwortung der Fragestellungen und einen nachvollziehbaren Bewertungsvorgang auf. Die in der Einleitung (Ziff. 1) von Hense/Widmer aufgeführte Vermutung, dass der Vorgang des Bewertens bisher kaum empirisch erforscht wurde, da es unter anderem in vielen Evaluationen an Transparenz mangle, kann für die ausgewählte Stichprobe bestätigt werden. Um dieses komplexe Thema fundiert zu analysieren, wären neben der Dokumentenanalyse weitere Erhebungsmethoden wie beispielsweise Interviews notwendig. Beleuchtet werden sollte der ganze Bewertungsprozess, da der Teilaspekt der Bewertungskriterien bei sehr vielen Auftraggebern und etlichen Evaluatoreninnen und Evaluatoren zu wenig bekannt sein dürfte.

### **7.2 Für die Evaluationspraxis**

Die Evaluationsberichte, welche die Bundesverwaltung extern in Auftrag gibt, schenken dem Thema Bewertungskriterien und ihrer Verwendung in den untersuchten Fallbeispielen sehr unterschiedlich und der Tendenz nach eher wenig Aufmerksamkeit. Die Bewertungskriterien auf übergeordneter Ebene werden zwar offengelegt, aber in mehr als der Hälfte der Berichte nicht explizit als Be-

wertungskriterien. Ihre Messbarmachung ist meist nicht transparent und nachvollziehbar, und sie werden in nur knapp der Hälfte der Berichte bei der Bewertung des Evaluationsgegenstands alle verwendet. Der methodische Schritt von der sozialwissenschaftlichen Studie hin zu einer professionell durchgeführten Evaluation und ihrer Berichtslegung gelingt in den wenigsten der untersuchten Evaluationsberichte.

Als Auftraggeberin, welche die Evaluationsberichte auch genehmigt, müsste die Bundesverwaltung das Thema mehr berücksichtigen; dies indem sie präzise Fragestellungen formuliert, Vorstellungen zu den Inhalten der Bewertungskriterien in eine Evaluation einbringt oder die Vorschläge der Evaluatorinnen und Evaluatoren diskutiert. Im Rahmen der Qualitätssicherung sollte sie zudem für die Transparenz der Bewertung und für die Qualität der Berichte sorgen. Falls das nötige Knowhow nicht vorhanden ist, müsste dieses aufgebaut werden. Geht man davon aus, dass der Bund als Auftraggeber Evaluationen nach Treu und Glauben an externe Experten und Expertinnen in Auftrag gibt, so müssen auch diese angesprochen werden. Das Bewerten ist ein allgemeines und zentrales Kennzeichen von Evaluation. Dies sollte von den Evaluatorinnen und Evaluatoren professionell gehandhabt werden.

Eine transparente Bewertung und eine hohe Qualität der Berichtslegung der Evaluationen ist ein Gewinn für die Bundesverwaltung als Auftraggeberin, für die Politik und für die interessierte Öffentlichkeit. Die weiterführende Forschung sollte die theoretische Auseinandersetzung mit dem Thema «(Be-)Werten» vorantreiben, damit Massnahmen aufgezeigt werden, die zur Professionalisierung der Evaluation beitragen können.

*Christine Heuer, lic. rer. soc., wissenschaftliche Mitarbeiterin,  
Bundesamt für Gesundheit  
E-Mail: Christine.Heuer@bag.admin.ch*

## Literaturverzeichnis

- Alkin, Marvin C. / Vo, Anne T. / Christie, Christina A., 2012, *The Evaluator's Role in Valuing: Who and With Whom*, *New Directions for Evaluation*, Special Issue, *Promoting Value in Evaluation*, no 133, S. 29–41.
- Beywl, Wolfgang / Niestroj, Susanne, 2009, *Das A-B-C der wirkungsorientierten Evaluation*, 2. Aufl., Köln.
- Geschäftsbericht des Bundesrates 2012 vom 20. Februar 2013 – Band I; [www.bk.admin.ch](http://www.bk.admin.ch) > Dokumentation > Politische Planung > Geschäftsbericht > Archiv Geschäftsbericht Band I
- Geschäftsbericht des Bundesrates 2011 vom 15. Februar 2012 – Band I; [www.bk.admin.ch](http://www.bk.admin.ch) > Dokumentation > Politische Planung > Geschäftsbericht > Archiv Geschäftsbericht Band I
- Hense, Jan / Widmer, Thomas, 2013, *Ein Überblick zum internationalen Stand der Forschung über Evaluation*, in: Hense, Jan / Rädiker, Stefan / Böttcher, Wolfgang / Widmer, Thomas (Hrsg.), *Forschung über Evaluation*, Münster, S. 252–278.
- House, Ernest R. / Howe, Keneth R., 1999, *Values in Evaluation and Social Research*, Thousand Oaks, London, New Delhi.
- IDEKOWI (Interdepartementale Kontaktgruppe Wirkungsprüfungen), 2004, *Wirksamkeit von Bundesmassnahmen. Vorschläge zur Umsetzung von Artikel 170 der Bundesverfassung bei Bundesrat und Bundesverwaltung*, Bern.
- Julnes, George, 2012, *Managing Valuation*, *New Directions for Evaluation*, Special Issue, *Promoting Value in Evaluation*, no 133, S. 3–15.
- Kromrey, Helmut, 2007, *Wissenschaftstheoretische Anforderungen an empirische Forschung und die Problematik ihrer Beachtung in der Evaluation – Oder: wie sich die Evaluationsforschung um das Evaluieren drückt*, *Zeitschrift für Evaluation*, H. 1, S. 113–123.
- OECD (2010): *Qualitätsstandards für die Entwicklungs-evaluierung. DAC-Reihe Leitlinien und Grundsatztexte*. OECDpublishing. Online: [www.oecd.org/dac/evaluation/publicationsanddocuments.htm](http://www.oecd.org/dac/evaluation/publicationsanddocuments.htm)
- Schweizerischer Bundesrat: *Beschluss des Bundesrates vom 3. November 2004: Umsetzung von Artikel 170 BV/Verstärkung der Wirksamkeitsüberprüfung*; [www.bj.admin.ch](http://www.bj.admin.ch) > Staat & Bürger > Wirksamkeitsprüfung / Evaluation
- Schweizerische Evaluationsgesellschaft (SEVAL) (2016): *Evaluationsstandards der Schweizerischen Evaluationsgesellschaft*. Von der Mitgliederversammlung am 9. September 2016 verabschiedet. Bern: SEVAL.
- Scriven, Michael, 2012, *The Logic of Valuing*, *New Directions for Evaluation*, Special Issue, *Promoting Value in Evaluation*, no 133, S. 17–28.
- Scriven, Michael (1991): *Evaluation Thesaurus*. 4th Edition. Thousand Oakes: SAGE.
- Scriven, Michael, 1959, *The Logic of Criteria*, *The Journal of Philosophy*, no 56/22, S. 857–868.
- SEVAL. Stake, Robert E., 2004, *Standards-Based & Responsive Evaluation*, Thousand Oaks, London, New Delhi.
- Stockmann, Reinhard, 2010, *Wissenschaftsbasierte Evaluation*, in: Stockmann, Reinhard / Meyer, Wolfgang, *Evaluation. Eine Einführung*, Opladen und Bloomfield Hill.
- Stockmann Reinhard (Hrsg.), 2007, *Handbuch zur Evaluation. Eine praktische Handlungsanleitung*, Münster, New York, München, Berlin.
- Wottawa, Heinrich / Thierau, Heike, 2003, *Lehrbuch Evaluation*, 3. Aufl., Bern, Göttingen, Toronto, Seattle.

## Résumé

*Les évaluations de mesures de la Confédération doivent être transparentes et de bonne qualité. Une des caractéristiques centrales des évaluations est l'appréciation basée sur des critères fixés à l'avance de manière transparente. La présente contribution expose dans les grandes lignes les résultats d'une étude concernant la publicité et l'utilisation des critères appliqués lors des évaluations externes commandées par l'administration fédérale. Cette étude, qui portait sur 29 rapports d'évaluation, est basée sur l'analyse de documents. L'auteur conclut que l'attention accordée aux critères d'évaluation est différente selon les rapports étudiés et plutôt faible dans l'ensemble.*