

Verena Friedrich / Kathrin Frey / Lisa Guggenbühl

## Evaluationssysteme in Organisationen

### Erkenntnisse aus der Literatur und Praxisbeispiele

---

Wenn Organisationen Projekte durchführen oder fördern und diese evaluiert werden sollen, so stellt sich die Frage: Braucht die Organisation dazu eine eigene Evaluationsabteilung? Wo ist diese am besten verortet? Der Beitrag befasst sich auf der Grundlage einer Literaturanalyse und von Experteninterviews damit, welche strukturellen Varianten es gibt und in welchem Zusammenhang sie mit den unterschiedlichen Zwecken der Evaluation stehen. Dabei werden das Spannungsfeld und die Trade-offs zwischen den Evaluationszwecken «Verbesserung» und «Rechenschaftslegung» beleuchtet. Der Beitrag schliesst mit der Diskussion verschiedener Möglichkeiten, wie die unterschiedlichen Evaluationszwecke im Rahmen von Evaluationssystemen verfolgt werden können.

---

Beitragsart: Wissenschaftliche Beiträge

Zitiervorschlag: Verena Friedrich / Kathrin Frey / Lisa Guggenbühl, Evaluationssysteme in Organisationen, in: LeGes 32 (2021) 1

## Inhaltsübersicht

1. Fragestellungen und Vorgehen
2. Begriffliche Erläuterungen
3. Gestaltung von Evaluationssystemen
  - 3.1. Summative Evaluation: Unabhängigkeit für glaubwürdige Rechenschaftslegung
  - 3.2. Formative Evaluation: Einbezug der Stakeholder, um Lernen zu ermöglichen
  - 3.3. Vor- und Nachteile, Trade-offs
  - 3.4. Mögliche Lösungsansätze
4. Praxisbeispiele
  - 4.1. Stiftung Mercator Schweiz
    - 4.1.1. Gegenstände und Evaluationspraxis
    - 4.1.2. Die Evaluationseinheit und deren Aufgaben
    - 4.1.3. Kultur und Selbstverständnis
    - 4.1.4. Vorteile und Herausforderungen
  - 4.2. Bertelsmann Stiftung
    - 4.2.1. Gegenstände und Evaluationspraxis
    - 4.2.2. Die Evaluationseinheit und deren Aufgaben
    - 4.2.3. Kultur und Selbstverständnis
    - 4.2.4. Vorteile und Herausforderungen
  - 4.3. Bundesamt für Gesundheit
    - 4.3.1. Gegenstände und Evaluationspraxis
    - 4.3.2. Die Evaluationseinheit und deren Aufgaben
    - 4.3.3. Kultur und Selbstverständnis
    - 4.3.4. Vorteile und Herausforderungen
  - 4.4. Independent Evaluation Group (Weltbank)
    - 4.4.1. Gegenstände und Evaluationspraxis
    - 4.4.2. Die Evaluationseinheit und deren Aufgaben
    - 4.4.3. Kultur und Selbstverständnis
    - 4.4.4. Vorteile und Herausforderungen
5. Diskussion

### 1. Fragestellungen und Vorgehen<sup>1</sup>

[1] Projekte, die durch die öffentliche Verwaltung, Organisationen im Bildungs- und Gesundheitswesen, durch private und öffentliche Stiftungen usw. durchgeführt oder gefördert werden, sind häufig Gegenstand von Evaluationen. Diese werden durchgeführt, weil eine gesetzliche Grundlage zur Wirkungsüberprüfung besteht, weil es eine Auflage des Geldgebers ist oder weil die Projektverantwortlichen ihre Arbeit selbst überprüfen und verbessern wollen.

[2] Sobald eine Evaluation keine einmalige Aktivität mehr darstellt, sondern die Evaluationstätigkeit eine stetige Aufgabe ist, stellt sich für eine Organisation die Frage, wie die Evaluationsfunktion in der Organisation eingebettet werden soll. Soll es eine eigene Evaluationsabteilung geben, und wo ist diese gegebenenfalls am besten verortet? Welche Funktionen und Aufgaben soll diese wahrnehmen? Wie gestaltet sich die Zusammenarbeit mit internen und externen Partnern? Dieser Beitrag befasst sich mit der Frage, welche Erkenntnisse sich aus Literatur und Praxis für die Gestaltung solcher Evaluationssysteme in Organisationen ziehen lassen. Dafür wurden eine

---

<sup>1</sup> Dieser Beitrag basiert auf einer Expertise, die durch Verena Friedrich und Kathrin Frey im Auftrag von Gesundheitsförderung Schweiz erarbeitet wurde.

Literaturrecherche<sup>2</sup> und Interviews mit Evaluationsexperten aus vier Organisationen (Stiftung Mercator Schweiz, Bertelsmann Stiftung, Bundesamt für Gesundheit, Independent Evaluation Group/Weltbank) durchgeführt. Der vorliegende Beitrag stellt die Ergebnisse dieser Untersuchung dar. Dazu werden zunächst die relevanten Begriffe erläutert, bevor die zentralen Rechercheergebnisse referiert werden. Danach werden die Evaluationssysteme der vier Beispielorganisationen zusammenfassend dargestellt. Eine Diskussion der Ergebnisse schliesst den Beitrag ab.

## 2. Begriffliche Erläuterungen

- Der Begriff «*Evaluationssystem*»<sup>3</sup> wird verwendet, um eine planvolle Verstetigung der Evaluationsaktivitäten in einer Organisation abzubilden: Es geht um überdauernde Strukturen und Prozesse, in die die Evaluationsfunktion eingebettet ist, mit dem Ziel der systematischen und kontinuierlichen Produktion von handlungsrelevanter Evidenz (LEEuw & FURUBO, 2008; LINDGREN et al. 2016).
- Damit ist auch die Frage verbunden, durch wen das Evaluationssystem selbst sowie die Evaluationsaktivitäten gesteuert werden. Hier kann der Begriff der «*Governance*» ergänzt werden: «The objective of governance is to ensure consistency, accountability and proper authorization to undertake [evaluation] activities. It ensures that the structure, oversight and management processes are in place for the delivery of the expected benefits from an [evaluation]» (Project Management Office, ohne Jahr, S. 2). Wie weiter unten gezeigt wird, sind bei dieser Frage die verschiedenen möglichen Beziehungen bzw. die Nähe und der Einfluss zwischen Projektverantwortlichen, Evaluierenden, Auftraggebenden und der Öffentlichkeit von Bedeutung. PICCIOTTO (2015) spricht von «evaluation governance models», um die verschiedenen möglichen Konfigurationen zu beschreiben.
- Die Aufgabe des «*Evaluationsmanagements*» ist es, die Umsetzung der einzelnen Evaluationen im Sinne des übergreifenden Evaluationssystems und in Übereinstimmung mit den vorgegebenen Governance-Strukturen zu managen. Werden Evaluationen extern in Auftrag gegeben, stellt das Evaluationsmanagement quasi die «Schaltstelle» dar zwischen der Fachabteilung, die innerhalb der auftraggebenden Organisation das zu evaluierende Projekt verantwortet, und den externen Evaluierenden. Zu den Aufgaben des Evaluationsmanagements gehört die Auftragsvergabe, die Begleitung der Evaluation sowie die Valorisation der Evaluationsergebnisse (SEVAL, 2014).

---

<sup>2</sup> Für die Literaturrecherche wurde in den Datenbanken «swissbib Basel Bern» und «Web of Science» mit folgenden Suchbegriffen auf Deutsch und Englisch gesucht: «Evaluationssystem», «M&E-System», «Evaluationsmanagement», «Wirkungsmanagement», «Evaluationsstrategie», «Evaluation Policy», «Institutionalisierung + Evaluation». Die Recherche hat zu vielen Treffern geführt, die teilweise für die Fragestellungen nicht relevant waren. So wurden Publikationen aus der näheren Analyse ausgeschlossen, die über einzelne Evaluationen spezifischer Projekte berichten oder sich mit technischen Systemen für die Datenerhebung oder -analyse befassen. Die ausgewertete Literatur ist im Literaturverzeichnis ersichtlich.

<sup>3</sup> In der Entwicklungszusammenarbeit wird der Begriff meist in Kombination mit «Monitoring» verwendet («monitoring and evaluation [M&E] system»). Dort sind damit die Indikatoren, Tools und Prozesse gemeint, die verwendet werden, um zu überprüfen, ob ein Projekt plangemäss implementiert wurde und die gewünschten Ergebnisse erzielt hat. Auch die Nutzung der Ergebnisse ist üblicherweise Teil dieses Systems (GÖRGENS & KUSEK, 2009; MORRA IMAS & RIST, 2009). Diese Definition wird hier nicht verwendet, da sie sich auf die Evaluation konkreter Projekte bezieht und Frage der institutionellen Einbettung kaum thematisiert.

- Der Begriff «*Evaluationskultur*» bezeichnet die in einer Organisation vorherrschende Praxis, die Traditionen und Einstellungen gegenüber Evaluation. Eine starke Evaluationskultur herrscht dann vor, wenn Evaluation von den Mitarbeitenden und der Leitung wertgeschätzt und unterstützt wird. Dies bedingt in der Regel, dass regelmässig Evaluationen durchgeführt und genutzt werden und dass Prozesse und Strukturen für eine gute Evaluationspraxis institutionalisiert sind (Eval-Wiki, 2016; MAYNE, 2012).

[3] Unserer Ansicht nach greifen diese Konzepte ineinander und müssen in der organisationalen Praxis aufeinander abgestimmt sein. Der Fokus dieses Beitrags liegt auf unterschiedlichen Evaluationssystemen, deren Governance und der damit verbundenen Evaluationskultur. Nicht näher berücksichtigt werden die Evaluation einzelner Projekte und die konkreten Aufgaben des Evaluationsmanagements.

### **3. Gestaltung von Evaluationssystemen**

[4] Bei der Frage, wie ein Evaluationssystem zu gestalten, zu organisieren und zu steuern ist, geht es letztlich darum, wie auf ressourcenschonende Weise eine gute Qualität und ein möglichst grosser Nutzen der Evaluationen sichergestellt werden kann. Aus der Zusammenschau der Literatur wird deutlich, dass die wichtigste Determinante für die Gestaltung des Evaluationssystems der Evaluationszweck ist. Im Folgenden wird die Ausgestaltung der Evaluationssysteme nach den zwei Hauptzwecken von Evaluation, der Rechenschaftslegung und der Verbesserung, dargestellt.

#### **3.1. Summative Evaluation: Unabhängigkeit für glaubwürdige Rechenschaftslegung**

[5] Steht die Rechenschaftslegung und Steuerung als Evaluationszweck im Vordergrund, so sollte die Steuerung der Evaluation unabhängig sein von den Projektverantwortlichen, damit keine Interessenskonflikte entstehen und die Glaubwürdigkeit der Ergebnisse sichergestellt werden kann. Die Evaluation wird in diesem Modell von einer unabhängigen Evaluationseinheit gesteuert (Abb. 1). Die Evaluationen sind auf die Informationsbedürfnisse der Instanz abgestimmt, der gegenüber die Rechenschaftslegung erfolgt. Häufig wird dabei angestrebt, dass die unterschiedlichen Projekte auf die gleiche Art und Weise evaluiert werden, um sie vergleichen und übergeordnete Entscheide treffen zu können. Dies führt dazu, dass die einzelnen Evaluationen nicht oder nur beschränkt auf die Besonderheiten von einzelnen Projekten und deren Kontexte eingehen können.

[6] In der Literatur wird festgehalten, dass die Unabhängigkeit nicht einfach dadurch sichergestellt werden kann, dass die Evaluationen als Auftrag an externe Evaluationsdienstleister vergeben werden. DE LAAT (2014) betont, dass ein externer Dienstleister nicht nur oft abhängig vom nächsten Auftrag ist, sondern im Konfliktfall auch relativ leicht ausgetauscht werden kann; er steht also in einem Abhängigkeitsverhältnis zur Auftraggeberin (Abb. 2). Vielmehr kommt es darauf an, wer festlegt, welche Projekte evaluiert werden sollen und wer die Evaluationen plant, vergibt und begleitet. Ist diese Instanz in irgendeiner Weise abhängig von den Projektverantwortlichen, so besteht die Gefahr, dass diese auf die Auswahl der Projekte, die Evaluationsplanung oder Berichtlegung Einfluss nehmen. Bringt die Evaluation negative Ergebnisse ans Licht, so kön-

nen diese möglicherweise nicht unbefangen aufgedeckt werden. Damit das System also sowohl «good news» als auch «bad news» produzieren kann, ist es wichtig, dass die Evaluatorinnen und Evaluatoren wie auch das Evaluationsmanagement vor möglichen Repressalien geschützt werden (MORRA IMAS & RIST, 2009). Dafür sind geeignete organisationale Massnahmen zu treffen, wie beispielsweise die finanzielle und personalmässige Unabhängigkeit der Evaluationseinheit, die im Idealfall auch nicht dem Management bzw. der Direktion, sondern einem Aufsichtsorgan (z. B. dem Stiftungsrat) untersteht. Die weiter unten beschriebene Independent Evaluation Group der Weltbank entspricht diesem Modell.

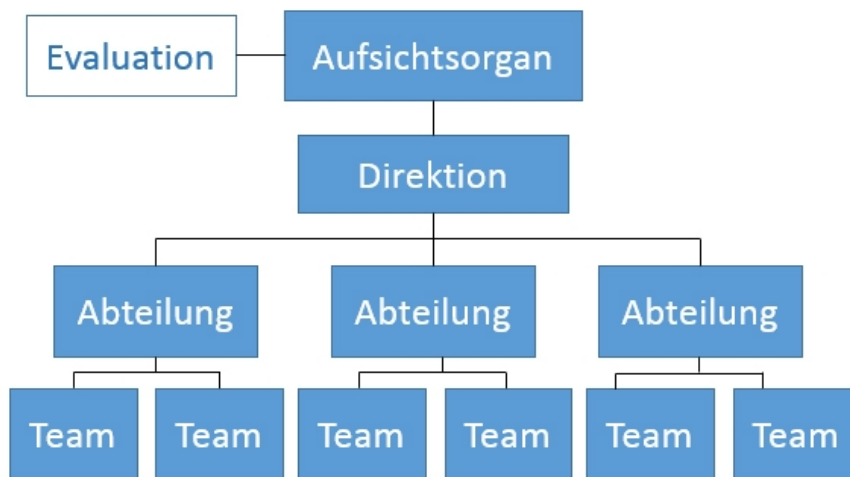


Abbildung 1: Zentralisiertes Modell (eine unabhängige Evaluationsabteilung erstellt das Arbeitsprogramm, das vom Aufsichtsorgan bewilligt wird, und führt die Evaluation durch bzw. gibt sie in Auftrag; Direktion und Abteilungen sind nur beratend einbezogen).

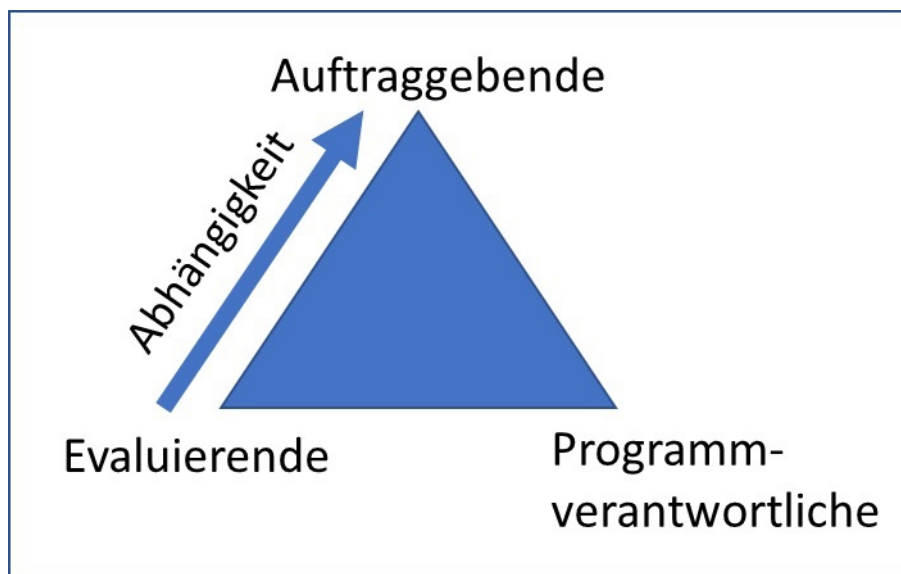


Abbildung 2: Vergibt eine Evaluationsabteilung einen Evaluationsauftrag an externe Auftragnehmer, befinden sich diese in einem Abhängigkeitsverhältnis (Abb. in Anlehnung an DE LAAT, 2014).

### 3.2. Formative Evaluation: Einbezug der Stakeholder, um Lernen zu ermöglichen

[7] Ist mit der Evaluation der Zweck verbunden, dass die Evaluationsergebnisse der Verbesserung der einzelnen Projekte dienen sollen, so ist es sinnvoll, die Projektverantwortlichen an der Steuerung der Evaluation zu beteiligen, beispielsweise indem sie selbst die Evaluation in Auftrag geben oder an deren Planung und gegebenenfalls an der Durchführung massgeblich mitwirken. Die Evaluationsabteilung ist in die operative Struktur integriert und steht den Projektverantwortlichen beratend zur Seite (gestrichelte Linien in Abb. 3). Die Evaluationen sind auf deren Informationsbedürfnisse abgestimmt und auf die jeweiligen Projekte individuell zugeschnitten.

[8] Die Projektverantwortlichen sind nicht nur wichtige Auskunftspersonen über die Projekte, deren Ziele und Kontext, sie sind auch die eigentlichen Adressaten der Evaluationsergebnisse. Sie sollen die Evaluationsergebnisse zur Verbesserung der Projekte nutzen. Die Forschung über die Nutzung der Evaluation zeigt, dass der Einbezug der Stakeholder bei der Evaluation dazu führt, dass die Evaluationsergebnisse eher genutzt werden. Dies ist zum einen darauf zurückzuführen, dass die Stakeholder Fragestellungen einbringen können, die für sie relevant sind; zum anderen entsteht durch die Mitwirkung eine höhere «Ownership» für die Evaluation und deren Ergebnisse. (FETTERMAN, RODRÍGUEZ-CAMPOS, & ZUKOSKI, 2017; JOHNSON et al., 2009; KING, 2007; PATTON, 1997).

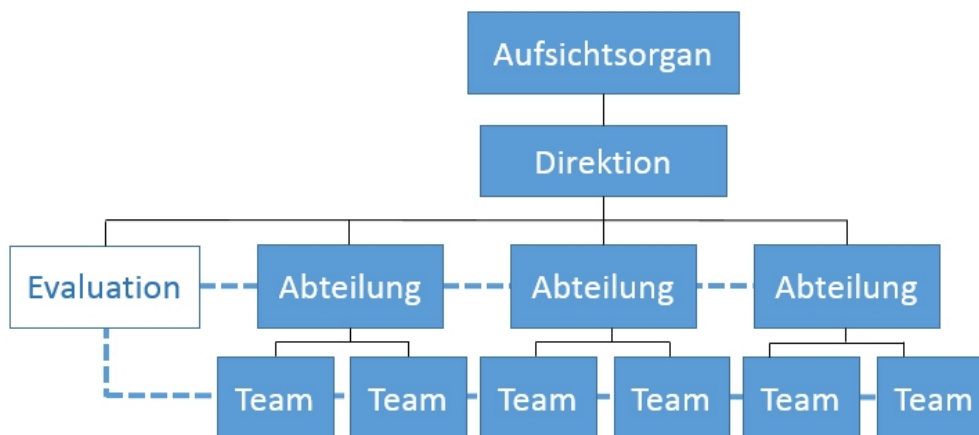


Abbildung 3: Dezentrales Modell (die Evaluationsabteilung ist in die operative Struktur integriert und steht den Projektverantwortlichen bei deren Evaluationsanliegen beratend zur Seite).

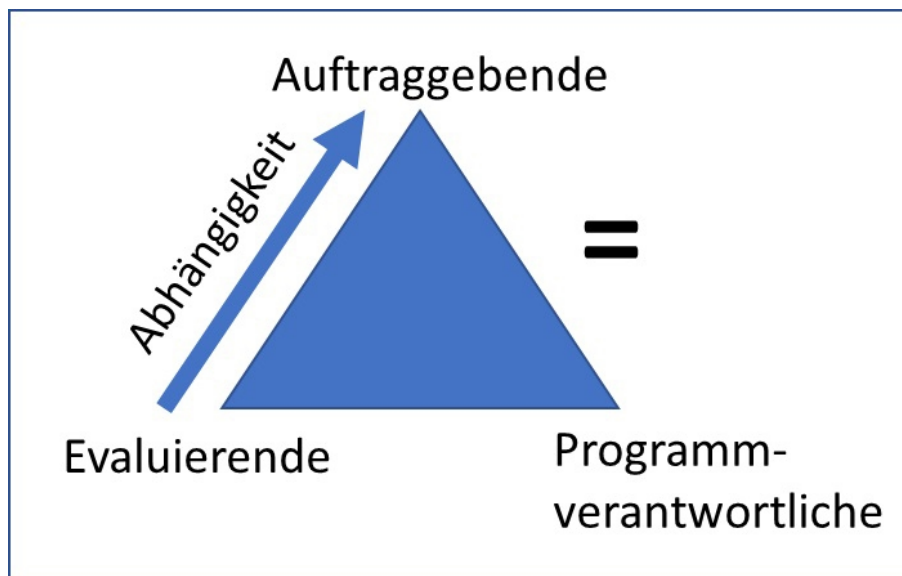


Abbildung 4: Im dezentralen Modell fallen die Rollen der Auftraggebenden und der Projektverantwortlichen mehr oder weniger stark zusammen (Abb. in Anlehnung an DE LAAT, 2014).

[9] Die weiter unten beschriebenen Beispiele der Bertelsmann Stiftung und der Stiftung Mercator Schweiz folgen diesem Modell. Bei der Bertelsmann Stiftung geht die Steuerung durch die Projektverantwortlichen sogar so weit, dass diese selbst entscheiden können, ob überhaupt eine Evaluation «ihres» Projektes durchgeführt werden soll, oder ob ein anderer Ansatz (z. B. eine Beratung) das nützlichere Instrument ist, um das Projekt zu verbessern. Beide Stiftungen stellen nicht das Instrument der Evaluation, sondern generell die «Wirkungsorientierung» in den Vordergrund. Dieses Konzept besagt, dass die gesamte Stiftungsarbeit von der Idee geleitet sein sollte, dass die Aktivitäten der Stiftung eine Wirkung erzielen.

### 3.3. Vor- und Nachteile, Trade-offs

[10] Aus der Literatur werden verschiedene Vor- und Nachteile sowie Trade-offs der oben beschriebenen idealtypischen Modelle deutlich. WIDMER und NEUENSCHWANDER (2004) fassen dies folgendermassen zusammen: «It is obvious that the decentralized type fits an evaluation whose purpose is improvement far better than does the centralized type, which is in turn more appropriate for an accountability-oriented evaluation.»

[11] Das «dezentrale» Modell, das die Projektverantwortlichen einbezieht, hat also den Vorteil, dass die Evaluationsergebnisse an die Projektpraxis gut anschlussfähig sind und unmittelbar für die Verbesserung der jeweiligen Projekte genutzt werden können. BALTHASAR (2008) hat gezeigt, «dass die Verwendung von Evaluationen von einer Institutionalisierung profitiert, die die Distanz zwischen den Evaluierenden und den Nutzenden von Evaluationen gering hält». Dadurch, dass die einzelnen Evaluationen aber auf die Informationsinteressen der Projektverantwortlichen zugeschnitten sind, ist es oft nur schwer möglich, auch projektübergreifende «lessons learned» als Grundlage für die Steuerung der Gesamtorganisation abzuleiten (PROM-JACKSON & BARTSIOTAS, 2014).

[12] Der Vorteil des «zentralisierten» Modells ist der unabhängige Blick von aussen, der zur unparteiischen und glaubwürdigen Einschätzung beitragen soll. Als Nachteil dieses Modells wird häufig genannt, dass durch die grössere Distanz zwischen den Evaluatorinnen und Evaluatoren und den Projektverantwortlichen die Nützlichkeit der Evaluation für Entscheide in der Projektpraxis nicht gegeben ist: «ensuring the independence of the process may weaken the potential for providing a strong linkage between the evaluation findings and decision-making» (OECD, 1991, S. 6). Auch BALTHASAR (2008, S. 243) kommt zu dem Schluss, dass es «kaum Belege dafür [gibt], dass sich die Chance auf Verwendung von Evaluationsergebnissen erhöht, wenn die Institutionalisierung Wert auf Distanz und Unabhängigkeit zwischen Evaluierenden und Evaluierten legt».

### 3.4. Mögliche Lösungsansätze

[13] Wenn nun eine Organisation sowohl summative wie auch formative Evaluationsanliegen hat, stellt sich die Frage, wie sie das erreichen kann und wie sie ihr Evaluationssystem dazu am besten gestalten sollte. Ein naheliegender Gedanke ist, dass eine Evaluationsabteilung beide Funktionen – formative und summative Evaluation – wahrnehmen könnte, oder dass sogar innerhalb einer konkreten Evaluation beide Zwecke verfolgt werden. Diese Vermischung kann aber sehr problematisch sein, da es bei der formativen Evaluation darum geht, aus Misserfolgen zu lernen. Dafür müssen diese offengelegt werden. Bei der Rechenschaftslegung geht es darum, Erfolge vorzuweisen – was eine grundlegend andere Auskunftsmotivation der Projektverantwortlichen nach sich zieht. Patton (persönliche Kommunikation) spricht daher davon, eine «Brandmauer» zwischen beiden Funktionen der Evaluation zu errichten. Auch WIDMER und FREY (2006) sowie WIDMER et al. (2015) sehen die Vermischung der beiden Evaluationsfunktionen kritisch: Wenn beide Zwecke verfolgt werden, besteht die Gefahr, dass die Verantwortlichkeit für die Nutzung unklar bleibt und die Ergebnisse letztendlich gar nicht genutzt werden.

[14] HAWKINS (2014) diskutiert, wie der «Spagat» zwischen der Unabhängigkeit einerseits und der Forderung nach Nützlichkeit, Relevanz und Responsivität andererseits zu schaffen ist: Sie fordert, dass die Evaluationsabteilung ein eindeutiges Mandat der obersten Leitung haben muss, dass eine klare Evaluationsstrategie und Guidelines für die Durchführung von Evaluationen vorhanden sein müssen und dass die Evaluationsabteilung unabhängig von (oder zumindest gleichberechtigt mit) anderen Abteilungen der Organisation sein muss. Gleichzeitig müssen Mechanismen und Kommunikationskanäle geschaffen werden, die es den Mitarbeitenden der Evaluationseinheit erlauben, «am Ball» und in engem Austausch mit den übrigen Mitarbeitenden der Organisation zu bleiben in Bezug auf inhaltliche Diskussionen, relevante Themen und aktuelle Trends.

[15] Wenn in einer Organisation beide Zwecke relevant sind, besteht eine weitere mögliche Lösung daher darin, die Funktionen auch organisational und personell zu trennen: Die Abteilung für formative Evaluation wäre nahe bei den Einheiten anzusiedeln, die die Strategie verantworten, die Projekte planen und umsetzen. Sie würde dort beratend und unterstützend zur Seite stehen, wobei der Fokus auf kurzen Feedbackschleifen zuhanden der Projektverantwortlichen läge (HAWKINS, 2014). Die Evaluation würde sich am Informationsbedürfnis der Projektverantwortlichen orientieren und aus deren Budget finanziert. Die Abteilung für summative Evaluation wäre demgegenüber bei einer Aufsichtsfunktion anzusiedeln, die im Auftrag und zuhanden des Aufsichtsorgans und mit eigenem Budget grössere Evaluationen wichtiger Projekte durchführt.



[16] Eine weitere mögliche Lösung, die beiden Funktionen zu kombinieren, liegt in der Idee der «accountability for learning». Dabei wird auf eine summative Evaluation im obigen Sinne verzichtet, stattdessen findet eine formative Evaluation der einzelnen Projekte statt. Der summative Aspekt besteht darin, dass die Projektverantwortlichen nachweisen müssen, dass ihre Projekte evaluiert wurden, was sie aus der Evaluation gelernt haben und wie sie die Erkenntnisse umsetzen werden. Die Rechenschaftspflicht verschiebt sich also vom eigentlichen Projekt auf die Evaluation und die Nutzung der Evaluationsergebnisse.

## **4. Praxisbeispiele**

[17] In den folgenden Absätzen werden die Experteninterviews zusammengefasst, um einen Einblick in die Evaluationssysteme verschiedener Organisationen zu erhalten. Ein Schwerpunkt wurde dabei auf Stiftungen gelegt, die Projekte fördern (Stiftung Mercator Schweiz) und Projekte selbst durchführen (Bertelsmann Stiftung). Zusätzlich wurden zwei Organisationen (Bundesamt für Gesundheit und Weltbank) ausgewählt, die über langjährige Erfahrung mit Evaluation verfügen und in denen die Evaluation bereits seit mehreren Jahrzehnten institutionalisiert ist. Zur besseren Einordnung werden bei jedem Beispiel auch die typischen Evaluationsgegenstände und die Evaluationspraxis vorgestellt. Alle Aussagen und Informationen in den folgenden Absätzen beruhen auf den Angaben der Interviewpartner.<sup>4</sup>

### **4.1. Stiftung Mercator Schweiz**

[18] Die Stiftung Mercator Schweiz ist eine private Familienstiftung, die – mit Förderpartnern und intern – Projekte in den Themenbereichen Umwelt, Bildung, Verständigung und Mitwirkung umsetzt. Das Interview wurde mit dem Leiter Wirkungsorientierung geführt.

#### **4.1.1. Gegenstände und Evaluationspraxis**

[19] Evaluationsgegenstände in der Stiftung sind die Projekte und Programme, die von den Förderpartnern und von der Stiftung selbst in den oben genannten Themenbereichen umgesetzt werden. Dabei besteht die Regel, dass sich alle Projekte ab einem Fördervolumen von 30 000 Franken selbst evaluieren müssen. Bereits im Förderantrag müssen dafür das Wirkungsmodell, die Projektziele sowie die Indikatoren für die Überprüfung der Zielerreichung genannt sein; der Abschlussbericht zuhanden der Stiftung enthält die Evaluationsergebnisse. Wenn aus Sicht der Stiftung spezifische Fragen zu einem Projekt bestehen (z. B. bei Pilotprojekten), kann auch eine externe Evaluation eine Auflage für die Förderung sein. In diesem Fall bringt die Stiftung ihre Fragestellungen ein; der Auftraggeber für die Evaluation und primäre Adressat der Ergebnisse bleibt aber der Förderpartner selbst.

---

<sup>4</sup> Bei der Weltbank wurde zusätzlich Literatur hinzugezogen, die im entsprechenden Absatz zitiert ist.

#### **4.1.2. Die Evaluationseinheit und deren Aufgaben**

[20] Aufgabe aller Projektmanager der Stiftung ist es, gemeinsam mit den Förderpartnern an der Wirkungsorientierung zu arbeiten. Dies bedeutet, dass die Projektmanager die Förderpartner darin unterstützen, die Wirkungsmodelle, Zielformulierungen und Indikatoren zu entwickeln, die Teil des Förderantrags sind. Der Leiter Wirkungsorientierung bietet zusätzlich mit externen Anbietern projektunabhängig interne Unterstützung für die Projektmanager an (Weiterbildung, Leitfäden etc.). Bei Projekten, die die Stiftung selbst umsetzt, unterstützt er zusätzlich bei der Konzeption der Wirkungsorientierung. Es gibt also nicht in diesem Sinne eine Evaluationseinheit, vielmehr befasst sich jeder Projektmanager mit dem Thema, da die Wirkungsorientierung Bestandteil jedes Projektes ist. Wie alle Mitarbeitenden ist auch der Leiter Wirkungsorientierung der Geschäftsführung unterstellt, die wiederum dem Stiftungsrat untersteht.

#### **4.1.3. Kultur und Selbstverständnis**

[21] In der Stiftung Mercator Schweiz steht das übergreifende Konzept der Wirkungsorientierung im Vordergrund. Dies ist der Geschäftsleitung und dem Stiftungsrat wichtig und für die Stiftung identitätsstiftend; auch die Mitarbeitenden tragen dies mit. Dabei geht es weniger um die Erhebung von möglichen Wirkungen oder um die Evaluation per se, sondern darum, dass die Förderpartner für ihre Projekte ein Wirkungsmodell entwickeln, realistische Ziele setzen und überlegen, wie man die Zielerreichung überprüfen kann. Im Vordergrund steht das formative Element und der Nutzen für die Förderpartner; die Ownership liegt bei ihnen. Der Leiter Wirkungsorientierung sieht seine Rolle in der Wissensvermittlung und Beratung nach innen (Entwicklung eines gemeinsamen Verständnisses und Harmonisierung der Praxis) und aussen (Aufbau von Evaluationskompetenzen).

#### **4.1.4. Vorteile und Herausforderungen**

[22] Ein Grossteil der Evaluationsverantwortung liegt bei den Förderpartnern; sie steuern die Evaluation massgeblich mit, dadurch steigt für sie der formative Nutzen der Evaluation. Dadurch, dass massgebliche Arbeiten bereits in den Projektantrag integriert sind, ist der Aufwand für die eigentliche Evaluation weniger hoch. Gemäss unserem Interviewpartner berichten viele Förderpartner, dass das wirkungsorientierte Arbeiten für sie von Vorteil sei, sowohl inhaltlich für das Projekt als auch für den Aufbau von Evaluationskompetenzen. Als Herausforderung werde beschrieben, dass die Antragstellung aufwendig ist und für manche Förderpartner eine «steile Lernkurve» bedeutet. Teilweise lässt sich, nach Auskunft des Interviewpartners, eine Verunsicherung bei Antragstellern beobachten, ob von vornherein nur solche Ziele formuliert werden sollen, die auch messbar sind. Dem versuchen die Projektmanager im direkten Kontakt entgegenzuwirken. Gemäss Interviewpartner ist die Qualität der Selbstevaluationen bei den verschiedenen Projekten sehr unterschiedlich. Wenn die Qualität von externen Evaluationen als problematisch eingeschätzt wird, hat die Stiftung – da sie nicht die Auftraggeberin ist – keine Interventionsmöglichkeit. Dadurch, dass alle Projektmanager in die Arbeiten zur Wirkungsorientierung einbezogen sind, ist für die Stiftung auch das gemeinsame Verständnis und die einheitliche Kommunikation gegen aussen eine Herausforderung, ebenso wie die Frage, wie die Stiftung die Ergebnisse aus den einzelnen Projektevaluationen systematisch übergreifend nutzen kann: Im Moment gibt es bei der Stiftung noch kein System, wie die einzelnen Erkenntnisse festgehalten und für die

Strategieentwicklung nutzbar gemacht werden können. Das Lernen für die Gesamtorganisation wird als zunehmend wichtig erachtet, gleichzeitig kann dies zum formativen Evaluationszweck auf Projektebene im Widerspruch stehen.

## **4.2. Bertelsmann Stiftung**

[23] Die Bertelsmann Stiftung ist eine operativ tätige Stiftung des privaten Rechts, die durch ihre Projekte einen Beitrag leisten will zur Weiterentwicklung der deutschen Gesellschaft. Die Stiftung ist in verschiedenen gesellschaftlichen Bereichen aktiv (u. a. Bildung, Gesundheit, Demokratie, Europa, Wirtschaft). Das Interview wurde mit dem Senior Expert Wirkungsorientierung geführt.

### **4.2.1. Gegenstände und Evaluationspraxis**

[24] Evaluationsgegenstände sind die Projekte und Programme, die die Stiftung selbst (teilweise mit Projektpartnern) in den oben genannten Bereichen umsetzt. Typischerweise bestehen diese aus einer Kombination von Interventionen mit unterschiedlichen Schwerpunkten: Angebote für Zielgruppen, Agenda Setting, Siegel und Preise, Netzwerken, Monitoring usw. Für die Evaluation gibt es zwei Ansätze: Zum einen werden Projekte zum gleichen Thema circa alle vier bis fünf Jahre im Rahmen einer Programmevaluation extern evaluiert. Zum anderen werden die Projektverantwortlichen mit verschiedenen Angeboten bei der wirkungsorientierten Projektarbeit unterstützt. Ob, wann und in welcher Form die Projektverantwortlichen zusätzlich zu den periodischen Programmevaluationen eine Evaluation ihres Projekts umsetzen oder ob sie einen anderen Ansatz wählen, bleibt ihnen überlassen. Während bei den Programmevaluationen Steuerung und Legitimation im Vordergrund stehen (wobei auch hier versucht wird, dem Zweck des Lernens verstärkt Raum zu geben), haben Projektevaluationen eine formative Funktion.

### **4.2.2. Die Evaluationseinheit und deren Aufgaben**

[25] Auch in der Bertelsmann Stiftung gibt es keine eigentliche Evaluationseinheit. Der Experte für Wirkungsorientierung, der dem Vorstand unterstellt ist, hat die Aufgabe, die Projektverantwortlichen zum Thema Wirkungsorientierung zu informieren und zu beraten (interne Weiterbildungen, Beratung bei der Erstellung von Wirkungsmodellen etc.). Bei Programmevaluationen gibt es in der Regel eine enge Abstimmung zwischen dem Experten Wirkungsorientierung, den Programmverantwortlichen und dem «Gremienbüro» (Stabsstelle des Vorstands, die bei Programmevaluationen die Aufgabe des operativen Evaluationsmanagements übernimmt); es gibt aber auch Programmverantwortliche, die auf den Einbezug des Experten für Wirkungsorientierung verzichten.

### **4.2.3. Kultur und Selbstverständnis**

[26] Wie bei der Stiftung Mercator Schweiz steht auch bei der Bertelsmann Stiftung das übergreifende Konzept der Wirkungsorientierung im Vordergrund. Dahinter steht die Auffassung, dass es nicht sinnvoll ist, eine Stabsstelle einzurichten, in der die Evaluation zentralisiert ist, und die Evaluation «von oben» zu verordnen. Vielmehr sollten das Wissen und die Einstellung breit bei

den Mitarbeitenden verankert sein. Der Experte Wirkungsorientierung versteht sich als Berater und Unterstützer und verfolgt das Ziel, die Mitarbeitenden zu befähigen, selbst zu entscheiden, welches Instrument für die Verbesserung ihres Projektes geeignet ist. Auch bei den externen Programmevaluationen soll die Ownership bei den Programmverantwortlichen liegen: obwohl formell der Vorstand die Programmevaluationen beauftragt, sind es de facto die Programmverantwortlichen, die diese Rolle ausfüllen.

[27] Generell wird davon ausgegangen, dass die Mitarbeitenden intrinsisch motiviert sind und Evaluation als im eigenen Interesse, als Teil eines professionellen Projektmanagements und einer ethischen Arbeitshaltung auffassen. Die Mitarbeitenden nehmen die Projektevaluationen entsprechend als Lerngelegenheit wahr. Zu den Programmevaluationen gibt es ein breiteres Spektrum an Meinungen, wobei auch hier eine Entwicklung auszumachen ist: Während sie anfangs eher als bürokratisches «Abarbeiten» wahrgenommen wurden, als «compliance activity», werden auch die Programmevaluationen inzwischen vermehrt als Chance wahrgenommen, besonders, wenn sich die Programmverantwortlichen selbst am Prozess beteiligen.

#### **4.2.4. Vorteile und Herausforderungen**

[28] Unser Interviewpartner berichtet, dass es durch den Ansatz gelingt, die Programm- und Projektverantwortlichen zu beteiligen, Wirkungsorientierung als Denk- und Arbeitsweise breit zu verankern und dem Leitbild der lernenden Organisation näher zu kommen. Die Herausforderung bestehe im «Spagat» zwischen der formativen und der summativen Evaluationsfunktion; mit dem gewählten Ansatz ist die summative Funktion eher gering ausgeprägt. Die Konzentration liegt stark auf der operativen Ebene der Projekte und Programme. Die organisationale Ebene ist noch wenig im Blick: Bislang werden die Ergebnisse des wirkungsorientierten Arbeitens noch wenig für die strategische Steuerung der Gesamtorganisation genutzt. Dies ist gemäss unserem Interviewpartner eine Herausforderung, auch, da die Themenvielfalt in den verschiedenen Projekten und Programmen in der Bertelsmann Stiftung so gross ist (was eine übergreifende Auswertung erschwert) und die Akteure grossen Wert auf ihre Autonomie legen.

### **4.3. Bundesamt für Gesundheit**

[29] Das Bundesamt für Gesundheit (BAG) gehört zum Eidgenössischen Departement des Inneren und setzt die schweizerische Gesundheitspolitik um. Zum Evaluationssystem des BAG hat der Leiter der Fachstelle Evaluation und Forschung Auskunft gegeben (vgl. BAG, 2020).

#### **4.3.1. Gegenstände und Evaluationspraxis**

[30] Evaluationsgegenstände im BAG sind Gesetze, Verordnungen, Strategien oder Projekte im Gesundheitsbereich, die durch das BAG selbst, durch Kantone, Leistungserbringer oder die Wirtschaft (häufig im Verbund) umgesetzt werden. Die Geschäftsleitung des BAG entscheidet jährlich darüber, welche Projekte evaluiert werden sollen; in circa 95 Prozent der Fälle existiert eine spezialgesetzliche Grundlage für die Evaluation. Die Fachstelle Evaluation und Forschung verantwortet das Evaluationsmanagement; sie vergibt die Evaluationsaufträge an externe Dienstleister.

[31] Beauftragt wird die Evaluation von der Geschäftsleitung des BAG. Die Projektleitung der Evaluation obliegt dem Evaluationsmanagement, das in der Fachstelle Evaluation und Forschung

angesiedelt ist. Die Organisation der Evaluationsprojekte sieht jeweils eine Steuer- und Begleitgruppe vor. Den Vorsitz der Steuergruppe übernimmt in der Regel die Leitung des betroffenen Direktionsbereichs. Die Gruppen setzen sich je nach Fall aus BAG-internen und -externen Stakeholdern zusammen.

#### **4.3.2. Die Evaluationseinheit und deren Aufgaben**

[32] Die Fachstelle Evaluation und Forschung gehört zum Direktionsstab – ein Bereich neben anderen thematischen Bereichen im BAG, in dem u. a. allgemeine Aufgaben angesiedelt sind. Die Leiterin des Direktionsstabes ist die Vorgesetzte im Rahmen der Grundorganisation. Die Fachstelle erbringt für alle Direktionsbereiche das Evaluationsmanagement als Dienstleistung; das Budget für die Evaluationen ist bei der Fachstelle angesiedelt. «Vorgesetzte» im Rahmen der einzelnen Evaluationsprojekte ist formal die Geschäftsleitung des BAG, wobei diese de facto von der Direktionsbereichsleitung, die auch das zu evaluierende Projekt verantwortet, vertreten wird.

#### **4.3.3. Kultur und Selbstverständnis**

[33] Im BAG ist die Evaluation zum einen formal eingebunden in die ordentlichen Planungs- und Steuerungsprozesse: Die Geschäftsleitung legt das jährliche Evaluationsprogramm fest und stützt ihre Entscheide systematisch auf vorliegende Evaluationsergebnisse. Evaluationen werden also gemäss unserem Interviewpartner von massgeblichen Leitungsgremien als wichtige Entscheidungsgrundlage wahrgenommen und genutzt. Zum anderen habe sich über die Jahre auch bottom-up eine Kultur entwickelt, in der deutlich werde, dass viele Mitarbeitenden den Wert der Evaluation erkennen: Die Evaluation werde wertgeschätzt und nicht als mühsame Pflicht wahrgenommen. Allerdings müssen neue Mitarbeitende mit jedem Projekt wieder überzeugt werden: Jede Evaluation ist wieder eine neue Referenz für das Instrument Evaluation.

[34] Wichtig für die Akzeptanz sei, dass die Evaluation hauptsächlich der Optimierung diene (in ca. 65 % der Evaluationen geht es um Verbesserung; in ca. 35 % auch um Entscheidungsfindung), dass der Evaluationszweck sowie die wichtigen Rechte und Pflichten im Pflichtenheft klar und für alle transparent ausgewiesen seien, dass die Evaluationsstandards allen bekannt seien, dass die Evaluation insgesamt ein «offenes Spiel» darstelle (jeder kann zu den Evaluationsresultaten und Empfehlungen Stellung nehmen) und dass die jeweils relevanten Entscheidungsträger massgeblich am Prozess beteiligt seien.

[35] Die Fachstelle Evaluation und Forschung versteht ihre Rolle darin, der Geschäftsleitung und den Direktionen Wissen an die Hand zu geben zur Weiterentwicklung und Verbesserung der gesundheitspolitischen Massnahmen. Ihre Aufgabe besteht darin, darauf zu achten, dass die «Spielregeln» eingehalten werden (bspw. zu verhindern, dass Projektverantwortliche unzulässigen Einfluss auf die Evaluation nehmen), und alles vorzubereiten, damit die externen Evaluierenden optimal arbeiten können und die Evaluationsergebnisse genutzt werden.

#### **4.3.4. Vorteile und Herausforderungen**

[36] Die Organisation und die Prozesse sind über 30 Jahre laufend optimiert worden; inzwischen sind die Prozesse standardisiert und gut implementiert. Dies habe, so der befragte Fachstellenleiter, den Vorteil, dass «das Rad nicht jedes Mal neu erfunden werden müsse» und alle Beteiligten

den Prozess kennen würden. Die Ergebnisse werden genutzt (auf der Ebene der einzelnen Projekte), da die Evaluation in den Rahmen der ordentlichen Planungs- und Steuerungsprozesse eingebunden ist und weil die Projektverantwortlichen massgeblich bei der Beauftragung der Evaluation und im Evaluationsprozess beteiligt sind. Ohne ihren Einbezug befände sich eine Evaluation im «Niemandland», und es könnten nicht die relevanten Diskussionen geführt und Entscheidungen getroffen werden. Als Herausforderung wird beschrieben, dass potenziell Interessenskonflikte entstehen können, da der Verantwortliche für den Evaluationsgegenstand im Rahmen der Projektorganisation gleichzeitig der «Vorgesetzte» des Evaluationsmanagers ist. Eine Vereinnahmung der Evaluation konnte bisher aber durch den Fachstellenleiter unter Rückgriff auf die Evaluationsstandards verhindert werden. Die Aufbereitung von Evaluationsresultaten aus mehreren Studien zu Synthesen findet im Moment aus Ressourcengründen nur partiell statt.

#### **4.4. Independent Evaluation Group (Weltbank)**

[37] Die Weltbank bezeichnet die in Washington, D.C., angesiedelte Weltbankgruppe, eine multinationale Entwicklungsbank. Die Weltbankgruppe umfasst die Internationale Bank für Wiederaufbau und Entwicklung, die Internationale Entwicklungsorganisation, die Internationale Finanz-Corporation, die Multilaterale Investitions-Garantie-Agentur und das Internationale Zentrum für die Beilegung von Investitionsstreitigkeiten. Deren Aufgabe ist es, durch finanzielle und technische Hilfe und Beratung die Wirtschaftsentwicklung der weniger entwickelten Mitgliedsstaaten zu fördern. Das Interview wurde mit einem Mitarbeiter der Independent Evaluation Group (IEG) geführt; ergänzt wurden die Angaben mit Informationen aus anderen Quellen (IBRD, 2019; VAESSEN, 2018).

##### **4.4.1. Gegenstände und Evaluationspraxis**

[38] Die IEG evaluiert die Aktivitäten aller oben genannten Institutionen mit Ausnahme des Zentrums für die Beilegung von Investitionsstreitigkeiten. Die Gegenstände umfassen einerseits interne Strategien und Prozesse, andererseits die Dienstleistungen der Weltbank, vom Einzelprojekt bis zu Länderprogrammen, thematischen Programmen und Sektor-Programmen (IBRD, 2019). Dabei werden drei Arten von Evaluationen unterschieden: unabhängige Evaluationen, verbindliche Selbstevaluationen und nachfrageorientierte Selbstevaluationen. Unabhängige Evaluationen werden durch die IEG gemanagt und durchgeführt. Sie dienen in erster Linie der Rechenschaftslegung und in zweiter Linie dem Lernen. Die Adressaten sind vorrangig das Board of Governors und die Boards of Directors der verschiedenen Institutionen. Welche unabhängigen Evaluationen durchgeführt werden sollen, wird durch die IEG unter Einbezug der Institutionen diesen Leitungsgremien vorgeschlagen, die den Arbeitsplan für die IEG verabschieden.

[39] Verbindliche Selbstevaluationen sind im Projekt- oder Programm-Zyklus eingebettet und werden durch die jeweils verantwortliche operative Einheit gemanagt (auch wenn von «Selbstevaluation» gesprochen wird, heisst das nicht, dass diese Evaluationen nicht extern in Auftrag gegeben werden können). Nachfrageorientierte Selbstevaluationen werden auf spezifische Nachfrage (z. B. von Geldgebern) durchgeführt. Sie sind strukturell in Managementprozesse eingebettet, werden jedoch häufig von funktional unabhängigen Evaluationseinheiten innerhalb der Institution durchgeführt, gemanagt oder beauftragt. Beide Arten der Selbstevaluation sind in der Regel eng mit der Entscheidungsfindung und den Lernprozessen in den Institutionen verbun-

den; Adressaten sind in erster Linie die operativen Einheiten und deren Management, Kunden und Entwicklungspartner (IBRD, 2019).

#### **4.4.2. Die Evaluationseinheit und deren Aufgaben**

[40] Im Folgenden wird nicht auf die Evaluationseinheiten innerhalb der verschiedenen Institutionen der Weltbank eingegangen, sondern nur auf die Independent Evaluation Group (IEG). Die IEG ist vom Management der Weltbank unabhängig und berichtet direkt den Boards of Directors, die dem Präsidenten der Weltbank vorgesetzt sind. Die IEG ist strukturell und funktional unabhängig, d. h. dass Entscheidungen über Personalressourcen und finanzielle Ressourcen durch die IEG unabhängig vom Management gefällt werden und dass die IEG unabhängig vom Management über die Evaluationsgegenstände und über das Vorgehen bei der Evaluation entscheidet (VAESSEN, 2018). Die IEG vergibt keine Evaluationsaufträge an externe Evaluatoreninnen und Evaluatoren, vielmehr wird ein hybrides Modell umgesetzt: Das Evaluationsteam setzt sich aus IEG-Mitarbeitenden und externen Consultants zusammen; die Leitung der Evaluationsprojekte liegt bei der IEG. Bei den Evaluationen selbst werden die relevanten Akteure einbezogen. Bei der Berichterlegung wird viel Wert darauf gelegt, die Formate an die Bedürfnisse der Zielgruppen anzupassen, Wissen in «konsumfertige» Formate zu verpacken und die relevanten Informationen an die «richtigen» Zielgruppen zu vermitteln.

#### **4.4.3. Kultur und Selbstverständnis**

[41] Obwohl das Schwergewicht bei der Arbeit der IEG auf der Rechenschaftslegung liegt, wird auch das Lernen als wichtiger Evaluationszweck genannt. Oft ist es so, dass sich die Evaluationen kleinerer Gegenstände (Projekte, Programme) eher an die operativen Mitarbeitenden richten und eher formativen Charakter haben. Grössere Evaluationen (thematische oder Sektorevaluationen) sowie Evaluationen interner Strategien und Prozesse richten sich eher an die Boards; sie sollen strategische Entscheidungen ermöglichen und sind eher summativer Natur. Grundsätzlich verfolgt die IEG das Ziel, die Weltbank dabei zu unterstützen, relevanter und effektiver zu werden: Management und Boards sollen die Ergebnisse nutzen können, um bessere Entscheidungen zu treffen. Dafür ist neben der strukturellen und der funktionalen auch die «verhaltensmässige» Unabhängigkeit wichtig, die das professionelle Selbstverständnis der IEG-Mitarbeitenden auszeichnet: Hierbei geht es darum, sich an ethische Verhaltensrichtlinien zu halten und sich nicht vereinnahmen zu lassen («speak truth to power») (VAESSEN, 2018).

[42] Das Weltbank-System ist sehr gross (vergleichbar mit dem UN-System) und divers, daher ist es schwer möglich, allgemeingültige Aussagen zur Evaluationskultur über die verschiedenen Institutionen hinweg zu machen. Tendenziell hat die Evaluation gemäss der Auskunft unseres Interviewpartners in den privaten Sektoren der Bank einen eher schwierigen Stand, da die Kultur dort eine andere sei. In den öffentlichen Sektoren hat die Evaluation eine lange Tradition; die Mitarbeitenden teilten ein Grundverständnis darüber, was Evaluation ist und welche Bedeutung sie hat. Die Leitungsgremien der Bank unterstützten die Evaluation und schätzten sie.

#### 4.4.4. Vorteile und Herausforderungen

[43] Der Vorteil des hybriden Modells, in dem Evaluationen durch interne und externe Mitarbeitende durchgeführt werden, ist aus Sicht unseres Interviewpartners, dass flexibel auf Kompetenzanforderungen reagiert werden könne (häufig ist z. B. externe Methodenexpertise und internes Wissen zum Evaluationsgegenstand gefragt). Eine Herausforderung dabei sei, die richtige Balance zwischen interner und externer Expertise zu finden. Damit noch mehr neue Ideen und Innovation umgesetzt werden, wäre mehr externe Expertise wünschenswert. Eine stetige Herausforderung sei auch, den neuen Mitarbeitenden, die von einer operativen Einheit zur IEG wechseln und deren Expertise auf einem inhaltlichen Gebiet liegt, die unabhängige Haltung zu vermitteln, die sie als Evaluierende einnehmen müssen.

[44] Eine Stärke der IEG liegt in ihrer Unabhängigkeit und darin, dass ihren Evaluationen hohe Glaubwürdigkeit zuerkannt wird. Die Kehrseite der Unabhängigkeit ist aus der Perspektive des Interviewpartners, dass die IEG als zentrale Einheit relativ weit weg ist von der Realität der Projekt- oder Programmarbeit; die Erkenntnisse ihrer Evaluationen seien daher oft relativ generisch und lassen sich teilweise nicht direkt in die operative Arbeit transferieren. Eine offene Frage ist für unseren Interviewpartner, ob so viele standardmässige Selbstevaluationen notwendig sind, wie aktuell durchgeführt werden, oder ob nicht mehr Ressourcen in strategisch bedeutsamere Evaluationen fliessen sollten. Allerdings sieht er letztendlich nicht die Anzahl an Evaluationen als entscheidend an, sondern ob es gelingt, die Ergebnisse zum Lernen und für die Entscheidungsfindung zu nutzen; hier liege auch bei der Weltbank noch Optimierungspotenzial.

## 5. Diskussion

[45] In der Literatur und von allen unseren Interviewpartnern wird das Spannungsfeld zwischen den beiden Evaluationszwecken Verbesserung und Rechenschaftslegung beschrieben. Die in der Literatur genannten Lösungsansätze für dieses Spannungsfeld lassen sich wie folgt zusammenfassen (vgl. Ziff. 3.4):

1. Es wird «nur» einer der beiden Zwecke verfolgt.
  - (a) Geht es um Verbesserung, so ist ein dezentrales Modell zu bevorzugen, bei dem die Evaluationsabteilung eine beratende Rolle einnimmt und die Ownership für die Evaluation bei den Projektverantwortlichen liegt.
  - (b) Geht es um Rechenschaftslegung, so sollte die Evaluationsabteilung von den Projektverantwortlichen unabhängig sein, damit Interessenskonflikte vermieden werden.
2. Es werden beide Zwecke verfolgt und dafür zwei parallele Strukturen geschaffen (a und b).
3. Wenn beide Funktionen von derselben Evaluationsabteilung wahrgenommen werden, ist es besonders wichtig, dass ein Mandat der obersten Leitung, eine klare Evaluationsstrategie und Guidelines für die Durchführung von Evaluationen vorhanden sind.
4. Schliesslich besteht eine vierte Möglichkeit darin, die Rechenschaftspflicht von den Projekten und deren Ergebnissen auf die Evaluation und die Nutzung der Evaluationsergebnisse zu verschieben («accountability for learning»).



[46] Jede dieser Varianten zum Umgang mit dem Spannungsfeld zwischen Verbesserung und Rechenschaftslegung hat Vor- und Nachteile. Interessant ist jedoch die Frage, welcher Ansatz sich in der Praxis unter welchen Bedingungen bewähren kann.

[47] Bei Organisationen, die den Evaluationszweck der Verbesserung ins Zentrum stellen (1a), dürfte es sich vor allem um nicht öffentlich finanzierte Organisationen handeln (wie die Beispiele der Stiftung Mercator und der Bertelsmann Stiftung zeigen), bei denen die Rechenschaftspflicht in Bezug auf Wirkungen nicht von zentraler Bedeutung ist. Dieses Modell setzt einen ausgeprägten Willen der Projektverantwortlichen zu stetiger Verbesserung voraus, der mit einer entsprechenden Evaluationskultur gestützt werden kann. Eine Schwäche des Ansatzes liegt darin, dass die Evaluationsergebnisse unterschiedlicher Projekte in Bezug auf Wirkung und Wirtschaftlichkeit nur schwer vergleichbar sind, was deren Nutzung für eine übergreifende Steuerung erschwert. Damit einher geht die Gefahr, dass viel Aufwand in zahlreiche kleine Evaluationen fliesst, welche Überprüfungsbedarf melden, dafür aber die Ressourcen für strategisch bedeutsamere, übergreifende Evaluationen fehlen, wie das von der IEG als Beobachtung berichtet wird. Schliesslich könnte eine Gefahr auch darin liegen, dass Projekte zwar verbessert, aber nicht grundsätzlich in Frage gestellt werden. Dazu wäre eine gewisse Distanz zwischen Evaluations- und Projektverantwortung hilfreich, die bei diesem Modell aber nicht vorgesehen ist.

[48] Organisationen, die ausschliesslich zur Rechenschaftslegung evaluieren und das Evaluationssystem entsprechend gestalten (1b), müssen einerseits einen ausgeprägten Bedarf für die Rechenschaftslegung haben und andererseits über ein so grosses Evaluationsvolumen verfügen, dass es eine eigenständige Evaluationsabteilung rechtfertigt, wie bspw. bei der Weltbank. Die Rechenschaftslegung auf Projektebene ermöglicht sowohl die Steuerung als auch das Lernen auf einer übergeordneten, strategischen Ebene. Es ist allerdings fraglich, ob diese Variante in Reinform tatsächlich existiert. Auch im Beispiel der Weltbank gibt es offenbar wenige rein summative Evaluationen. Die Regel scheint vielmehr ein Sowohl-als-auch: Neben der Steuerung und Rechenschaftslegung sollen auch auf Projektebene Lernprozesse unterstützt und Verbesserungen angestossen werden.

[49] Der kombinierte Lösungsansatz mit parallelen Strukturen (2) lässt keine Wünsche offen. Hier bleibt einzig zu erwähnen, dass dieser wohl nur für Organisationen mit grossem Evaluationsvolumen und einem hohen Stellenwert der Evaluationstätigkeit in Frage kommt. Auf die Weltbank als Gesamtorganisation dürfte dies zutreffen, die neben der IEG auch Evaluationseinheiten innerhalb der verschiedenen Institutionen unterhält.

[50] Vermutlich am weitesten verbreitet ist der gemischte Ansatz (3), in dem beide Evaluationsfunktionen von derselben Evaluationsabteilung wahrgenommen werden. Das Beispiel des Bundesamtes für Gesundheit geht in diese Richtung. Dieser Ansatz birgt das grösste Konfliktpotenzial, weil das Spannungsfeld bestehen bleibt. Die Evaluationszwecke stehen in Konkurrenz zueinander und gefährden dadurch die Nutzung der Ergebnisse. Bei dieser Variante ist es deshalb eine zentrale Aufgabe der Direktion und der Evaluationsabteilung, Spielregeln zu definieren und auf deren Einhaltung zu achten. Ein Einbinden der Leitung sowie weiterer wichtiger interner Stakeholder in Steuer- oder Begleitgruppen kann hilfreich sein, um sowohl Schwachstellen zu untersuchen wie auch steuerungsrelevante Informationen zu generieren. Das setzt jedoch voraus, dass diese Personen auch Zeit dafür aufbringen können und wollen – was wiederum einen hohen Stellenwert der Evaluation und eine ausgeprägte Evaluationskultur voraussetzt.

[51] Verlockend ist schliesslich der Ansatz der «accountability for learning», der das Spannungsfeld zwischen summativer und formativer Herangehensweise ganz aufzulösen scheint (keine der

hier beschriebenen Organisationen verfolgt diesen Ansatz). Abgesehen von den Nachteilen eines rein formativen Ansatzes (analog 1a) besteht eine zentrale Herausforderung dieses Modells in der Qualitätssicherung der Evaluation und der Form der Berichterstattung, die nicht nur die Evaluationsergebnisse, sondern auch die Verbesserungsmassnahmen und deren Umsetzung umfassen muss. Das Aufsichtsorgan, dem die Rechenschaftslegung gilt, muss sich somit auch für das «Wie» der Massnahmenumsetzung interessieren, was eine gewisse Nähe zur operativen Ebene erfordert. Gleichzeitig setzt der Ansatz voraus, dass das Aufsichtsorgan auf die Steuerung der Massnahmen aufgrund der Evaluationsergebnisse verzichtet. Seine Rolle beschränkt sich auf die Überwachung der Lernprozesse. Werden die Ergebnisse dennoch als Entscheidungsgrundlage auf Massnahmenebene genutzt, bspw. für die Beendigung von Massnahmen, so werden die Lernprozesse auch mit diesem Modell «bedroht».

[52] Diese Arbeit hat einige Limiten: So wäre es wünschenswert gewesen, mehr Literaturdatenbanken zu berücksichtigen, eine grössere Zahl an Organisationen zu untersuchen und mehrere Personen pro Organisation um Auskunft zu bitten. Auch die Auswahl der Organisationen erfolgte gestützt auf pragmatische und nicht theoretische Überlegungen. Nichtsdestotrotz bieten die Beispiele einen aufschlussreichen Einblick in die Evaluationspraxis, in der, wie zu erwarten war, das Spannungsverhältnis zwischen summativer und formativer Evaluation nicht immer und überall gelöst ist. Auch bei den in der Literatur genannten Lösungsansätzen bleibt – mit Ausnahme des zweiten Lösungsansatzes, der zwei parallele Strukturen vorsieht – das Spannungsfeld stets bestehen. Der Praxis bleibt einstweilen nichts anderes übrig, als dies anzuerkennen und sorgfältig flankierende Massnahmen zu entwickeln und umzusetzen, die das Spannungsfeld abmildern. Die Forschung über Evaluation sollte sich vermehrt der Frage widmen, wie und unter welchen Bedingungen dies gelingen kann.

---

DR. VERENA FRIEDRICH ist wissenschaftliche Mitarbeiterin beim Tabakpräventionsfonds.

DR. KATHRIN FREY ist Partnerin bei KEK – CDC, Zürich.

LISA GUGGENBÜHL leitet den Bereich Wirkungsmanagement bei Gesundheitsförderung Schweiz.

Wir danken unseren Interviewpartnern sehr herzlich dafür, dass sie sich die Zeit genommen und uns einen Einblick in ihre Organisationen gewährt haben.

---

## Literatur

BALTHASAR, ANDREAS (2008): Der Einfluss unterschiedlicher Institutionalisierungsmodelle auf die Verwendung von Evaluationen: Erfahrungen aus der Schweiz. *Zeitschrift für Evaluation*, 7(1), 243–271.

Bundesamt für Gesundheit (2020): Evaluation im BAG. Retrieved from [www.bag.admin.ch/evaluation](http://www.bag.admin.ch/evaluation).

DE LAAT, BASTIAAN (2014): Evaluator, Evaluand, Evaluation Commissioner: A Tricky Triangle. In M. Läubli Loud & J. Mayne (Eds.), *Enhancing Evaluation Use. Insights from Internal Evaluation Units* (pp. 15–36). Thousand Oaks: Sage.

Eval-Wiki (2016): [Eval-wiki.org](http://eval-wiki.org). Retrieved from [https://eval-wiki.org/w\\_glossar/index.php?title=Eval-wiki.org&oldid=2516](https://eval-wiki.org/w_glossar/index.php?title=Eval-wiki.org&oldid=2516).

FETTERMAN, DAVID M./RODRÍGUEZ-CAMPOS, LILIANA/ZUKOSKI, ANN P. (2017): *Collaborative, participatory, and empowerment evaluation: Stakeholder involvement approaches*. New York: Guilford Publications.

- GÖRGENS, MARELIZE/KUSEK, JODY (2009): Making Monitoring and Evaluation Systems Work. A Capacity Development Toolkit. Washington, D.C. The World Bank.
- HAWKINS, PENNY (2014): Evaluation Management: Lessons From New Zealand and International Development Evaluation. In M. Läubli Loud & J. Mayne (Eds.), *Enhancing Evaluation Use. Insights from Internal Evaluation Units* (pp. 37–54). Thousand Oaks: Sage.
- HØJLUND, STEVEN (2014): Evaluation use in evaluation systems – the case of the European Commission. *Evaluation*, 20(4), 428–446.
- IBRD (2019): *World Bank Group Evaluation Principles*. Washington, D.C.
- JOHNSON, KELLI/GREENSEID, LIJA O./TOAL, STACIE A./KING, JEAN. A./LAWRENZ, FRANCES P./VOLKOV, BORIS (2009): Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, 30(3), 377–410.
- KING, JEAN A. (2007): Making sense of participatory evaluation. *New Directions for Evaluation* (Issue 114), 83–86.
- LEEUEW, FRANS/FURUBO, JAN-ERIC (2008): Evaluation Systems. What Are They and Why Study Them? *Evaluation*, 14(2), 157–169.
- LINDGREN, LENA/HANBERGER, ANDERS/LUNDSTRÖM, ULF (2016): Evaluation systems in a crowded policy space: Implications for local school governance. *Education Inquiry*, 7(3).
- MAYNE, JOHN (2012): Independence in Evaluation and the Role of Culture. In J.C. Barbier & P. Hawkins (Eds.), *Evaluation cultures: Sense-making in Complex Times*. (pp. 105–138). New Brunswick, NJ: Transaction.
- MORRA IMAS, LINDA G./RIST, RAY C. (2009): *The Road to Results. Designing and Conducting Effective Development Evaluations*. Washington, D.C.: The World Bank.
- OECD, Organisation for Economic Co-operation and Development (1991): *Principles for Evaluation of Development Assistance*. Paris: OECD.
- PATTON, MICHAEL Q. (1997): *Utilization-focused evaluation. The new century text* (3rd ed.). Thousand Oaks: Sage.
- Project Management Office (ohne Jahr): *Governance of Evaluation, Proof of Concept and Pilot Projects: Government of Newfoundland and Labrador*.
- PROM-JACKSON, SUKAI/BARTSIOTAS, GEORGE A. (2014): *Analysis of the Evaluation Function in the United Nations System*. Geneva: United Nations.
- SEVAL (2014): *Kompetenzrahmen für Evaluationsmanagerinnen und Evaluationsmanager: Aufgaben, Herausforderungen, Kompetenzen*. Bern: SEVAL.
- STUFFLEBEAM, DANIEL L. (1997): *Strategies for Institutionalizing Evaluation: Revisited*. Retrieved from Kalamazoo.
- VAESSEN, JOS (2018): Five ways to think about quality in evaluation. Retrieved from <https://ieg.worldbankgroup.org/blog/five-ways-think-about-quality-evaluation>.
- WIDMER, THOMAS/FREY, KATHRIN (2006): Evaluation von Mehrebenen-Netzwerkstrategien. *Zeitschrift für Evaluation*, 2/2006, 287–316.
- WIDMER, THOMAS/FREY, KATHRIN/RÜEGG, NADJA/STADTER, CORNELIA/BUNDI, PIRMIN/STEIN, JEFFREY (2015): *Qualität der IFES-Schulevaluationen und deren Nutzung im Kanton Zürich. Schlussbericht*. Retrieved from Zürich:
- WIDMER, THOMAS/NEUENSCHWANDER, PETER (2004): Embedding Evaluation in the Swiss Federal Administration. Purpose, Institutional Design and Utilization. *Evaluation*, 10(4), 388–409.